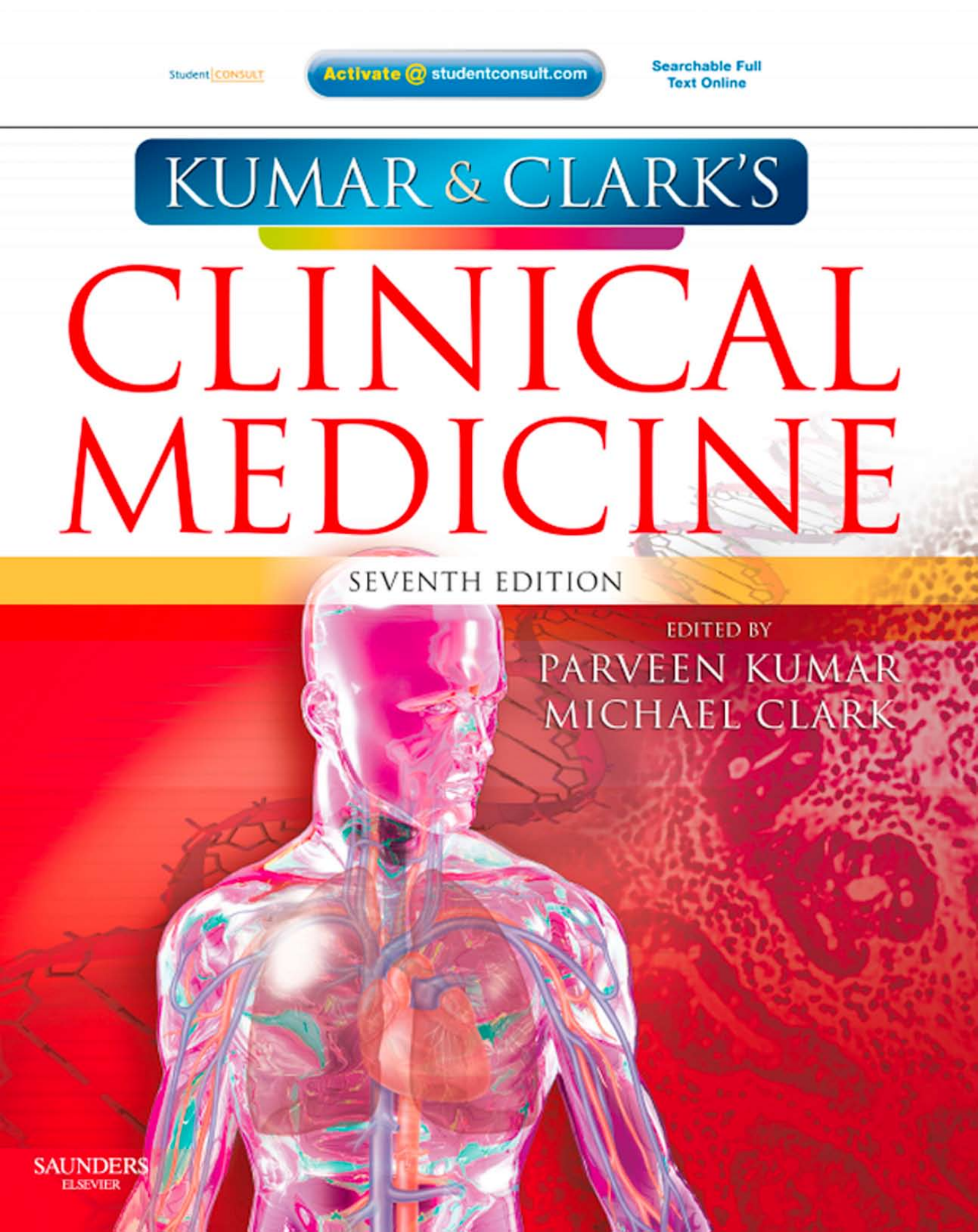
Student CONSULT
Activate @ studentconsult.com
Searchable Full Text Online

KUMAR & CLARK'S CLINICAL MEDICINE

SEVENTH EDITION

EDITED BY
PARVEEN KUMAR
MICHAEL CLARK



SAUNDERS
ELSEVIER

KUMAR & CLARK'S

CLINICAL MEDICINE

SEVENTH EDITION

For Elsevier
Commissioning Editor: **Pauline Graham**
Development Editors: **Janice Urquhart, Ruth Swan**
Project Manager: **Frances Affleck**
Designer: **Stewart Larking, Sarah Russell**
Illustration Manager: **Gillian Richards**
Illustrators: **Richard Morris, Ethan Danielson**

KUMAR & CLARK'S

CLINICAL MEDICINE

SEVENTH EDITION

EDITED BY

PROFESSOR PARVEEN KUMAR

CBE BSc MD DM(HC) FRCP FRCP (EDIN)

Professor of Medicine and Education,
Barts and The London School of Medicine and Dentistry,
Queen Mary University of London,
and Honorary Consultant Physician and Gastroenterologist,
Barts and the London NHS Trust and Homerton University
Hospital NHS Foundation Trust, London, UK

DR MICHAEL CLARK

MD FRCP

Honorary Senior Lecturer,
Barts and The London
School of Medicine and Dentistry,
Queen Mary University of London, London, UK

SAUNDERS

ELSEVIER

Edinburgh London New York Oxford Philadelphia St Louis Sydney Toronto 2009

SAUNDERS
ELSEVIER

© 2009, Elsevier Limited. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permissions may be sought directly from Elsevier's Rights Department: phone: (+1) 215 239 3804 (US) or (+44) 1865 843830 (UK); fax: (+44) 1865 853333; e-mail: healthpermissions@elsevier.com. You may also complete your request online via the Elsevier website at http://www.elsevier.com/permissions.

Sixth Edition 2005
Fifth Edition 2002
Fourth Edition 1998
Third Edition 1994
Second Edition 1990
First Edition 1987

ISBN 9780702029936
International ISBN 9780702029929

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
A catalog record for this book is available from the Library of Congress

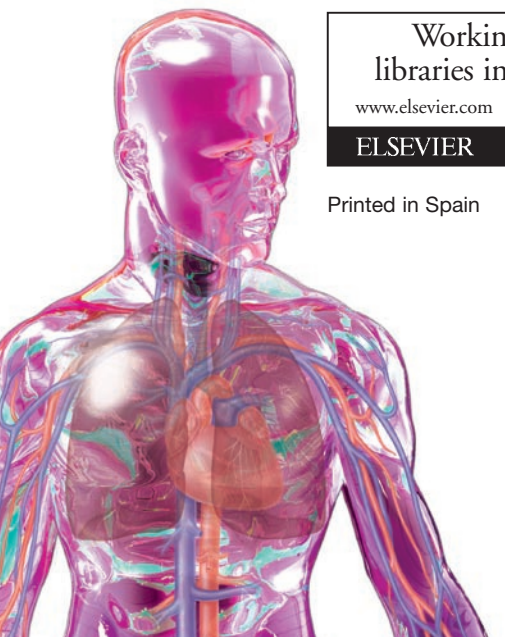
Notice
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our knowledge, changes in practice, treatment and drug therapy may become necessary or appropriate. Readers are advised to check the most current information provided (i) on procedures featured or (ii) by the manufacturer of each product to be administered, to verify the recommended dose or formula, the method and duration of administration, and contraindications. It is the responsibility of the practitioner, relying on their own experience and knowledge of the patient, to make diagnoses, to determine dosages and the best treatment for each individual patient, and to take all appropriate safety precautions. To the fullest extent of the law, neither the Publisher nor the Editors assume any liability for any injury and/or damage to persons or property arising out of or related to any use of the material contained in this book.

The Publisher

ELSEVIER your source for books, journals and multimedia in the health sciences
www.elsevierhealth.com

Working together to grow libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

Printed in Spain



The publisher's policy is to use **paper manufactured from sustainable forests**

Contents

Online contents vii
Contributors ix
International Advisory Panel xi
Preface to the Seventh Edition xiii
Acknowledgements xv

1. **Ethics** 1
L Doyal
Communication 8
A Cushing and C Mallinson

2. **Molecular cell biology and genetic disorders** 19
RK Iles

3. **The immune system and disease** 53
M Peakman

4. **Infectious diseases, tropical medicine and sexually transmitted infection** 79
RG Finch, WL Irving, P Moss and J Anderson

5. **Nutrition** 207
M Elia

6. **Gastrointestinal disease** 241
PD Fairclough and DBA Silk

7. **Liver, biliary tract and pancreatic disease** 319
AK Burroughs and D Westaby

8. **Haematological disease** 387
MF Murphy, J Wainscoat and KJ Pasi

9. **Malignant disease** 449
CJ Gallagher, TA Lister, MJ Johnson and A Davies

10. **Rheumatology and bone disease** 499
M Shipley, A Rahman, D O'Gradaigh and JE Compston

11. **Renal disease** 571
MM Yaqoob

12. Water, electrolytes and acid–base balance **649**
MM Yaqoob

13. Cardiovascular disease **681**
AJ Camm and N Bunce

14. Respiratory disease **811**
AJ Frew and ST Holgate

15. Intensive care medicine **889**
CJ Hinds and D Watson

16. Drug therapy and poisoning **923**
M Rawlins and JA Vale

17. Environmental medicine **953**
ML Clark, CRA Clarke and A Geddes

18. Endocrine disease **963**
TA Howlett and MJ Levy

19. Diabetes mellitus and other disorders of metabolism **1029**
EAM Gale and JV Anderson

20. The special senses **1077**
M Harries, F Vaz and S Verma

21. Neurological disease **1095**
CRA Clarke

22. Psychological medicine **1185**
PD White and AW Clare

23. Skin disease **1225**
DG Paige

Index 1273

Online Contents

Student CONSULT

www.studentconsult.com

The seventh edition of *Kumar & Clark's Clinical Medicine* has an online version that can be accessed via www.studentconsult.com. The online version consists of the full text and illustrations for the seventh edition as well as other additional features, some of which are unique to Kumar & Clark. To access the online version, readers should follow the instructions on the book's inside front cover. The list below provides a guide to some of the online features, and you can also find them listed online in the 'extras' folder.

ONLINE APPENDIX

edited by Professor Janaka de Silva

This appendix, which will be updated regularly, provides supplementary information on diseases and problems that are more prevalent in certain parts of the world, up to date for the seventh edition. The need for this material – which enlarges on what can be found in the print version – became plain to the Editors from their travels to various parts of the world. We are grateful to members of the Kumar & Clark International Advisory Panel and other colleagues for their contributions to this unique resource.

- Brucellosis *Professor S S Fedail*
- Dengue fever *Dr R Sood*
- Diarrhoeas including amoebiasis *Professor S A Azer*
- Human avian influenza infection *Professor M Ellis*
- HIV in resource-limited settings *Dr M Mendelson and Professor G Maartens*
- Infections caused by Rickettsiae, Orientiae and Coxiella *Dr R Premaratna*
- Leprosy *Dr S A Kamath*
- Malaria *Dr S A Kamath*
- Rabies *Professor K N Viswanathan*
- Rheumatic fever *Professor K N Viswanathan*
- Rift Valley fever *Professor O. Khalafallah, Dr A A K Jebriel, Dr M El Sadig, Dr M Mirghani, Professor S S Fedail*
- Severe acute respiratory syndrome (SARS) *Professor K N Lai*
- Soil-transmitted helminths *Professor N de Silva*
- Vaccination for adults *Dr R Dewan*
- Fluorosis *Professor F S Hough*
- Tropical sprue *Professor S A Azer*
- Liver transplantation *Dr J T Wells, Professor M R Lucey*
- Thrombotic thrombocytopenic purpura (TTP) associated with HIV *Professor V J Louw*
- Vitamin B_{12} and folic acid deficiency *Professor V J Louw*
- IgA nephropathy *Professor K N Lai*
- Arsenic poisoning *Dr A K Kundu*
- Drug-drug and drug-food interactions *Professor S G Carruthers*
- Poisoning – a Sri Lankan perspective *Professor J de Silva*
- Snake bite *Professor Dr CA Gnanathasan and R Sheriff*
- Pyogenic meningitis *Dr S Shafqat and Dr A Zaidi*
- Tropical neurology *Dr U K Ranawaka*

INTERACTIVE SURFACE ANATOMY

from Drake et al: Gray's Anatomy for Students

This programme relates anatomy to the skills of clinical examination. Over 40 interactive animations take the student through the seven body regions, demonstrating major anatomical features, surface landmarks and other palpable structures. By working through each module students see their role when performing a general physical examination, listening for heart sounds, examining movements of the eye, and many others. We are grateful to the authors of *Gray's Anatomy for Students* for the opportunity to provide Kumar & Clark readers with the benefits of this unique interactive learning tool and to introduce them to this outstanding anatomy textbook.

Back

- How to identify specific vertebral spinous processes
- Visualizing the inferior ends of the spinal cord and subarachnoid space

Thorax

- How to count ribs
- Surface anatomy of the breast in women
- Visualizing structures at the T4/5 vertebral level
- Visualizing structures in the superior mediastinum
- Visualizing the margins of the heart
- Where to listen for heart sounds
- How to visualize the pleural cavities and lungs
- Where to listen to right lung sounds
- Where to listen to left lung sounds

Pelvis and perineum

- Identification of structures in the urogenital triangle of women
- Identification of structures in the urogenital triangle of men

Abdomen
- How to find the superficial inguinal ring
- How to determine lumbar vertebral levels
- Visualizing structures at the LI vertebral level
- Visualizing the position of major blood vessels
- Using abdominal quadrants to locate major viscera
- Defining surface regions to which pain from the gut is referred
- Where to find the kidneys and spleen

Lower limb
- Avoiding the sciatic nerve
- Finding the femoral artery in the femoral triangle
- Identifying structures around the knee
- Visualizing the contents of the popliteal fossa
- Finding the tarsal tunnel
- Identifying tendons around the ankle and in the foot
- Finding the dorsalis pedis artery
- Identifying major superficial veins
- Where to take peripheral arterial pulses in the lower limb

Upper limb
- Visualizing the axilla, its contents and related structures
- Locating the brachial artery in the arm
- Locating the triceps brachii tendon and the position of the radial nerve
- Identifying the cubital fossa and its contents
- Identifying tendons and locating major vessels and nerves in the distal forearm
- Identifying the position of the flexor retinaculum and the recurrent branch of the median nerve
- Visualizing the positions of the superficial and deep palmar arches
- Where to take peripheral arterial pulses in the upper limb

Head and neck
- Visualizing structures at the CIII/CIV and CVI vertebral levels
- How to outline the anterior and posterior triangles of the neck
- How to locate the cricothyroid ligament (membrane)
- How to find the thyroid gland
- Where to take peripheral arterial pulses in the head and neck
- Testing eye movements – demonstration illustrating the muscle involved in each activity.

ANIMATIONS OF PRACTICAL PROCEDURES
This feature brings to life a selection of key practical procedures in a way not usually possible outside the clinic or skills laboratory. Once again, the existence of the online version has made it possible to provide this unique feature, where the animations build on the descriptions given in the 'practical boxes' in Kumar & Clark.

- Arterial cannulation
- Arterial puncture
- Bladder catheterization: female
- Bladder catheterization: male
- Central venous catheterization (CVC): jugular vein
- Central venous catheterization (CVC): subclavian vein
- Joint aspiration
- Lumbar puncture
- Venepuncture

HEART AND LUNG SOUNDS
by Dr Salvatore Mangione

This feature provides readers with the opportunity to improve their clinical skills by listening to 10 pulmonary and 12 cardiac sounds. The sounds can be heard on their own or accompanied by an instructive narrative. This workshop has been prepared by Dr Salvatore Mangione, and we are grateful for the opportunity to make it available to readers of Kumar & Clark.

- CARDIAC AUSCULTATION:
  Opening snap
  Aortic regurgitation and systolic flow murmur
  Pericardial friction rub
  S3 gallop
  Mid-systolic click
  Patent ductus arteriosus
  Early-systolic ejection sound
  Aortic stenosis
  S4 gallop
  Mitral regurgitation
  Aortic regurgitation
  Mitral stenosis and opening snap
- PULMONARY AUSCULTATION:
  Bronchial breath sounds
  Crackles
  Wheezes
  Vesicular breath sounds
  Rhonchi
  Pleural friction rub
  Bronchial breath sounds and late-inspiratory crackles
  Late-inspiratory squeak
  Amphoric breath sounds
  Whispered pectoriloquy

ASK THE AUTHORS
This popular feature continues on into the seventh edition of Kumar & Clark. It gives readers the opportunity to submit questions to the editors and to read the answers online. Readers can view the most recent questions, or can access the complete list of questions and answers that have been logged on the site.

Contributors

Jane Anderson BSc PhD MB BS FRCP
Consultant Physician, The Centre for the Study of Sexual Health and HIV, Homerton University Hospital NHS Foundation Trust; Professor, The Institute of Cell and Molecular Science, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK
Infectious diseases, tropical medicine and sexually transmitted infection

John V Anderson MD MA MBBS FRCP
Consultant Physician, Homerton University Hospital NHS Foundation Trust and Barts and The London NHS Trust; Honorary Senior Lecturer at Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK
Diabetes mellitus and other disorders of metabolism

Nicholas Harry Bunce BSc MBBS MD
Consultant Cardiologist, St George's Healthcare NHS Trust, London, UK
Cardiovascular disease

Andrew Kenneth Burroughs MBChB (Hons) FRCP
Consultant Physician and Hepatologist, Professor of Hepatology, The Royal Free Sheila Sherlock Liver Centre, and Department of Surgery, Royal Free Hospital, London, UK
Liver, biliary tract and pancreatic disease

A John Camm MD FRCP
Professor of Clinical Cardiology, St George's, University of London, London, UK
Cardiovascular disease

Anthony W Clare MD FRCP FRCPI FRCPsych MPhil (deceased)
Formerly Associate Professor of Psychiatry, Dublin University; Consultant Psychiatrist, St Patrick's Hospital, Dublin, Ireland
Psychological medicine

Michael L Clark MD FRCP
Honorary Senior Lecturer, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK
Environmental medicine

Charles Richard Astley Clarke MA MB FRCP
Honorary Consultant Neurologist, National Hospital for Neurology and Neurosurgery, London, UK
Environmental medicine
Neurological disease

Juliet Compston MD FRCPath FRCP FMedSci
Professor of Bone Medicine, Department of Medicine, University of Cambridge, School of Clinical Medicine, Cambridge; Honorary Consultant Physician, Addenbrookes' NHS Trust, Cambridge, UK
Bone disease

Annie Cushing PhD FDSRCS (Eng) BDS (Hons)
Reader in Clinical Communication Skills, Institute of Health Sciences Education, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK
Communication

Andrew Davies MRCP PhD
Senior Lecturer in Medical Oncology and Honorary Consultant, Cancer Services Division, Southampton General Hospital, Southampton, UK
Malignant disease

Len Doyal BA MSc
Emeritus Professor of Medical Ethics, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London; Professor of Healthcare Ethics, University of East Anglia, Norwich, UK
Ethics

Marinos Elia BSc (Hons) MD FRCP
Professor of Clinical Nutrition and Metabolism, University of Southampton, Southampton, UK
Nutrition

Peter D Fairclough MD FRCP
Honorary Consultant Gastroenterologist, Barts and The London NHS Trust, London, UK
Gastrointestinal disease

Roger Finch FRCP FRCP (Edin) FRCPath FFPM
Professor of Infectious Diseases, Division of Infectious Diseases, School of Molecular Medical Sciences, Faculty of Medicine and Health Sciences, The University of Nottingham, Nottingham, UK
Infectious diseases, tropical medicine and sexually transmitted infection

Anthony J Frew MA MD FRCP
Professor of Allergy and Respiratory Medicine, Brighton and Sussex Medical School, Brighton, UK
Respiratory disease

Edwin AM Gale MB FRCP
Professor of Diabetic Medicine, University of Bristol, Bristol, UK
Diabetes mellitus and other disorders of metabolism

Christopher J Gallagher MBChB PhD FRCP
Consultant Medical Oncologist, St Bartholomew's Hospital, Barts and The London NHS Trust, London, UK
Malignant disease

Alasdair MacIntosh Geddes CBE MBChB FRCP (Lond & Edin) FRCPath FFPH FMedSci
Emeritus Professor of Infectious Diseases, School of Medicine, University of Birmingham, Birmingham, UK
Environmental medicine

Meredydd Harries FRCS MSc
Consultant ENT Surgeon, Royal Sussex County Hospital, Brighton, UK
The special senses

Charles Hinds FRCP FRCA
Professor of Intensive Care Medicine, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, and Barts and The London NHS Trust, London, UK
Intensive care medicine

Stephen T Holgate BSc MB BS MD DSc FRCP FRCP (Edin) FRC Path FIBiol FMedSci
MRC Clinical Professor of Immunopharmacology, School of Medicine, University of Southampton, Southampton, UK
Respiratory disease

Trevor A Howlett MD FRCP
Consultant Physician and Endocrinologist, Leicester Royal Infirmary, University Hospitals of Leicester NHS Trust, Leicester, UK
Endocrine disease

Raymond Kruse Iles BSc MSc PhD CBiol MIBiol FRSC
Associate Dean, Research, and Professor of Biomedical Sciences, School of Health and Social Sciences, Middlesex University, Hendon, UK
Molecular cell biology and genetic disorders

William L Irving MA PhD MB BChir MRCP FRCPath
Professor of Virology, The University of Nottingham; Honorary Consultant Virologist, Nottingham University Hospitals NHS Trust, Nottingham, UK
Infectious diseases, tropical medicine and sexually transmitted infection

Miriam J Johnson MD FRCP MRCGP MB ChB (Hons)
Senior Lecturer in Palliative Medicine, Hull York Medical School; Honorary Consultant to St Catherine's Hospice and the Acute Scarborough and North East Yorkshire Health Care Trust, North Yorkshire, UK
Palliative care

Miles J Levy MD FRCP
Consultant Physician and Endocrinologist, Leicester Royal Infirmary, University Hospitals of Leicester NHS Trust, Leicester, UK
Endocrine disease

T Andrew Lister MD FRCP FRCPath FRCR FMedSc
Professor of Medical Oncology, St Bartholomew's Hospital, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK
Malignant disease

Christopher Niels Mallinson BA MB BCh FRCP
Emeritus Consultant Physician, University Hospital Lewisham, The Lewisham Hospital NHS Trust, London; Founder Board Member of European Association for Communication in Health Care (EACH); Former President of Communication in Healthcare Forum of the Royal Society of Medicine, London, UK
Communication

Peter J Moss MB ChB MD FRCP DTMH
Consultant in Infectious Diseases, Department of Infection and Tropical Medicine, Hull and East Yorkshire Hospitals NHS Trust; Honorary Senior Lecturer, Hull York Medical School, UK
Infectious diseases, tropical medicine and sexually transmitted infection

Michael F Murphy MD FRCP FRCPath
Professor of Blood Transfusion Medicine, University of Oxford; Consultant Haematologist, NHS Blood & Transplant and Department of Haematology, Oxford Radcliffe Hospitals NHS Trust, Oxford, UK
Haematological disease

Donncha O'Gradaigh MB PhD FFSEM MRCPI
Consultant Rheumatologist, Department of Rheumatology, Waterford Regional Hospital, Ireland
Bone disease

David Paige MA MBBS FRCP
Consultant Dermatologist, Barts the London NHS Trust; Honorary Senior Lecturer, Queen Mary University of London; Barts and the London School of Medicine and Dentistry, London, UK
Skin disease

K John Pasi PhD MBChB FRCP FRCPath FRCPCH
Professor of Haemostasis and Thrombosis, Centre for Haematology, Institute of Cell and Molecular Science, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK
Haematological disease

Mark Peakman MB BS BSc MSc PhD FRCPath
Professor of Clinical Immunology, King's College London School of Medicine; Honorary Consultant in Immunology, King's College Hospital, London, UK
The immune system and diseases

Anisur Rahman MA PhD BM BCh FRCP
Professor of Rheumatology, University College London; Honorary Consultant Rheumatologist, University College London Hospital, London, UK
Rheumatology and bone disease

Sir Michael Rawlins MD FRCP FMedSci
Chairman, National Institute for Health and Clinical Excellence; Honorary Professor, London School of Hygiene and Tropical Medicine, London, UK
Drug therapy and poisoning

Michael Shipley MA MD FRCP
Consultant Rheumatologist, University College Hospital and King Edward VII's Hospital, London, UK
Rheumatology

DBA Silk MD AGAF FRCP
Professor of Clinical Nutrition, Department of Biosurgery and Surgical Technology, Division of SORA, Imperial College London, London, UK
Gastrointestinal disease

Allister Vale MD FRCP FRCPE FRCPG FFOM FAACT FBTS Hon FRCPSG
Director, National Poisons Information Service (Birmingham Unit) and West Midlands Poisons Unit, City Hospital, Birmingham; Honorary Professor of Toxicology, School of Biosciences and College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK
Drug therapy and poisoning

Francis Vaz MBBS BSc (Hons) FRCS (ORL-HNS)
Consultant ENT/Head and Neck Surgeon, Department of ENT/Head and Neck Surgery, University College London Hospital, London, UK
The special senses

Seema Verma MBBS MD FRCOphth
Consultant Ophthalmic Surgeon and Director of A&E, Moorfields Eye Hospital, London, UK
The special senses

James Stephen Wainscoat FRCP FRCPath
Consultant Haematologist, Department of Haematology, Oxford Radcliffe Hospital, Oxford Radcliffe Hospitals NHS Trust, Oxford, UK
Haematological disease

David Watson BSc (Hons) FRCA
Consultant and Senior Lecturer in Intensive Care Medicine, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, and Homerton University Hospital NHS Trust, London, UK
Intensive care medicine

David Westaby MA Cantab FRCP
Consultant Physician and Gastroenterologist, Director of Endoscopic Service, Lead for Hepatobiliary Medicine at the Imperial College Healthcare NHS Trust, Hammersmith Hospital, London, UK
Liver, biliary tract and pancreatic disease

Peter D White MD FRCP FRCPsych
Professor of Psychological Medicine, Centre for Psychiatry, Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK
Psychological medicine

Muhammad Magdi Yaqoob MD FRCP
Professor, William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London; Consultant, Department of Renal Medicine and Transplantation, Barts and the London NHS Trust, London, UK
Renal disease
Water, electrolytes and acid–base balance

International Advisory Panel

AUSTRALIA

Associate Professor Peter Katelaris
Senior Consultant, Gastroenterology Department, Concord Hospital, University of Sydney, Sydney

Professor N Thomson
Professor and Head, Department of Medicine and Central and Eastern Clinical School, Monash University Medical School, The Alfred Hospital, Melbourne, Victoria

BANGLADESH

Professor Salimur Rahman
Professor of Hepatology, Department of Hepatology, Faculty of Medicine, Bangabandhu Sheikh Mujib Medical University, Bangladesh

CHINA

Professor Kar Neng Lai
Yu Chiu Kwong Chair of Internal Medicine and Chief of Medicine, Department of Medicine, University of Hong Kong, Queen Mary Hospital, Hong Kong

Professor Debing Wang
Institute of Hematology, Peking University People's Hospital, Beijing

EGYPT

Professor Rashad Barsoum
Emeritus Professor of Medicine, Department of Internal Medicine, Cairo University, Cairo

GREECE

Professor Athanasios G Papavassiliou
Professor and Head, Department of Biological Chemistry, University of Athens Medical School, Athens

INDIA

Professor G Abraham
Professor, Department of Medicine, Sri Ramachandra Medical College and Research Institute, Chennai

Professor Debabrata Bandyopadhyay
Professor and Head, Department of Dermatology, STD and Leprosy, R G Kar Medical College, Kolkata

Professor Alaka Deshpande
Professor of Medicine and Head, Department of Internal Medicine, Grant Medical College and Sir J J Group of Govt. Hospitals, Mumbai

Dr Richa Dewan
Professor and Head, Department of Medicine, Maulana Azad Medical College, New Delhi

Dr R Gupta
Associate Professor, Department of Medicine, All India Institute of Medical Sciences, New Delhi

Professor Dr Sandhya Kamath
Professor and Head, Department of Medicine, Topiwala National Medical College and B Y L Nair Charitable Hospital, Mumbai

Dr Arup K Kundu
Associate Professor of Medicine and In-Charge, Division of Rheumatology, Department of Medicine, R G Kar Medical College, Kolkata

Professor (Dr) B D Mankad
Professor and Head, Department of Medicine, BJ Medical College, Ahmedabad

Dr Dilip Mathai
Professor and Head, Department of Medicine, Christian Medical College, Vellore, Tamil Nadu

Dr V K Rajamani
Professor, Internal Medicine, Madras Medical College, Chennai

Dr S M Rajendran
Professor and Head, Department of Medicine, Rajah Muthiah Medical College Hospital, Annamalai University, Chidambaram

Dr S K Sharma
Director, Professor and Head, Department of Medicine, Lady Hardinge Medical College and SSK Hospital, New Delhi

Dr S Shivakumar
Professor and Head, Department of Medicine, Stanley Medical College and Hospital, Chennai

Dr Rita Sood
Professor, Department of Medicine, All India Institute of Medical Sciences, New Delhi

Dr K N Viswanathan
Professor and Senior Consultant, Internal and Tropical Medicine, Department of Medicine, Sri Manakula Vinayagar Medical College and Hospital, Pondicherry

JORDAN

Professor K Ajlouni
Professor and President, Department of Internal Medicine, National Center for Diabetes, Endocrinotogy and Genetics, Jordan University Hospital, Amman

Professor Abdullah Saadeh
Consultant Physician and Cardiologist, Department of Internal Medicine, King Abdullah University Hospital/Jordan University of Science and Technology, Irbid

KUWAIT

Professor Nabila Abdella
Professor and Chairperson, Department of Medicine, Faculty of Medicine, Kuwait University, Kuwait City

MALAYSIA

Professor Samy A Azer
Professor of Medical Education and Chair of Medical Education Research and Development Unit, Faculty of Medicine, Universiti Teknologi Mara, Shah Alam

MALTA

Professor JM Cacciottolo
Professor and Head, Department of Medicine, University of Malta Medical School

OMAN

Professor N Woodhouse
Professor and Head, Department of Endocrinology, Department of Medicine, Sultan Qaboos University, Muscat

PAKISTAN

Professor Dr Muhammad Akbar Chaudhry
Principal, Professor and Head, Department of Medicine, Continental Medical College, Lahore

Dr Saeed S Hamid
The Karim Jiwa Professor of Medicine (Gastroenterology), Associate Dean and Medical Director, Aga Khan University Hospital, Karachi

Professor M Ata Khan
Professor, Department of Medicine, The Aga Khan University Hospital, Karachi

Professor Masood Hameed Khan
Vice Chancellor (Professor of Medicine), Department of Medicine, Dow University of Health Sciences, Karachi

Dr S Shafqat
Associate Professor, Department of Neurology, Aga Khan University Medical College, Karachi

SAUDI ARABIA

Dr M M Al-Nozha
Professor and Consultant Cardiologist, Department of Cardiology, King Fahad Hospital (Teaching), Madinah Munawwarah

SOUTH AFRICA

Professor F S Hough
Professor, Division of Endocrinology and Metabolism, Department of Medicine, Tygerberg Academic Hospital and University of Stellenbosch, Cape Town

Professor Vernon J Louw
Associate Professor and Head, Division of Clinical Haematology, Department of Internal Medicine, University of the Free State, Bloemfontein

Professor Janet L Seggie
Professor, Department of Medicine, University of Cape Town and Groote Schuur Hospital, Cape Town

Professor Yosuf Veriaya
Professor and Head, Department of Internal Medicine, Faculty of Health Sciences, University of Witwatersrand, Johannesburg

SRI LANKA

Professor H Janaka de Silva
Professor of Medicine, Department of Medicine, Faculty of Medicine, University of Kelaniya, Ragama

Professor Nilanthi de Silva
Professor, Department of Parasitology, Faculty of Medicine, University of Kelaniya, Ragama

Professor Rezvi Sheriff
Senior Professor of Medicine and Director Postgraduate Medicine, Department of Clinical Medicine, Faculty of Medicine, University of Colombo, Colombo

SUDAN

Professor Suliman S Fedail
Professor of Medicine and Gastroenterologist, Chairman of Fedail Medical Centre, Department of Gastroenterology, University of Khartoum, Khartoum

UNITED ARAB EMIRATES

Professor S George Carruthers
Professor of Internal Medicine and Dean, Faculty of Medicine & Health Sciences, United Arab Emirates University, Al Ain

Professor Michael Ellis
Professor and Chair, Department of Medicine, Faculty of Medicine and Health Sciences, United Arab Emirates University, Al Ain

Dr J M Muscat-Baron
Professor of Medicine, Clinical Dean, Department of Medicine, Dubai Medical College, Dubai

United States of America

Dr Michael R Lucey
Professor of Medicine and Chief, Section of Gastroenterology and Hepatology, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin

Preface

It is almost 25 years since we first had the idea of writing and editing this book. Now in its 7th edition, *Kumar & Clark's Clinical Medicine* continues to play a primary role in the education of medical students and doctors worldwide, enabling us all to provide better medical care to our patients. We are committed to ensuring that the book continues to change and be updated constantly in line with evolving methods of learning and, particularly, the increasing uses of the internet and advent of new technologies.

When we sat down to think about the 7th edition we were determined the book would not get bigger in size. Thanks to an innovative new design, we have managed to reduce the number of pages whilst accommodating a huge amount of new material. This includes increased coverage and emphasis on new technologies, and more material on areas such as rehabilitation medicine. We have brought in new authors to many chapters, and this edition has once again been thoroughly revised and updated, clearly and colourfully illustrated, and attractively presented. We hope its clarity will help students learn more effectively.

Kumar & Clark's Clinical Medicine was one of the first textbooks to go online and is the progenitor of *Student Consult*. In this new edition our International Advisory Panel has expanded the online appendix of topics. There are 14 new chapters written by experts working in areas where the diseases covered are common. They provide more detailed advice and local emphasis than is possible in the printed text. *Kumar & Clark's Clinical Medicine* has always aimed to be a clinical book (with a very strong basis in science) and the view of doctors worldwide who see these clinical problems every day is most valuable.

In addition, the online book comes with one year's free access to *Cecil's Medicine* online; a selection of animated procedures (such as lumbar puncture, catheterisation, cannulation, aspiration and other procedures); heart and lung sounds, and an interactive surface anatomy resource.

Kumar & Clark's Clinical Medicine continues to be a book read and used in many parts of the world. The 6th edition has been translated into Turkish, Greek, Italian and most recently Chinese. In this new edition we remain totally committed to providing an accessible account of the current state of evidence on the prevention and treatment of diseases worldwide.

Kumar & Clark's Clinical Medicine is now the hub of a large body of related publications. The family of books includes *Pocket Essentials of Clinical Medicine* by Ballinger and Patchett, *Pass Finals* by Smith, Carty and Langmead, and our own *Acute Clinical Medicine* (based on clinical cases). There is also a series of pocket books on other medical specialities (paediatrics, psychiatry, clinical surgery, obstetrics and gynaecology, and general practice). Our readers worldwide include not only medical students and doctors but also members of the nursing, pharmacy and other allied healthcare professions. Since the book has been online, readers frequently email questions to the editors who do their best to provide sensible answers to often very difficult questions. These questions have now been collected into a new book in the Kumar and Clark family, *1000 Questions and Answers from Clinical Medicine*.

This book would not continue to be so successful without the constant input, help and guidance we get from colleagues, be they students, young doctors or senior medical experts. Everyone likes to engage with the content of the book, which means that it is constantly being reviewed so that it is always useful for the understanding and solving of clinical problems. This indeed is what *Kumar & Clark's Clinical Medicine* has always striven for; to be a book that allows professionals to understand how the problems occurred, how they can be diagnosed and finally how they can be treated.

As always we would like to thank our families, our own and those of our authors, for their patience and support when 'book work' interferes with family life.

The editors are supporting the charity Water Aid with a percentage of their royalties from the 7th edition.

PJK
MLC

This page intentionally left blank

Acknowledgements

A book of this size could never be completed without the help and advice of many. We would like to thank many of our colleagues who have helped in the preparation of this edition. Many have given us useful advice, helped us to collect photographs and indeed read the manuscripts to make sure that the contents are totally up-to-date. In particular we would like to acknowledge: Mr Peter Hamilton (for his help with the eye section of the diabetes chapter); Professor Graham Foster and Professor Mark Thursz (for their help with the hepatitis section); Dr Paul Jarmon, Professor Adrian Dixon and the X-Ray department of Addenbrooke's Hospital in Cambridge (for pictures in the neurology chapter); Mr Trevor Beedham (prenatal diagnosis); Mr Fausto Palazzo (surgical aspects of endocrinology); Mr Paul Hamilton (rheumatology); Professor David van Heels (genetics); and Dr Susannah Leaver for general help with the book from a junior doctor's perspective.

We are delighted to welcome several new contributors to this edition. They replace some of our authors who have stepped down from the book after many years of commitment and loyalty. We would like to thank Dr Jacqueline Parkin and Professor Anthony Pinching (immunology); Professor Don Jeffries (virology); Dr Teresa Tait (palliative care); Professor Carol Black and Dr Christopher Denton (rheumatology); and Professor Brian Colvin (haematology).

We were extremely saddened by the death of Professor Anthony Clare who had been an author from the very beginning of the book and who is much missed.

Our ward rounds and outpatient reviews are a continuing source of evidence-based education and we are grateful to our specialist registrars, and junior trainees, including Foundation officers, as well as our own medical students who continue to stimulate us by asking penetrating questions.

Our travels around the world give us much insight into the practice of medicine and we would like to thank the many who have escorted us through their hospitals and medical schools. We are grateful to those who write to us, and we are extremely grateful to our International Advisory Panel. Members provide very helpful advice about their regions and contribute to the online appendix articles. In particular we would like to thank Professor Janaka de Silva for his advice and expertise and careful editing of these online contributions.

We are extremely grateful for the skill and support of our publisher, Elsevier, whose staff have maintained a commitment and loyalty to the book. We would like to acknowledge: Pauline Graham, our commissioning editor, who has taken us through this new edition; the production team of Frances Affleck (project manager), Sarah Russell, Gillian Murray and Stewart Larking (design and illustration), and Janice Urquhart (development editor), who have all contributed to the production of this extremely high quality edition. We are also grateful for the meticulous work of the copy editors, Ruth Swan and Pat Croucher, and to Caroline Cockrell. Caroline has played a major part in the day to day activities of the book in arranging endless couriers and parcels, but most particularly we are grateful for her careful management of the International Advisory Panel and the 'Ask the Authors' questions and answers included in the book's website. We would also like to express our sincere thanks to the many other people behind the scenes who have contributed in so many ways; we thank them for their loyalty and commitment.

Finally, we would like to thank Ms Jillian Linton for her meticulous preparation and co-ordination of all the chapters, and also Ms Sophie Rambaud and the Princess Grace Hospital for general assistance.

This page intentionally left blank

1 Ethics and communication

ETHICS

Why study ethics and law applied to medicine? 1

The nature of medical mistakes 2

Respect for autonomy 3

Ethical and legal boundaries of the duty to protect life and health 6

COMMUNICATION

Communication in healthcare 8

The medical interview 11

Interviews in which difficulties in communication can be expected 13

Complaints: what to do if the relationship breaks down? 14

Culture and communication 15

Patients with impaired faculties for communication 16

New influences on communication 16

ETHICS

WHY STUDY ETHICS AND LAW APPLIED TO MEDICINE?

Professional concern about ethics and law

Many clinical choices created by advances in medical technology are essentially ethical rather than scientific. Doctors may be expert in understanding and applying clinical science, yet this expertise does not in itself answer many ethical questions about the circumstances in which such science should or should not be applied. For example, they may know a great deal technically about advanced life-support systems or the termination of pregnancy. Their knowledge, however, will not tell them whether or not it is ethical to withdraw ventilatory support from a severely brain-damaged patient who will not otherwise die or to perform a termination on a 13-year-old girl who does not want her parents to know that she is pregnant. Answers to these questions derive from moral beliefs and arguments.

Patients are increasingly aware of what they believe to be their human rights and expect doctors to respect them. These may differ in different countries but the basic principles are the same.

- Rights are claims for specific types of goods or services that individuals are believed to be entitled to make on others (e.g. free speech, access to primary and secondary education, access to an acceptable standard of medical care).
- In the UK, if patients believe their rights have been ignored by doctors, they may formally complain to the General Medical Council (GMC) or seek legal redress.
- As it pertains to medicine, the law establishes boundaries for what government and the courts – through statute and common law – have deemed to be acceptable professional practice. The GMC – the UK regulatory body – sets boundaries for acceptable standards of conduct and care through guidance documents formulated by its Committee on Standards of Professional Conduct and on Medical Ethics and decisions in cases heard under the jurisdiction of its Fitness to Practice Committee.
- Whether or not legal actions or formal complaints against doctors are successful, they contribute toward bringing the medical profession into disrepute.

The three duties of clinical care

The rights of patients may be summarized by three corresponding duties of care, which apply to all patients for whom doctors have clinical responsibility.

1. **Protect life and health**. Clinicians should practise medicine to a high standard, taking care not to cause unnecessary harm or suffering. Patients should only be given treatments which they need. Treatments should not be prescribed, for example, just because patients request them.
2. **Respect autonomy**. Competent humans have autonomy – the ability to reason, plan and make choices about the future. Respect for these attributes goes hand in hand with respect for human dignity. Doctors should respect the autonomy, and thus the dignity, of their patients. This duty to respect the autonomy of patients leads to two further rights – informed consent and confidentiality. Competent patients should be able to choose to accept proposed treatments and to control personal information which they divulge concerning such treatments. Denying patients such choice and control robs them of their human dignity.

Ask the authors

www.studentconsult.com

Box 1.1 General Medical Council (UK), Good Medical Practice

The importance of protecting life and health:

- 'Adequately assessing the patient's conditions, taking account of the history . . . the patient's views, and where necessary examining the patient'
- 'Prescribe drugs or treatment, including repeat prescriptions, only when you have adequate knowledge of the patient's health, and are satisfied that the drugs or treatment serve the patient's needs'
- 'Provide effective treatments based on the best available evidence'

The importance of respect for autonomy:

- 'Listen to patients, ask for and respect their views about their health and respond to their concerns and preferences'
- 'Share with patients, in a way they can understand, the information they want or need to know about their condition, its likely progression, and the treatment options available to them, including associated risks and uncertainties'
- 'Respect patients' privacy and right to confidentiality'

The importance of fairness and justice:

- 'You must treat your patients with respect whatever their choices and beliefs . . . (and must not) . . . discriminate against them by allowing your personal views to affect adversely your professional relationship with them or the treatment you provide or arrange'
- 'You must not refuse or delay treatment because you believe that a patient's actions have contributed to their condition'

Similar statements are made by a variety of professional bodies in many countries.

3. **Protect life and health and respect autonomy with fairness and justice**. In the conduct of public and professional life, it is generally thought that people have the right to expect to be treated equally. Medicine is no exception and doctors have a duty to practise accordingly. The access to, and quality of, clinical care should be based only on the dictates of need rather than arbitrary prejudice or favouritism.

FURTHER READING

British Medical Association. *The Impact of the Human Rights Act 1998 on Medical Decision Making*. London: BMA, 2000.

General Medical Council. *Good Medical Practice*. London: GMC, 2006.

World Medical Association. *Medical Ethics Manual*. Ferney-Voltaire, France: WMA, 2005.

Why should doctors take the duties of care seriously?

Professional regulation

Within the UK, the General Medical Council (GMC) plays a crucial role in regulating the practice of medicine. It is responsible for the registration of doctors, setting and monitoring the quality of their undergraduate education and disciplining them for unprofessional conduct. Doctors have no professional choice but to conform to the standards laid down by the GMC, which are based on the three duties of care (Box 1.1). These same duties are confirmed by other professional bodies like the Royal College of Physicians, the Royal College of Surgeons, the British Medical Association and the Medical Research Council.

National practices differ but all countries have similar regulatory bodies.

The law

The three duties of care are also enshrined in statute and common law, which also regulate medical practice. Doctors

Box 1.2 European Convention on Human Rights

Substantive rights which apply to evaluating good medical practice

Right to life (Article 2)
Prohibition of torture, inhuman or degrading treatment or punishment (Article 3)
Prohibition of slavery and forced labour (Article 4)
Right to liberty and security (Article 5)
Right to a fair trial (Article 6)
No punishment without law (Article 7)
Right to respect for private and family life (Article 8)
Freedom of thought, conscience and religion (Article 9)
Freedom of expression (Article 10)
Right to marry (Article 12)
Prohibition of discrimination (Article 14)

can be sued in civil law for financial compensation for any harm that they cause while failing in their professional duties. If it can be shown that this failure is intentional, or reckless causing extreme harm or death, doctors may face prosecution in the criminal courts and, if found guilty, imprisonment.

UK law includes the Human Rights Act 1998. This law incorporates the European Convention on Human Rights, making it legally binding in the UK (Box 1.2). The provisions of the Act impose duties on doctors to ensure that their clinical practice is not in violation of these rights – the same duties that all other doctors in Europe must observe. The Act also thus helps to ensure that what constitutes legal practice in the UK is judged by broader, trans-national moral standards.

Rational self-interest

The most rational way for doctors to ensure that their own medical treatment (as well as the treatment of those for whom they have deep personal feelings) meets high standards of care is to support, through the usual professional and legal channels, the right of all patients to the same high standards of care.

The clinical importance of trust

Patients will not trust doctors whom they suspect may ignore their human rights. Without trust, patients will not cooperate in their diagnosis and treatment, undermining the prospect for clinical success. Lack of trust also engenders a defensive and impersonal approach to medicine by both clinicians and patients, potentially spoiling the quality of patient care and professional life.

The doctor–patient relationship

Doctors are expected to treat patients, and their carers where appropriate, as active partners in the healing process. Patients and doctors should make decisions together about the patients' care and management.

THE NATURE OF MEDICAL MISTAKES

Doctors have a duty to protect the life and health of patients to an acceptable professional standard. What does this mean in practice and what are the penalties for not doing so?

Clinical negligence

If clinicians are suspected not to have provided an acceptable standard of clinical care, they may be sued for negligence – a breach of their professional duty. To win a legal action and financial damages for negligence, patients/claimants must provide convincing evidence to a judge that:

- they were harmed
- the harm was caused by the accused doctor
- the action which caused the harm was a breach of professional duty.

However, in practice, demonstrating that these three things have occurred poses problems. For example, the alleged harm may have occurred against the background of a complex medical condition or course of treatment, making it difficult to establish the actual cause.

When has a breach of professional duty occurred?

In the UK, whether or not a doctor has acted inconsistently with the duty to protect life and health to an acceptable standard is ordinarily decided in civil cases by a judge on the basis of testimony from expert witnesses.

- Such experts are selected to represent a responsible body of professional opinion. The testimony of these experienced clinicians will be used to help the court determine a professional standard which doctors working in their specializations should meet in clinical practice of the kind under dispute. Both the patient/claimant and doctor/defendant will try to produce experts to support their case.
- There will be clinical conduct so unreasonable that its negligence speaks for itself (e.g. a dramatic and harmful mistake about drug dosage).
- If acceptable expert witnesses are found for the doctor/defendant, who will state that under similar circumstances, they would have clinically responded in the same way as the doctor then the patient/claimant will probably lose the case. Provided that the doctor's testimony is logically consistent and compatible with other widely accepted professional beliefs about good clinical practice, it is likely that the doctor's actions will be found to be reasonable. This will not necessarily depend on how representative the clinical opinions given in court are of other doctors practising in the same field. Experts can differ about what is and is not clinically acceptable. In the face of such conflict, what a judge deems reasonable will partly depend on how probable the evidence presented against the doctor is judged to be (e.g. about any harm alleged by the patient). Such a civil standard of proof differs from the criminal standard which is more stringent and where juries must be convinced of guilt 'beyond reasonable doubt'. Equally, the more serious the allegation of clinical negligence, the higher will be the quality of evidence that will be demanded by judges before deciding the outcome of the case.

Inexperience is irrelevant

Lack of experience is not taken into account in legal determinations of negligence or the GMC's formal hearings. All doctors are expected to work to a professional standard of expertise based on similar clinical work of experienced clinicians. This does not mean that doctors have to be experts in everything. In the face of doubts about their ability or training to provide treatment to a reasonable standard, doctors should seek appropriate supervision and refuse to proceed otherwise.

Mistakes are not necessarily to be feared

Poor professional practice – for example, clinical care which is overly defensive – can result from unfounded fears of patients' readiness to make formal complaints or take legal action.

- All doctors make clinical mistakes in their professional careers. Yet it should be clear that a clinical error is not necessarily a negligent error. Responsible and experienced clinicians may testify that the erroneous action was and is unavoidable in that type of clinical practice. Under similar circumstances, they too sometimes make – or might make – the same error.
- Just because doctors can make professionally defensible mistakes does not mean that they should in any way relax their professional standards.
- Professional bodies within medicine, and the defence associations which insure doctors against negligence, recommend that doctors should be honest and apologetic about their mistakes, remembering that to do so is not necessarily an admission of negligence. As a result of such honesty and humility, there is mounting evidence that patients will feel that they have been respected and are less likely to take legal action or to make formal complaints.

For all of these reasons, there is less to fear legally from patients than doctors sometimes believe. The law offers wide protection for clinicians provided that they do their best to act reasonably and responsibly in protecting the lives and health of their patients.

RESPECT FOR AUTONOMY

Legally valid consent

Obtaining the consent of patients to treatment is just as fundamentally a part of good medical care, as is proper clinical diagnosis and good therapeutic management. Doctors seeking consent for a particular procedure must be competent in the knowledge of how the procedure is performed and its problems.

For agreement to treatment to be legally acceptable it must meet three conditions:

- Consent must be informed to an adequate standard.
- Patients must be competent to consent to treatment.
- Patients must not be coerced into accepting treatment against their wishes.

Patients need to weigh up the pros and cons of proposed treatments with other objective interests in their personal lives. They cannot do so without the basic ability to reason about information concerning what is wrong with them, what their doctors propose to do about it, and with what potential benefits and risks. If patients are coerced into making choices

FURTHER READING

Brazier M, Cave E. *Medicine, Patients and the Law*. Harmondsworth: Penguin, 2007.

Mason JK, Laurie G. *Mason and McCall Smith's Law and Medical Ethics*. London: Butterworths, 2006.

about treatment, the ethical and legal right to exercise control over their personal life is ignored. In these circumstances, such choices become more those of clinicians who unduly pressure patients rather than of patients themselves.

Battery

It is an unlawful battery to intentionally touch a competent person without their consent.

- For competent adult patients to be touched lawfully, they must be given information in broad terms about the proposed treatment – what it is and why it is being suggested. Thus a clinician will commit a battery if such a patient is given an injection without permission, irrespective of the need for it.
- Treatment can be given legally to adult patients without consent if they are temporarily or permanently incompetent to provide it and they have not previously refused such future treatment in a legally acceptable way or such treatment is not refused on their behalf by a representative with the legal authority to do so. Further, given the condition of such patients, treatment must be necessary to save their life, or to prevent them from incurring serious and permanent injury. Otherwise, consent must be obtained from patients expected to regain competence when they do so, however inconvenient this may be for the patient or clinician.

Negligence: information about risks

Clinicians may also be in breach of their professional duty to obtain adequate consent through not providing a reasonable amount of information about the risks of proposed treatment. Here the legal claim of a patient/claimant would be for negligence.

The reasonable doctor standard of disclosure

- Success will depend on the court being convinced that the patient/claimant had been harmed by the treatment and would not have agreed to it if they had been given more information about its risks.
- The patient/claimant must also show the amount of information provided about risks was unreasonable. If the clinician/defendant can find expert witnesses deemed to constitute a reasonable body of medical opinion who will say that they too (at that point in time) would have provided no more information about risks, the patient/claimant can face serious difficulties in convincing a judge of the unreasonableness of the amount of information, just as we saw in the case of litigation for clinical negligence.
- This 'professional standard' of disclosure of information about risks has been increasingly disputed in the courts.

The reasonable patient standard of disclosure

Patients may disagree with clinicians about how much information they require to protect their personal interests. Indeed, in deciding what information to disclose to patients about risks, clinicians may know little about how they perceive their interests. Since it is the health and lives of patients that are potentially at risk, the moral focus of such disclosure should be on what is acceptable to them rather than to the medical profession.

- It is increasingly recognized that clinicians obtaining consent to treatment should ask what sort of information about risks a 'reasonable person' in the position of the patient would want before agreeing to treatment. To do otherwise constitutes an unacceptable threat to the moral rights and dignity of patients and entails a potential loss of trust in the medical profession.
- Practically, clinicians should interpret this standard of disclosure of the reasonable person as meaning that they should ask what sort of information someone in the position of the patient needs, in order to make an informed choice. When in doubt, clinicians should also ask what sort of information they would want for themselves, their families and friends. They should also remember that the graver the risk, the more necessary it is to disclose information about it, even when the chance of it occurring is small.

Consent to treatment

This may be obtained in the following ways:

- ***Express consent*** may be verbal or written, usually through the patient signing a consent form. Here consent is given explicitly in relation to specific information about the proposed treatment. Consent to surgical treatment is usually 'express' and 'written'. Consent to lesser forms of physical intervention (e.g. venepuncture) is ordinarily 'express' and 'verbal'. Clinicians should always remember that a signed consent form is not legal or professional proof that proper express consent has been obtained. Informed consent should be viewed as an educational process. Signed consent forms are symbols of the completion rather than the success of this process.
- Consent may also be ***implied*** by the fact that the patient accepts treatment without question, protest or any other physical sign that might be associated with rejection. Implied consent is ordinarily given against the background of an express consent to a specific treatment, which has already been obtained. For example, once patients have given their express consent for being connected to a drip, there is no need for them to give further consent whenever its contents are replenished.
- Patients have not given their implied consent to a specific treatment simply because they have presented themselves for care in a hospital. They must be given appropriate information about the proposed care and provide express consent to it.
- Patients and doctors should make decisions in a spirit of partnership.
- Medical students or their supervisors should always obtain the explicit consent of patients to provide case histories or to be examined for purely educational purposes. Students should always make it clear to patients that they are not qualified doctors.

Confidentiality

If clinicians violate the privacy of their patients, they risk causing harm rather than protecting patients from it. Through violating the right of patients to control information which

they divulge as part of their medical care, such clinicians disrespect autonomy, undermine trust and call the medical profession into disrepute. These rights are protected by common and statute law. Doctors who breach the confidentiality of patients may face severe professional and legal sanctions.

Respecting confidentiality in practice

Patients should be informed of the ways in which information about them will need to be shared with other clinicians and healthcare workers involved in their treatment. On the basis of this information, unless patients state the contrary, they give their implied consent for this information to be shared. Where patients object to particular information being shared, their wish should be respected unless this interferes with the successful execution of their treatment. Information should be sought from patients as early as possible about whom they would wish to be given information about their condition and treatment. A note should be made of their wishes in the clinical record. Clinicians or healthcare workers not involved in the care of a patient have no right of access to related information without the patient's consent, simply because such information is considered useful for other purposes. In almost all clinical circumstances, therefore, the confidentiality of patients must be respected.

When confidentiality must or may be breached

The principle of privacy in medicine is not absolute. Sometimes, the law dictates that clinicians must reveal private information about patients to others in contexts that they may or do object to. At other times, they have the discretion to do so, in accordance with good professional practice. Both circumstances highlight the difficult ethical tension, which can be posed between the rights of individual patients and the interests of the public. The right to privacy does not entail the right to harm others in exercising it.

Therefore, clinicians must breach confidentiality when (among others):

- patients have infectious diseases which must be notified, through informing the relevant local authority officer
- they are presented with a court order by a judge or asked to do so by a judge in judicial proceedings
- police request information about patients who have been involved in a traffic accident
- patients are suspected of engaging in terrorist activity in the UK.

Clinicians have the discretion to breach confidentiality when they become aware of, for example:

- past or potential criminal and violent behaviour that has resulted, or is likely to result, in serious harm to the patient or others, such as child abuse – note that in these circumstances, the ordinary expectation would be in favour of discretion
- refusal of patients to comply with their own legal obligations to provide information about their medical condition to the relevant authority (e.g. to the Driver and Vehicle Licensing Agency); again, it would usually be expected for this discretion to be exercised
- infectious patients who pose a threat to specific individuals through undisclosed risks.

In all these circumstances, patients should be informed of any intent to breach their confidentiality, unless doing so may place the clinician or others at risk of serious harm. Clinicians should always remember that they are professionally accountable for discretionary breaches and may be asked to justify in court, or by the GMC in the UK, their decision either to disclose or not to disclose.

Respect for autonomy in the treatment of vulnerable patients

We have seen that for consent to treatment to be valid, patients must be competent to give it. In England, the legal criteria for the determination of competence are stated in the Mental Capacity Act 2005.

What is competence?

Competence should be understood as task-oriented. People may be competent to do some things but incompetent to do others. This means that they should not be judged to be either competent or incompetent in absolute terms. Legally, if patients are competent they must be able to:

- understand information about their condition and treatment
- remember this information
- deliberate about the therapeutic choices posed by the information
- communicate their choices to others.

Competence to consent to treatment may be compromised by many things, e.g. age, mental illness, congenital disease, accident and injury.

Children

In the UK, the legal age of presumed competence to consent to treatment is 16. Below this age, those with parental responsibility are the legal proxies of their children and usually consent to treatment on their behalf. Yet:

- The age of 16 is somewhat arbitrary and many children below this age clearly possess the abilities associated with competence. For example, they may be mature enough to understand and reason about information given to them about their condition and treatment.
- Clinicians should ordinarily respect the dignity of such children through asking them if they agree to the proposed treatment, even when the consent of parents is also obtained.
- If such children wish to have clinical consultations and clinically indicated care without the knowledge of their parents then this – along with their confidentiality – should be respected.
- Young children may be incompetent to make decisions about their medical care, although they may have developed some degree of autonomy in this regard. They should be consulted about their care and due consideration given to their wishes.
- In England, unlike Scotland, young people do not acquire the right to refuse medical treatment thought to be in their best interest until the age of 18. In practice,

the legal ability to force treatment on mature children against their wishes should only be contemplated in circumstances where life is at risk or there is a risk of serious and permanent injury. Doing otherwise is not in the best interests of young people because of the potentially dangerous alienation that it may create toward doctors and medicine.

Those with parental responsibility should always be consulted about their child's care. However, they have questionable legal authority to direct clinicians to administer or withdraw treatments deemed necessary to protect children from death or serious harm. The court should be consulted about such disagreements, unless an emergency dictates otherwise.

Psychiatric illness

The vast majority of patients being treated for psychiatric illness are competent to consent and to refuse treatment. There is no difference between their rights and those of other competent patients. Because of the danger of stigma associated with mental illness, great care should be taken to protect their confidentiality. If patients with severe psychiatric illness are incompetent to understand the nature and consequences of their illness, they will be unable to provide valid consent to treatment. Here, the ethical duty of care shifts from respect for autonomy of such patients to protecting them.

The Mental Health Act 2007

Mental illness may so compromise the autonomy of patients that they become a danger to themselves and/or to others. Patients with 'any disorder or disability of the mind' may be detained under the Mental Health Act 2007 for further examination and treatment in England and Wales. This abolishes the 'treatability' provision in the previous Act. Because of the terrible risks and ethical significance of denying someone their ordinary civil liberties, the conditions of detainment under the Act are highly specific: the longer the period of detainment, the more safeguards there are to ensure the clinical need for it (see Table 22.29). Relevant legislation in other legal jurisdictions have similar provisions.

Psychiatric treatment without consent

Under certain circumstances, detained patients may be given psychiatric treatment without their consent. However, attempts should be made to obtain consent, if possible. Unless their mental illness has made them incompetent to do so, such patients must still give their consent to proposed treatments for physical disorders. Again, incompetence in one respect does not entail incompetence in all respects. However, if patients are unable competently to consent because of the severity of their psychiatric condition – and treatment is required to save their life or to prevent serious and permanent disability – they can be given necessary care without it. If treatment can wait, without seriously compromising their interests, then patients should be asked to consent to it when they become competent to do so.

Other forms of incompetence to provide consent to treatment

Patients may also be incompetent because of congenital, developmental or accidental brain damage. Here, good legal and professional practice is again determined by the Mental Capacity Act 2005, along with its Code of Practice. The Act makes provision for adults to act as the legal representatives for adult patients who are permanently incompetent. Other legal jurisdictions (e.g. Scotland) also endorse forms of representative authority, some stronger than others. In England, if no appropriate representative is nominated by the patient or by the appropriate court, relatives *do not* have the final say about treatment decisions. In the absence of such delegated authority, not allowing relatives the authority to make such decisions can be morally justified by the fact that some people might not be motivated by protecting the best interests of their family members. Even so, relatives should be asked their views about the wishes or concerns of the patient and consulted about medically relevant information, which might be used to optimize the success of the patient's care. However, unless this is legally appropriate, they should not be given the impression as a result of such consultation that it is they who are determining treatment decisions. Therefore, often, doctors and no one else will continue to decide what is and is not in the best clinical interests of patients who are permanently incompetent. Without the intervening authority of legally designated representatives and/or legally acceptable advance decisions of refusal, all forms of treatment can be administered to such adults on this basis. Of course, the moral coherence of these arguments depends on the assumption that clinicians are best placed to protect the interest of vulnerable patients and this, in turn, rests on consistently high standards of clinical care and reflection about ethics and law.

ETHICAL AND LEGAL BOUNDARIES OF THE DUTY TO PROTECT LIFE AND HEALTH

Generally speaking, clinicians are professionally obligated to intervene to save the lives of patients for whom they have clinical responsibility. However, there are a range of circumstances where the duty of care to provide life-sustaining treatment is superseded by other ethical and legal responsibilities.

Competent refusal

The right of competent refusal supersedes the ordinary duty clinicians have to try to save the lives of such patients, along with any preferences others might have that they should be forced to accept treatment (Box 1.3). Patients may refuse life-sustaining treatment explicitly when it is offered or, as already suggested, they may formulate an advance decision refusing treatment. This should stipulate the circumstances under which they refuse it, should they become incompetent to do so in the future. Competent refusals of life-sustaining treatment (e.g. refusal of blood transfusion by Jehovah's Witnesses) must be made on the basis of clear information about the types of treatment refused. Equally, decisions not to attempt to sustain life medically should also only be made with the consent of competent patients (e.g. a decision not to provide cardiopulmonary resuscitation). Such patients may have arrangements, which they need to make that their

FURTHER READING

British Medical Association. *Consent, Rights, and Choices in Healthcare for Children and Young People*. London: BMJ Press, 2001.

British Medical Association. *The Mental Capacity Act 2005 – Guidance for Health Professionals*. London: BMA, 2007.

General Medical Council. *Confidentiality: Protecting and Providing Information*. London: GMC, 2004.

General Medical Council, *Confidentiality: Protecting and Providing Information: Frequently Asked Questions*. London: GMC, 2004.

Royal College of Psychiatrists. *Good Psychiatric Practice*. London: RCPsychiatrists, 2004.

Box 1.3 Competent refusal

'Prima facie every adult has the right and capacity to decide whether or not he will accept medical treatment, even if a refusal may risk permanent injury to his health or even lead to premature death. Furthermore, it matters not whether the reasons for the refusal were rational or irrational, unknown or even non-existent. This is so notwithstanding the very strong public interest in preserving the life and health of all citizens.'

Re T (adult: refusal of medical treatment) [1992]

clinicians will know nothing about and it would be ethically wrong to pre-empt this.

The best interests of the incompetent

There will be some situations where the provision of life-sustaining treatment will not be regarded as being in the best interests of permanently incompetent patients, even though not providing it will lead to their death occurring before it otherwise would. This is when prolongation of patients' lives will be of no benefit to them, against the background of their dire clinical circumstances. For example, it is acceptable not to use medical means to prolong the lives of such patients when:

- it is believed on good evidence that further treatment will not save life
- patients are already imminently and irreversibly close to death
- patients are so permanently or irreversibly brain damaged that they will always be incapable of any future self-directed activity or intentional social interaction.

One ethical justification for the latter two provisions is that such patients have no further need for medical treatment because they have no further objective interest in continuing to live. Because of their clinical condition, they can no longer do or achieve anything in life. However, because of potential differences in moral and religious belief, decisions not to provide or continue life-sustaining treatment should always be made with as much consensus as possible among the clinical team responsible for the patient's care and people close to the patient who are consulted about their care. In the absence of an advance decision refusing treatment, efforts should still be made to determine what the patients would have wished under these circumstances. Where there is unsolvable conflict between those involved in decision-making, especially in relation to children and their parents, the relevant court should be consulted. In some circumstances, the court may appoint a deputy to represent the best interest of the patient. Healthcare staff should appoint an independent healthcare advocate in the absence of other forms of appropriate representation. In emergencies, judges are always available in the relevant court. As already mentioned, other legal jurisdictions have similar arrangements. Key legal differences usually focus on the extent of decision-making authority automatically given to relatives of the patient.

Clinicians may decide not to prolong the lives of imminently dying and/or extremely brain-damaged patients for the legally acceptable reason that they are acting in the best interests of their patients to attempt to minimize their suffering rather than intending to kill them. To do the latter would be murder. Clearly, when contemplating not providing or withdrawing life-sustaining treatments, doctors should do their best to act within the law. However, it should also not be forgotten that much debate continues about whether or not there is a much closer ethical link between clinical decisions not to save life and decisions actively to end it. The law concerning such matters may change at some point in the future.

The duty as a doctor to be fair and just

The principle of equality of persons should dominate the way in which the duties of care are discharged in clinical practice. Clinicians may show technical mastery of treatment options and great ability and sensitivity in obtaining informed consent to treatment and maintaining confidentiality. However, if in the process of doing so, they actively discriminate against individual or groups of patients, they are still acting unprofessionally and should be penalized accordingly. For example, discrimination on the basis of race, age, social worth, class, intelligence or occupation should not be tolerated (Box 1.4).

Scarce resources

As a matter of right in the UK, the National Health Service (NHS) provides equal access to appropriate medical care on the basis of need alone. This right is mitigated by scarce resources and the courts have made it clear that they will not force health authorities or trusts to provide treatments which are beyond their means. However, the courts also demand that decisions about such means must be made on reasonable grounds and that patients have a clear right to expect this. Thus on both ethical and legal grounds, prejudice or favouritism is not acceptable in the allocation of scarce resources. If rationing healthcare is inevitable in the NHS, as many believe to be the case, then ways must be found to do so which are fair and just. In practice this means that scarce resources should be allocated to patients on the basis of the similarity and extremity of their need (e.g. triage) and the time at which they presented themselves for treatment (e.g. the randomness imposed by the 'lottery of nature'). To the degree that this procedure of allocation is followed, all

Box 1.4 The duty as a doctor to be fair and just

According to the General Medical Council, injustice can occur in medicine through treating patients unequally according to, e.g.:

- Age
- Colour
- Ethnic or national origin
- Gender
- Lifestyle
- Marital or parental status
- Race
- Religion or beliefs
- Sexual orientation
- Social and economic status

patients will be said to have had an equal opportunity to be treated on the basis of equal need. It is respect by clinicians and healthcare teams for these principles of equality – equal need and equal chance – that guarantees fairness and justice in the delivery of healthcare. In the UK, the National Institute for Health and Clinical Excellence (NICE) is a body set up to evaluate treatments on a clinical and cost-effective basis, and to ensure equal access to care across the country. In circumstances of scarcity, waste and inefficiency are clearly matters of great ethical concern.

More specifically, any well-run A&E department within the UK and many other countries will follow precisely these moral principles in the way in which decisions are made about who to treat first and with what. The treatments on offer will be those that have been shown to deliver optimal results for minimal expense. Patients who are in the most need – as determined by the triage nurse – will be seen first. Everyone else will receive treatment in relation to the assessment of the urgency of their need and when they initially register for treatment. In this way a waiting list will be constructed for each clinical session. So long as no one attempts to jump the queue, patients traditionally respect these rules – sometimes to a point of waiting for long periods to be seen without complaint. Waiting lists in other clinical settings should be similarly constructed, although here other variables will come into play (e.g. the accuracy with which general practitioners communicate the urgency of need to hospital consultants). What is unacceptable in any aspect of the allocation of scarce healthcare resources is for prejudice or favouritism – treating equals unequally – to influence their construction.

Lifestyle

Patients should not be denied potentially beneficial treatments on the grounds that their lifestyles have been more unhealthy than others with whom they compete for the same treatments. Decisions to do so are almost always prejudicial. For example, why single out smokers or the obese for blame as opposed to those who engage in dangerous sports? Patients are not equal in their abilities to lead healthy lives and to make correct healthcare choices for themselves. Some are better educated and informed, more emotionally confident and more supported by their social environment. To ignore this, to regard all patients as equal competitors and to reward the already better off is unjust and unfair. In this situation it is incorrect to treat unequals equally.

FURTHER READING

General Medical Council. *Withholding and Withdrawing Life-Prolonging Treatments*. London: GMC, 2002.

Newdict C. *Who Should We Treat? Rights, rationing and resources in the NHS*. Oxford: OUP, 2005.

GENERAL READING

Ashcroft R, Lucassen A, Parker M et al. *Case Analysis in Clinical Ethics*. Cambridge: CUP, 2005.

British Medical Association. *Medical Ethics Today: the BMA's handbook of ethics and law*. London: BMJ Books, 2004.

Tallis R. *Hippocratic Oaths: medicine and its discontents*. London: Atlantic Books, 2004.

COMMUNICATION

COMMUNICATION IN HEALTHCARE

The aim of every healthcare professional is to provide care that is evidence based, unconditionally patient centred and shared in a collaborative partnership with the patient. Good communication between a patient and the professional forms the basis of such care. Effective communication has been shown to develop the trust between patient and clinician and produces better outcomes. Fortunately the effective skills in healthcare communication are learnable.

What is patient-centred communication?

Patient-centred communication involves reaching a common ground about the illness, its treatment, and the roles that the clinician and the patient will assume (Fig. 1.1). This means discovering and connecting both the biomedical facts in detail and the patient's ideas and feelings. This information is essential for diagnosis and appropriate management and also to gain the confidence, the trust and the involvement of the patient.

The traditional approach of 'doctor knows best' with patients' views not being considered is now very outdated. This change is not just societal but driven by evidence about improved health outcome and is spreading world-wide.

There are two main reasons for this:

- Patients increasingly expect information about their condition and treatment options and want their views taken into account in deciding treatment. They also like to be equally involved if they wish to reduce risks from their habits and lifestyles. This does not mean clinicians totally abdicate power. Patients want their doctors' opinions and expertise and may still prefer to leave matters to the clinician.
- Clinicians can usually offer practical help with patients' feelings, ideas and concerns but, if not, they can always listen supportively. Patients may also hold misconceptions, which can profoundly affect their symptoms and ability to recover; these can be corrected with information. A distinguishing feature of the healthcare professions is that patients expect humanity and empathy from their doctors as well as competence.

Patient-centred communication tries to achieve a good balance between:

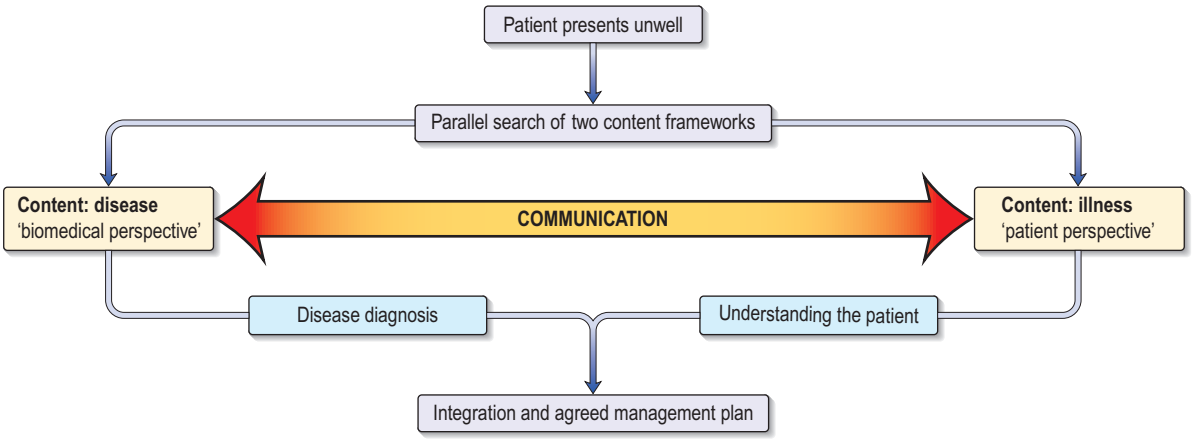


Fig. 1.1 **The patient-centred clinical interview.** Adapted from Levenstein JH. In: Stewart M, Roter D (eds) *Communicating with Medical Patients.* Thousand Oaks, California: Sage Publications, 1989, with permission.

- Clinicians asking all the questions needed to include or exclude diagnoses
- Patients being asked to express their ideas, concerns, expectations and feelings
- Clinicians explaining and advising in ways patients can understand and be involved in decisions.

What are the effects of communication

There are enormous benefits to be gained from good communication (Box 1.5). When clinicians communicate effectively with patients they identify patients' problems more accurately and efficiently. When they do not they can miss problems and concerns. The failures in identifying problems are shown in Box 1.6.

(i) Health outcomes

Several widely different conditions have been found to be improved by good communication (Box 1.7). The converse also applies. For example, the main predictive factor for patients developing full-blown psychiatric disorders on learning of the diagnosis of cancer depends upon the way their bad news is broken. Hospital visits, admissions and length of stay were dramatically reduced where clinicians used a biopsychosocial approach to managing patients with medically unexplained symptoms. The mortality rate was also significantly reduced.

Box 1.5 Benefits of good communication

- Improved health outcomes
- Increased patient adherence to therapies
- Increased patient and clinician satisfaction
- Reduced litigation
- Improved time management and costs
- Patient safety

Box 1.6 Failures in identifying problems

Patients report the following problems in interviews:

- 54% of complaints and 45% of concerns were not elicited
- 50% of psychological problems not elicited
- 80% of breast cancer patients' concerns remain undisclosed
- In 50% of visits, patients and doctors disagreed on the main presenting problem
- In 50% of cases, patient's history was blocked by interruption within 24 seconds

Simpson M, Buckman R, Stewart M et al. The Toronto Consensus Statement. *British Medical Journal* 1991; **303**: 1385–1387

Box 1.7 Improvements in health outcomes with good communication

- Postoperative pain control
- Chronic headache relief
- Blood pressure control
- Diabetic control
- Asthma control
- Emotional health
- Symptom resolution
- Function improvement

Stewart MA. Effective physician–patient communication and health outcomes. *Canadian Medical Association Journal* 1995; **152**: 1423–1433

(ii) Adherence to treatment

Some 45% of patients are not following treatment advice. Costs of medicines prescribed and not used runs into millions of pounds wasted in the UK. Patients may not understand or remember what they were told, whilst others actively decide not to follow advice and commonly do not tell their doctors.

Errors in use of medications are also a common problem, which risk patient safety. Research on consultations shows clinicians rarely check patients' understanding or their views and doubts. Communication is a major contributing factor to adherence (Practical Box 1.1)

(iii) Patient satisfaction

Satisfaction, as reported to the Patients' Association in the UK, is largely a result of patients knowing they are:

- getting the best appropriate biomedical healthcare
- being treated humanely as individuals and not as items on a conveyer belt.

Practical Box 1.1 Communication factors which improve adherence to clinical advice

Clinician

- Listens to and understands the patient
- Tone of voice
- Elicits all the patient's health concerns

Patient

- Is comfortable asking questions
- Feels sufficient time is spent with the clinician

Stewart M, Brown JB, Boon H et al. Evidence on patient communication. *Cancer Prevention and Control* 1999; **3**(i): 25–30

Practical Box 1.2 Behaviours leading to good relationships

Primary care physicians who have never been sued:

- Orientated patients, e.g. 'We are going to do this first and then go on to that'
- Used facilitative comments, e.g. 'uh huh, I see'
- Used active listening
- Checked understanding
- Asked patients their opinions
- Used humour and laughter appropriately
- Conducted slightly longer visits (18 versus 15 minutes)

Levinson W. Physician–patient communication. The relationship with malpractice claims among primary care physicians and surgeons. *Journal of the American Medical Association* 1997; **277**: 553–559

Satisfaction affects psychological well-being and adherence to treatment both of which have a knock on effect on clinical outcomes. It also reduces patient complaints and litigation.

(iv) Clinician satisfaction

Healthcare professionals have a very high rate of occupational stress and burnout, which is costly to both them and health services. Notwithstanding pressure from staffing shortages and inadequate resources, it is the quality of their relationships with their patients and colleagues which are the most reliable global indicator of clinician satisfaction and happiness.

(v) Dissatisfaction between patients and clinicians

Failures in communication of information about illness and treatment are the most frequent source of patient dissatisfaction. Some 70% of lawsuits are a result of poor communication rather than failures of biomedical practice. Patients felt deserted and devalued by their clinician and thought the information to aid their understanding was delivered poorly. Authoritarian and paternalistic styles revealed in questioning style and tone of voice increase the chances of being sued. By contrast, the behaviours of primary care physicians who have never been sued are shown in Practical Box 1.2. Complaints and lawsuits represent only the tip of the iceberg of discontent as revealed by surveys of patients in hospital and general practice.

(vi) Time and costs

Clinicians commonly interrupt patients, after an average 24 seconds, whether or not a patient has finished explaining their problem. Uninterrupted patients will talk for 90 seconds on average (maximum 2.5 minutes). If clinicians wait they avoid the situation of the late-arising problems, which then take longer. Those who integrate *patient-centred communication* into all interviews actually save time and also reduce non-essential investigations and referrals, which waste resources. Time skimped deprives the patient and actually involves more time later either for themselves or colleagues.

Patients given the latest evidence on treatment options commonly choose more conservative management with no adverse effects on health outcomes. This has potential for considerable savings in health budgets.

Box 1.8 Strategies that doctors use to distance themselves from patients' worries

Selective attention to cues	'I have this headache and I'm worried . . .'
	'What is the pain like?'
Normalizing	'It's normal to worry. Where is the pain?'
Premature reassurance	'Don't worry. I'm sure you'll be fine'
False reassurance	'Everything is OK'
Switching topic	'Forget that. Tell me about . . .'
Passing the buck	'Nurse will tell you about that'
Jollying along	'Come on now, look on the bright side'
Physical avoidance	Passing the bedside without stopping

Maguire P. *Communication Skills for Doctors*. London: Arnold, 2000

Box 1.9 Common barriers and difficulties in communication

Clinician factors:

- Lack of knowledge of:
  influence and prevalence of psychosocial matters in illness and its recovery
  skills that can help communication
- Attitude:
  authoritarian manner and negative attitude to shared care
  unwillingness to examine own communication skills
- Skills:
  using distancing tactics to avoid difficult topics
  using jargon
  lack of empathy

Patient factors:

- Anxiety
- Feeling powerless
- Reticence to disclose concerns
- Misconceptions
- Conflicting information
- Forgetfulness
- Hearing/visual and speech impairment

Shared factors:

- Different first language
- Lack of privacy
- Lack of time
- Different cultural backgrounds

Barriers and difficulties in communication

Only recently has healthcare education included training to help clinicians communicate well. Hence many remain

unaware of the influence of psychosocial issues on illness (25–50% of primary care visits are for medically inexplicable complaints), or their costly impact on healthcare. They will avoid such discussions if they fear it will unleash emotions too difficult to handle, upset the patient or take too much time (Box 1.8).

Patients for their part won't disclose concerns if they sense clinicians aren't interested. Anxiety or embarrassment may also cause them to withhold information. Other common barriers are listed in Box 1.9.

Clinicians are human and can be rushed, unhappy, and have ignorant patches. They all work against the clock and in fallible systems. But as professionals it is they, together with healthcare managers, who bear the responsibility for dealing with these difficulties and problems; not the patient.

THE MEDICAL INTERVIEW

Structure and skills for effective interviewing

Clinicians conduct some 200 000 medical interviews during their careers. There is no single prescription for a medical interview, and flexibility is a skill in itself because each patient is different. However a framework helps clinicians use their time and that of patients' most productively.

The example below is in the context of a first medical interview in a consulting room and will obviously be different in a follow-up interview, or an emergency.

There are seven essential steps in the medical interview:

1. Building a relationship

Because patients are frequently anxious and may feel unwell, building a relationship is the first stage in the interview. Introductions and first impressions are critical to create rapport and trust. Well-organized arrangements for appointments, reception and punctuality put patients at ease. Non-verbal communication is a powerful element, e.g. making eye contact, facial expression and tone of voice should all convey friendliness, interest and respect. Sitting at a slight angle, not behind a desk or computer, helps communication.

Clinicians' own non-verbal messages, body language and unspoken attitudes have a huge impact on the emotional tone of the interview.

2. Opening the discussion

The aim is to obtain all the patients' concerns, remembering that they usually have at least three (range 1–12). Ask 'What would you like to discuss today?' or 'What problems have brought you to see me today?'

Listen attentively without interrupting. Ask 'And is there something else?' to screen for any other problems before exploring the history in detail. Clinicians are failing if a serious point is raised only as the patient has a hand on the doorknob, preparing to leave.

Only when all concerns are identified can the agenda be prioritized, balancing the patients' concerns with the clinician's main points of interest.

Table 1.1 Components of a medical interview

The nature of the key problems
Clarification of these problems
Date and time of onset
Development over time
Precipitating factors
Help given to date
Impact of the problem on patient's life
Availability of support
Patient's ideas, concerns and expectations
Patient's attitude to similar problems in others
Screening question

3. Gathering information

The components of a complete history are shown in Table 1.1.

Listening skills

Attentive listening is a necessary communication skill. Patients will recognize clinicians are listening if they look at them and not the notes or computer. Clinicians should nod occasionally, smile appropriately and avoid interrupting before the patient has finished talking. Ask the patient to tell the story of the problem in their own words from when it first started up to the present.

Questioning styles

The way a clinician asks questions determines whether the patient speaks freely or just gives one word or brief answers (Practical Box 1.3). Start with open questions ('What problems have brought you in today?') and move to screening, focused and closed questions. Open methods allow clinicians more time to generate their problem-solving approach and provide more information on which to base their theories and hypotheses. Closed questions are necessary to check specific symptoms but if used too early on may narrow down and lead to inaccuracies by missing patients' problems.

Leading questions which imply the expected answer ('You've given up drinking haven't you?') risk inaccurate responses as patients may want to go along with the clinician rather than disagree.

4. Understanding the patient

Finding out the patient's perspective is an essential step towards achieving common ground.

Ideas, concerns and expectations (ICE)

Patients seek help because of their own ideas or concerns about their condition. If these aren't heard they may think the clinician has not got things right and then not follow advice. Moreover any misconceptions will continue to prevail. A patient's view can emerge if the clinician listens carefully and picks up on cues. But if they do not, it is still a very necessary task to ask specific questions, e.g.:

- 'What do you think is wrong?'
- 'Are there any particular concerns you have about . . .?'
- 'Was there anything you were hoping we might do about this?'

FURTHER READING

Ambady N, LaPlante D, Nguyen T et al. Surgeon's tone of voice: a clue to malpractice history. *Surgery* 2002; **132**(1): 5–9.

Coulter A, Ellins J. Effectiveness of strategies for informing, educating, and involving patients. *British Medical Journal* 2007; **335**: 24–27.

Margalit A, El-Ad A. Costly patients with unexplained medical symptoms. A high-risk population. *Patient Education and Counseling* 2008; **70**: 173–178.

Stewart M, Brown JB, Weston WW et al. *Patient-centred Medicine: transforming the clinical method*, 2nd edn. Abingdon: Radcliffe Medical Press, 2003.

Practical Box 1.3 Questioning style

Closed questioning style:

Dr. 'You say chest pain, where is the pain?' (*closed Q*)
Pt. 'Well just here' (pointing to sternum)
Dr. 'And is it a sharp or dull pain?' (*closed Q*)
Pt. 'Quite a sharp pain'
Dr. 'Does it go anywhere else?' (*closed Q*)
Pt. 'No, just there'
Dr. 'And you don't smoke now do you?' (*leading Q*)
Pt. 'Well. . .just the occasional one'

Open questioning style:

Dr. 'Tell me about the pain you've been having.' (*open/focused Q*)
Pt. 'Well it's been getting worse over the past few weeks and waking me up at night. It's just here (points to sternum), it's very sharp and I get a burning and bad acid taste in the back of my throat. I try to burp to clear it. I've taken antacids but they don't seem to be helping now and I'm a bit worried about it. I'm losing sleep and I've got a busy workload so that's a worry too.'
Dr. 'I see. So it's bothering you quite a lot. Anything else you've noticed?' (*empathic statement, open screening Q*)
Pt. 'I've noticed I get it more after drinking and I have been drinking and smoking a bit more recently. Actually I've been getting lots of headaches too which I've just taken Ibuprofen for.'
Dr. 'You say you are worried, is there anything in particular that concerns you?' (*picks up on patient's cue and uses reflecting question*)
Pt. 'I wondered if I might be getting an ulcer.'
Dr. 'I see. So you've had this sharp pain under your breastbone with some acidity and burping for several weeks which is worse at night and aggravated by drinking and smoking but not relieved by antacids. You are busy at work, getting headaches, drinking and smoking more and not sleeping well. You're concerned this could be an ulcer.' (*summarizing*)
Pt. 'Yes, a friend had problems like this.'
Dr. 'I can appreciate why you might be thinking that then.' (*validation*)
Pt. 'Yes and he had to have a 'scope' so I wondered whether I would need one?'
Dr. 'Well, let me explain first what I think this is and then what I would recommend next . . .' (*signposting*)

Hearing patients' ideas, acknowledging their concerns and empathizing, are essential steps in engaging a patient's trust, and beginning to 'treat the whole patient'. The example in Practical Box 1.3 also shows how information that may be biomedically relevant emerges while listening to patients.

Non-verbal communication

In adult conversation some 5% of meaning derives from words, 35% from tone of voice and 65% from body language and non-verbal communication. When there is mismatch between words and tone the non-verbal communication elements hold the truer meaning. Patients who are anxious, uneasy, puzzled or confused are more likely to communicate this through expression and/or restless activity, for example of the feet and hands, than to tell the clinician outright. The observant clinician can pick up on this 'You seem uneasy about what I have said . . .' thus inviting the patient to share their concerns.

Empathizing

Empathy has been described as 'imagination for others'. It is different from sympathy (feeling sorry for the patient), which rarely helps. Empathy is a key skill in building the patient–clinician relationship and is highly therapeutic. The starting point is attentive listening and observing patients to try and understand their predicament. This understanding then needs to be conveyed back in a supportive way. For example, by saying: 'It sounds like . . .' (patient heard), 'I can see you are upset' (patient seen), 'I realize that this is a shock' (acknowledgement), 'I can tell you that most people in your circumstances get angry' (accepting the patient).

Like other communication skills, empathy can be taught and learnt but it has to be genuine and cannot be counterfeited by a repertoire of routine mannerisms.

Whilst empathy is about trying to understand, the phrase 'I understand' may be met with 'how could you!' It is more helpful to reflect back using some of the patient's own words and ideas.

5. Sharing information

The key is to tailor the information to both what the patient wants to know and the detail they prefer as this helps them understand.

Patients generally want to know; 'is this problem serious and how will it affect me; what can be done about it; and what is causing it?' Research has shown it is difficult for patients to take in explanations about cause if they are still worrying about the first two concerns.

There is a large amount of research to show that:

- Information must be related not only to the biomedical facts, but to patients' ideas about their condition and tailored to what they want to know.
- Most patients want to be fully informed irrespective of socioeconomic group.
- Most patients of all ages will understand and recall 70–80% of even the most unfamiliar, complex or alarming information if it is provided along the guidelines given below:

  Find out what patient already knows or thinks is wrong. Then confirm, correct or add new information.

  Chunk and Check – give information in assimilable chunks, one thing at a time. Check after each chunk that the patient is understanding.

  Signpost. 'I'll explain first of all . . .'. 'Can I move on now to explain treatment options . . .' Explain the cause and effect of the condition in the context of the patient's symptoms, e.g. 'The reason you are experiencing . . . is because . . .'

  Use *simple language*; translate any unavoidable medical terms and write them down.

  Make your *information and instructions direct*, detailed and concrete.

  Tell patients you will *write down* the key words in case they worry they will forget. Use simple diagrams, and offer them other aids to memory. Provide information to take away, or a source of it (Practical Box 1.4). Write out the risks of operations for patients, which tend to be rapidly forgotten after operations, however well understood before.

  Check understanding. Ask patients what they have understood – explain that this is not to test them but to make sure you have explained it clearly enough.

  Always remember that information about the patient belongs to the patient, not the clinician and not, without permission, the relatives or anyone else.

Practical Box 1.4 Possible aids to giving information

Method	Example
Write it down/use a diagram	As you go – make a copy for your records
Send the patient a copy of your report	With a patient's paragraph as a routine
Make an audio recording	
Use prepared information	Leaflets, audio or video
Provide address of support organization	Local cancer services, e.g. Macmillan Cancer Support
Provide a website reference	Local website in your country

6. Reaching agreement on management

Once the patient has understood the information, the clinician and patient need to agree on the best course as regards possible investigations and treatments.

Negotiating the next steps – enlisting the patient's collaboration

The clinician's opinion should be given and the patient's views sought. Check if they want to be involved or leave decisions up to the clinician.

Patients will adhere best to suggestions about investigation and treatment when they are thus enlisted as partners as a result of:

- a frank exchange of information
- a negotiation of options
- involving the patient in decisions.

Summarizing

This allows a patient to correct or add information and correct any misunderstandings. It is a feature of shared decision making and good practice.

7. Providing closure

Summary

A final summary of agreed course of action and next steps indicates the closing stage of the interview. Plans for follow-up and informing other healthcare professionals involved in the patient care are confirmed.

Closing

Some clinicans close the interview by asking whether the consultation has been useful before saying goodbye.

INTERVIEWS IN WHICH DIFFICULTIES IN COMMUNICATION CAN BE EXPECTED

The skills above form the basis of any patient interview but some interviews are particularly difficult:

- breaking bad news
- complaints and lawsuits
- cultural differences
- patients with impaired faculties for communication.

Breaking bad news

Bad news is any information which is likely to alter drastically a patient's view of the future. The way news is broken has an immediate and long-term effect. When skillfully performed the patient and family are enabled to understand, cope and make the best of even very bad circumstances. These interviews are difficult because biomedical measures may be of little or no help, patients are upset and can react unpredictably, and the clinician may be upset as well – the more so if there is an element of medical mishap. The most difficult things clinicians report are how to be honest with the patient whilst not destroying hope and how to deal with the patient's emotions.

Withholding information from patients or telling just the family is becoming a thing of the past even in those parts of the world where traditionally disclosure did not occur. Although truth can hurt, deceit hurts more. It erodes trust and deprives patients of information to make choices. Most people now express the wish to be told the truth and the evidence is that patients:

- usually know more than anyone realizes and may imagine things are worse than they are
- appreciate clear information even about the worst news and want the opportunity to talk openly and ask questions rather than join in a charade of deception
- differ in how much they can take in at a time.

A framework

Having a plan helps clinicians to present bad news in a factual, unhurried, balanced and empathic fashion that responds to each patient. The S-P-I-K-E-S strategy is:

S – Setting

- See the patient as soon as possible once the current information has been gathered.
- Ask not to be disturbed and hand over bleeps to colleagues.
- If possible, the patient should have someone with them.
- Choose a quiet place with everyone seated, introduce everyone.
- Indicate your status and the extent of your responsibility toward the patient and the time available.

P – Perception

Before telling, ask what has happened to the patient since the last appointment, what have they been told or construed so far and how have they reacted. This stage helps gauge the patient's perception, but should not be too drawn out.

I – Invitation

- Indicate you have the results and ask if the patient wants you to explain.
- Assess how much the patient would like to know.
- If patients do not want details, offer to answer any questions they may have later or to talk to a relative or friend.

FURTHER READING

Lloyd M, Bor R. *Communication Skills in Medicine*, 2nd edn. Edinburgh: Churchill Livingstone, 2004.

Silverman J, Kurtz Z, Draper J. The Calgary–Cambridge guide: a guide to the medical interview. In: *Skills for Communicating with Patients*, 2nd edn. Abingdon: Radcliffe Medical Press, 2005. www.skillscascade.com/index.html

K – Knowledge

If the patient wishes to know:

- Give a warning to help the patient prepare: 'I'm afraid it looks more serious than we hoped.' Then give the details.
- At this point WAIT: allow the patient to think, and only continue when the patient gives some lead to follow. This pause may be a long one, commonly a matter of minutes, while thoughts go round in the patient's head, and is often accompanied by shutdown which makes patients unable to hear anything further until these thoughts settle down.
- Give the information without fudging, in small chunks. Avoid technical terms. Check understanding frequently 'Is this making sense so far?' before moving on. Watch for signs the patient can take no more.
- Emphasize that some things, for example pain and other symptoms, are fixable and others are not. A broad time frame can be given for the fixable aspects of care.
- Be prepared for the question: 'How long have I got?' Avoid the trap of providing a figure which is bound to be inaccurate. There are no easy answers and this is where training can help rehearse responses that work for each clinician. Common faults are to be overly optimistic. Stress the importance of ensuring that the quality of life is made as good as possible from day to day.
- Provide some positive information and hope tempered with realism.

E – Empathy

Responding to the patient's emotions is about the human side of medical care and also helps patients to take in and adjust to difficult information. A range of emotions are experienced in seriously and terminally ill patients (Box 1.10).

- Be prepared for the patient to have disorderly emotional responses of some kind. Acknowledge them early on as being what you expect and understand and wait for them to settle before continuing.
- Crying can be a release for some patients. Allow time if the patient needs it rather than rushing in to stop the crying.
- Learn to judge which patients wish to be touched and which do not. You can always reach out and touch their chair.
- Always watch out for shutdown.
- Keep pausing to allow patients to think and frame their questions.
- Stop the interview if necessary and arrange to resume later.

Box 1.10 Emotional responses to the fact or threat of loss: some or all of these can be expected but in no particular pattern

- Despair
- Denial
- Anger
- Bargaining
- Depression
- Acceptance

Kubler Ross E. *On Death and Dying*. New York: Macmillan, 1970

S – Strategy and summary

Patients who have a clear plan for the next and future steps are less likely to feel anxious and uncertain. The clinician must ensure that:

- the patient has understood what has been discussed because thought processes are disorganized at times of emotion and misconceptions can take root
- the patient writes down crucial information to take away
- the patient knows how to contact the appropriate team member and thus has a safety net in place, the next interview date – preferably soon – has been agreed, for what purpose and with whom
- family members are invited to meet the clinicians as the patient wishes and written material or further sources of information are readily available
- everyone is bid goodbye, starting with the patient.

Follow-up

Bad news is a process and not a one-off. Patients may well not remember everything from the last visit and recapping is necessary. Always start by asking patients what they have understood so far. It is extremely distressing for patients to hear different things from different clinicians. Keep colleagues informed and document accurately what was said.

The move from active treatment to palliative care is a difficult stage in bad news. Patients want to know what happens next. They want to know 'Will I be in pain?', 'Can I stay at home?' and 'How long do I have?' Give clear answers with acknowledgement of any uncertainty. The priorities in patient care now are relief of symptoms, quality of life and enabling the patient to settle family matters or unfinished business.

The clinician's role is to mediate between the patient, other medical staff and the patient's relatives whilst continuing to be an empathic and caring doctor.

COMPLAINTS: WHAT TO DO IF THE RELATIONSHIP BREAKS DOWN?

Much of the enormous increase in complaints and medical lawsuits is related to failures in communication.

Clinicians tend not to deal effectively with grumbles and complaints at the appropriate time, which is as soon as they are made. They use avoiding strategies as for other forms of difficulty and tend to become defensive. The consultation that occurs after an adverse experience is crucial in influencing any decision to sue.

A national commission in the UK on patients' complaints recommended that all healthcare professionals should have training in communication and be aware of their responsibility for dealing promptly and personally with complaints made by patients in their care. They should not automatically dump complaints on managers.

Complaints

The majority of complaints come from the exasperation of patients who:

FURTHER READING

Baile WF, Lenzi R, Kudelka AP et al. SPIKES – A six-step protocol for delivering bad news: application to the patient with cancer. *Oncologist* 2000; **5**: 302–311.

Fallowfield LJ, Jenkins VA, Beveridge HA et al. Truth may hurt but deceit hurts more: communication in palliative care. *Palliative Medicine* 2002; **16**: 297–303.

Maguire P, Pitceathly C. Managing the difficult consultation. *Clinical Medicine* 2003; **3**: 532–537.

Tse CY, Chong A, Fok SY. Breaking bad news: a Chinese perspective. *Palliative Medicine* 2003; **17**: 339–343.

- have not been able to get clear information
- feel that they are owed an apology
- are concerned that other patients will go through what they have.

Many complaints are resolved satisfactorily once these points are dealt with, and the sooner the better. An apology, as an expression of regret, as medical defence organizations emphasize, is not an expression of guilt; it is a courtesy. Nor should it be wrapped up so cautiously as to become a worthless token. Clinicians should work in a professional culture, which regards complaints as a valuable source of feedback which deserve to be noted, collected systematically and acted upon if need be.

When clinicians learn of a complaint they should:

- remember the complainant is still a patient
- listen to the complainant and acknowledge distress
- be objective, not resentful or defensive – remember that there is a duty of care to the patient in this as every other way
- allow all the facts to speak through a clear account, verbal or written
- express regret
- explain the reasons and circumstances behind the facts
- explain how things will improve
- leave the medical records strictly unaltered.

Honest disclosure in the case of medical injury increases trust and reduces the likelihood of legal action.

Lawsuits

Lawsuits, the extreme form of complaint, are commonly rooted in poor communication or miscommunication, aggravated by a sense of grievance. Some 17% of patients affected by medical injury in the UK want financial compensation or disciplinary action. Any clinician faced with a lawsuit must seek specialist advice.

CULTURE AND COMMUNICATION

Whilst doctors strive to treat all patients equally, those from minority cultures get poorer healthcare than others of the same socioeconomic status even when they speak the same language. They experience fewer expressions of empathy, shorter consultations and fewer attempts to include them in shared decision-making. They also tend to say less in consultations.

Clinicians commonly express anxiety and uncertainty about how to respond to cultural diversity, how to use interpreters, and how to avoid causing offence.

Beliefs

We all take our culture for granted but it can profoundly affect notions about symptoms, sickness and appropriate behaviour and treatment.

It influences practices about when to seek medical assistance and what patients and doctors expect of the consultation. In some cultures for example it is very difficult for a woman to see a male doctor. Sometimes family members may think it is their duty to talk for the patient whilst the doctor might expect to talk directly with the patient. Sensitive topics may be difficult to address but failure to do so could jeopardize care. It helps to apologize if offence is inadvertently caused and to explain why such questions are required. Clinicians vary too and those from more traditional cultures may have a more paternalistic style than some patients want.

Language

Patients sometimes bring a family member, friend or even child to interpret. This can be problematic because they may not have sufficient vocabulary, they could be censoring sensitive matters or saying things that suit their agenda. Confidentiality cannot be guaranteed and patients may feel restricted in what they can say. Children should not be used to interpret although often still are.

When communicating through an interpreter/advocate, ask for the correct pronunciation of a patient's name and if there are any cultural differences in body language. Arrange the seating to see both the patient and advocate but look and speak directly to the patient. Speak in short phrases, avoid jargon and find out the patient's ideas, concerns and expectations. Watch for non-verbal communication and check the patient's understanding. Whilst examining the patient, ask the advocate to stand outside the curtain and always thank the advocate at the end of the interview. If professional interpreters are not available use telephone language lines or ask colleagues who speak the patient's language. Advocates are people from the patient's culture who can do more than translate. They can explain beliefs and concerns that could be relevant in the patient's culture and help patients to understand the workings of the healthcare systems.

Clinicians sometimes worry that interpreters are editing when long exchanges are followed by only a short summary back to them. It helps to ask interpreters to translate exactly what has been said, e.g.:

Patient: Long explanation.
Interpreter: She says she is sick.
Doctor: Could you please translate exactly what she said so that I can understand everything.
Interpreter: I have been feeling dizzy, my head hurts all the time and this morning I was sick.

Non-verbal communication

Awareness of cultural taboos, for example handshaking, eye contact, personal space and subjects that would be offensive to talk about, can help in maintaining dignity and respect.

Paraverbal communication also varies across cultures. We infer things from tone of voice, stress on words and phrases, silence, pace, and politeness conventions used. Some cultures are more open, direct and assertive than others. Some languages do not differentiate gender in common nouns and pronouns, so 'he' and 'she' may be used interchangeably. It is hardly surprising that misunderstandings occur and it can be much harder to create rapport. It is worth remembering that smiling is a universal expression of kindness and warmth.

Patients may be more or less traditional so check out assumptions.

Training materials are available to help health professionals to respond to the needs of ethnically diverse patients (http://info.cancerresearchuk.org/proceed/).

FURTHER READING

Johnson RL, Roter D, Powe R et al. Patient race/ethnicity and quality of patient–physician communication during medical visits. *American Journal of Public Health* 2004; **94**: 2084–2090.

Kai J (ed). *Ethnicity, Health and Primary Care*. Oxford: Oxford University Press, 2003.

Schouten BC, Meeuwesen L. Cultural differences in medical communication: a review of the literature. *Patient Education and Counselling* 2006; **64**(1–3): 21–34.

Ask the authors
www.studentconsult.com

PATIENTS WITH IMPAIRED FACULTIES FOR COMMUNICATION

All healthcare professionals need patience, ingenuity and willingness to learn to be able to communicate effectively with patients who have impaired communication faculties.

Impaired hearing

Patients with severely impaired hearing may be accompanied by a signer in one of the two sign languages but less than 1% of hearing-impaired people sign. Many hard of hearing people lip-read and some commonsense tips not always followed are listed in Practical Box 1.5.

Conversation aids are available which involve a microphone and amplifier. Specially adapted textphones, Minicoms, allow people to type messages which the operator relays to the person on the phone. Mobile textphones are also now available.

Impaired vision

Patients with visual impairment can miss non-verbal cues in communication. Whilst it may sound obvious, it helps to make more conscious efforts to use the patient's name so they know they are being spoken to. Clinicians should avoid sudden touch, explain what they are about to do and say what they are doing as they go along. Large print information sheets should be available for those with limited vision; audiotapes and Braille versions for the blind (www.rnib.org.uk).

Patients with limited understanding or speech

Aphasia is a communication disorder (see p. 1099). Even though hearing and thought processes are unaffected, patients find it hard to understand or they know what they want to say but cannot find the words; they are literally 'lost for words'. This also affects their ability to write, gesture, draw or mime their thoughts.

Patients will have a strength in one area, for example understanding, with a weakness in the other area, i.e. expression, or vice versa.

It helps to find a quiet place without distractions, to make eye contact and get the person's attention. Speak slowly and clearly, using simple phrases and leave plenty of time between sentences to allow for the extra processing time it takes with aphasia. Make it obvious when changing the subject.

Closed questions requiring yes or no answers are easier. Write down key words or headings that both the patient and clinician can refer to. This helps because the auditory memory needed to 'hold on to' the spoken word taxes patients language system. Use pictures and have pen and paper at hand if the patient can use them.

Much can be learnt from carers or watching speech and language therapists – http://www.ukconnect.org/index.aspx.

Practical Box 1.5 Communicating with people who are deaf or hard of hearing

- Ask if they need to lip-read you
- Position yourself on the better hearing side
- Smile and use eye contact
- Put your face in a good light
- Do not cover your face or mouth
- Use plain language
- Speak clearly but not too slowly
- Don't shout
- If stuck, write it down
- Check for understanding
- Never say 'forget it'

www.rnid.org.uk/information_resources/factsheets

NEW INFLUENCES ON COMMUNICATION

The internet

The internet has revolutionized ready access to information and is a fast growing source of health information and support for patients. A recent UK survey found 30% of 45 year olds and over had searched for health information on the web. Research on its impact on the doctor–patient relationship is emerging.

Some clinicians welcome it whilst others fear longer consultations, demands which can't be met, commercially biased information and patients knowing more than them.

Public trust in information is essential and far from undermining the doctor–patient relationship it can enhance communication. Quality control organizations exist such as Health on the Net Foundation (www.HON.ch), Royal College of Physicians (http://www.rcplondon.ac.uk/patientcarer.asp) and the National Patient Safety Association (http://www.npsa.nhs.uk/pleaseask/).

The web also has sites where patients share experiences of illness and treatment. These provide support and again enable patients to bring informed questions to the consultation (www.dipex.org.uk, www.patientsvoices.org.uk).

Decision aids

Decision aids written in non-technical language, and which are evidence based, are available to help patients weigh up risks and benefits on healthcare choices. They are reliable and from independent sources. There are currently over 400 listed on the Cochrane register (www.ohri.ca/decisionaid). The numbers are growing and many are available on the internet.

Studies show they do not increase patients' anxiety and their use results in 21–44% reductions in choice of invasive surgical options over more conservative treatments without adverse effects on health.

Communication and patient safety

Clear handovers and accurate signed and dated, legible records are the responsibility of everyone in the healthcare team. They are vital in providing best care, reducing error and ensuring patient safety.

Clinical records

The content of records is listed in Table 1.2 and essential criteria listed in Table 1.3. In addition to documenting an

Table 1.2 Records should include

Relevant clinical and psychosocial information – history and examination
Relevant findings, both positive and negative
Diagnosis including uncertainties
Investigations arranged
Test results
Decisions made
Information given to patients
Consent
Drugs or other treatments prescribed
Follow-up and referrals

Table 1.3 Criteria for good records

Clear, accurate, legible and contemporaneous
Dated and signed with printed name
Written in pen
First hand
Original – never altered – but an additional note, signed and dated, should be made alongside any mistake
Kept secure

GMC Good Medical Practice 2006; MPS Medical Records 2006. www.mps.org.uk/gpregistrar

individual patient's care they are valuable as part of audit aimed at improving standards of healthcare.

Patients are also entitled to see their records in the UK and they provide essential information when a complaint or claim for negligence is made.

Computer records are increasingly replacing written records and research suggests they are more understandable. They include more information and overcome problems of legibility. A recent study found prescription errors are reduced by a striking 66% when electronic systems replace handwritten prescriptions.

E-mails and texts need to be lodged in medical records as part of the record and care needs to be taken with confidentiality.

With adequate data protection the internet holds immense potential for unifying patient record systems and allowing access electronically by members of the healthcare team in primary, secondary and tertiary care sectors.

Team communication

Effective team communication is essential if patients are to get the best care. This is never more important than when a patient is critically ill and patient safety is vitally dependant on teamwork. Lessons from industries such as aviation and energy have been useful in understanding how to reduce error from communication. Frameworks such as the SBAR (Situation-Background-Assessment-Recommendation) technique provide an easy and focussed way to set expectations for what will be communicated and how between members of the team (www.ihi.org/IHI/Topics/PatientSafety/SafetyGeneral/Tools/).

Training in communication skills

There is now evidence that communication ability is not necessarily innate and that it can be improved and subsequently used in everyday practice. However it cannot be learned from books. As with learning any skill the opportunity to practise and receive constructive feedback on performance is essential together with feedback.

GENERAL READING

Di Blasi Z, Harkness E, Ernst E et al. Influence of context effects on health outcomes: a systematic review. *Lancet* 2001; **357**: 757–762.

Healthcare Commission Spotlight on Complaints (UK) 2007. www.healthcarecommission.org.uk/_db/_documents/spotlight_on_complaints.pdf

SIGNIFICANT WEBSITES

http://www.bma.org.uk/ethics
British Medical Association ethics

http://www.each.nl/
European Association for Communication in Healthcare

http://www.gmc-uk.org/
General Medical Council

http://www.pickereurope.org
Picker Institute Europe

http://www.wma.net/e/policy
World Medical Association policy

http://www.hkma.org/index.htm
Hong Kong Medical Association

The medical interview
Patients with impaired faculties for communication
New influences on communication

FURTHER READING

Broom A. Virtually he@lthy: the impact of internet use on disease experience and the doctor–patient relationship. *Qualitative Health Research* 2005; **15**: 325–345. http://204.187.39.28/decaids.html

Fallowfield L, Jenkins V, Farewell V et al. Enduring impact of communication skills training: results of a 12-month follow-up. *British Journal of Cancer* 2003; **89**: 1445–1449.

McMullan M. Patients using the Internet to obtain health information: how this affects the patient–health professional relationship. *Patient Education and Counselling* 2006; **63**(1–2): 24–28.

O'Connor AM, Stacey D, Entwhistle V et al. *Decision Aids for People Facing Health Treatment or Screening Decisions.* (Cochrane Review) The Cochrane Library (4). Chichester, UK, 2005.

Sood R, Adkoli BV. Medical education in India – problems and prospects. *Journal, Indian Academy of Clinical Medicine* 2000; **1**(3): 210–212.

This page intentionally left blank

2 Molecular cell biology and genetic disorders

THE CELL

The cell membrane 19
Cytoplasm 22
Secondary messengers 24
The cytoskeleton 24
Intercellular connections 26
Cell adhesion molecules 26
The nucleus and its responses 27

MOLECULAR BIOLOGY

DNA structure and function 29
Genes 29
Transcription and translation 29
The control of gene expression 30
Chromosomes, introns and the size of the human genome 33
Tools for molecular biology 33
The biology of chromosomes 36
Human chromosomes 36
The X chromosome and inactivation 37
Telomeres and immortality 37
The mitochondrial chromosome 37

GENETIC DISORDERS

Chromosomal disorders 38
Gene defects 40
Population genetics 44
Clinical genetics and genetic counselling 44
Genetic anticipation 45
Prenatal diagnosis 45
Applications of molecular genetics 46
Gene therapy 46
The genetic basis of cancer 49
Oncogenes 49
Tumour suppressor genes 49
Cancer aetiology: inheritance or environment? 51

THE CELL

The cell (Fig. 2.1) is a highly organized structure and consists of various common organelles held by an adaptive internal scaffolding (the cytoskeleton), which radiates from the nuclear membrane to the cell plasma membrane. The number and finer detail of these common organelles varies according to the specialized function a cell might perform. For example, the mitochondria of muscle cells contain many infolded cristae, producing large amounts of ATP (from the electron transport chain that sits on the inner membrane), whereas the mitochondria of liver cells are smaller and rounded as they produce low levels of ATP. However, many products from the inner matrix feed other metabolic pathways. Each cell is a component unit but 'talks' to adjacent cells via specific channels and receptors. Within a cell there is a constant flow of traffic between the organelles.

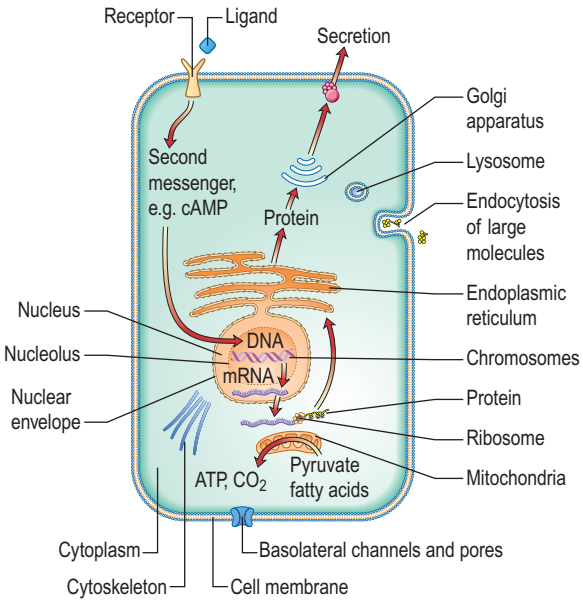


Fig. 2.1 **Diagrammatic representation of the cell**, showing the major organelles and the sites of receptor activation (see Fig. 2.3), intracellular messengers, protein formation and secretion, endocytosis of large molecules and production of ATP (see also Fig. 2.4). For basolateral channels and pores, see Figure 2.2.

THE CELL MEMBRANE

- The *plasma cell membrane* is a bilayer of amphipathic phospholipids that consist of a polar hydrophilic head (e.g. phosphatidyl choline) and an insoluble non-polar lipid hydrophobic tail (commonly two long-chain fatty acids). The phospholipids spontaneously form bilayers (as complete circular structures) that form an effective barrier that is impermeable to most water-soluble molecules. This barrier defines the interior environment of the cell. The exchanges across the plasma membrane

are regulated by various proteins that are embedded in the lipid bilayer (Fig. 2.2). The lipids' hydrophobic structure means that relatively weak bonds hold the plasma membrane together. However, this strongly opposes the transverse movement of hydrophilic molecules but allows considerable freedom for lateral 'fluid' movement by molecules such as membrane proteins that are embedded in it. The plasma membrane is thus a very dynamic structure. Cell-to-cell communication is the key to many diseases and chemotherapeutic interventions.

- The *membrane proteins* embedded in the lipid bilayer either traverse the whole membrane or are associated only with the outer or inner leaflet of the bilayer (Fig. 2.2). These proteins are responsible for the cell's interaction with its external environment. The molecules involved are either *channels*, *receptors* or *cell adhesion molecules* (see below).

Cell dynamics

Cell components are continually being formed and degraded, and most of the degradation steps involve ATP-dependent multienzyme complexes. Old cellular proteins are mopped up by a small cofactor molecule called '*ubiquitin*', which interacts with these worn proteins via their exposed hydrophobic residues. *Ubiquitin* is a small 8.5 kDa regulating protein present universally in all living cells. Cells mark the destruction of a protein by attaching molecules to the protein. This 'ubiquitination' signals the protein to move to lysosomes or proteosomes for destruction. A complex containing more than five ubiquitin molecules is rapidly degraded by a large proteolytic multienzyme array termed '26S proteosome'. Ubiquitin also plays a role in regulation of the receptor tyrosine kinase in the cell cycle and in repair of DNA damage. The failure to remove worn proteins can result in the development of chronic debilitating disorders. For example, Alzheimer and frontotemporal dementias are associated with the accumulation of ubiquinated proteins (prion-like proteins), which are resistant to ubiquitin-mediated proteolysis. Similar proteolytic-resistant ubiquinated proteins give rise to inclusion bodies found in myositis and myopathies. This resistance can be due to point mutation in the target protein itself (e.g. mutant *p53* in cancer; see p. 49) or as a result of an external factor altering the conformation of the normal protein to create a proteolytic-resistant shape, as in the prion protein of variant Creutzfeldt–Jakob disease (vCJD). Other conditions include von Hippel–Lindau syndrome (p. 666) and Liddle's syndrome (p. 666).

Free radicals

A free radical is any atom or molecule which contains one or more unpaired electrons, making it more reactive than the native species. The major free radical species produced in the human body are the hydroxyl radical (OH), the superoxide radical (O_2^-) and nitric oxide (NO).

Free radicals have been implicated in a large number of human diseases. The hydroxyl radical is by far the most reactive species but the others can generate more reactive species as breakdown products. When a free radical reacts with a non-radical a chain reaction ensues which results in direct tissue damage by membrane lipid peroxidation. Furthermore hydroxyl radicals can cause genetic mutations by attacking purines and pyrimidines. Superoxide dismutases (SOD) convert superoxide to hydrogen peroxide and are thus part of an inherent protective antioxidant mechanism. Patients with dominant familial forms of amyotrophic lateral sclerosis (motor neurone disease) have mutations in the gene for Cu–Zn SOD-1 catalases. Glutathione peroxidases are enzymes that remove hydrogen peroxide generated by SOD in the cell cytosol and mitochondria.

Free radical scavengers bind reactive oxygen species. Alpha-tocopherol, urate, ascorbate and glutathione remove free radicals by reacting directly and non-catalytically. Severe deficiency of α-tocopherol (vitamin E deficiency) causes neurodegeneration. There is evidence that cardiovascular disease and cancer can be prevented by a diet rich in substances that diminish oxidative damage (p. 224). The principal dietary antioxidants are vitamin E, vitamin C, β-carotene and flavanoids.

Heat shock proteins

The heat shock response is a highly conserved and ancient response to tissue stress (chemical and physical) that is mediated by activation of specific genes leading to the production of specific heat shock proteins (HSPs). The diverse functions of HSPs include the transport of proteins in and out of specific cell organelles, acting as molecular chaper-

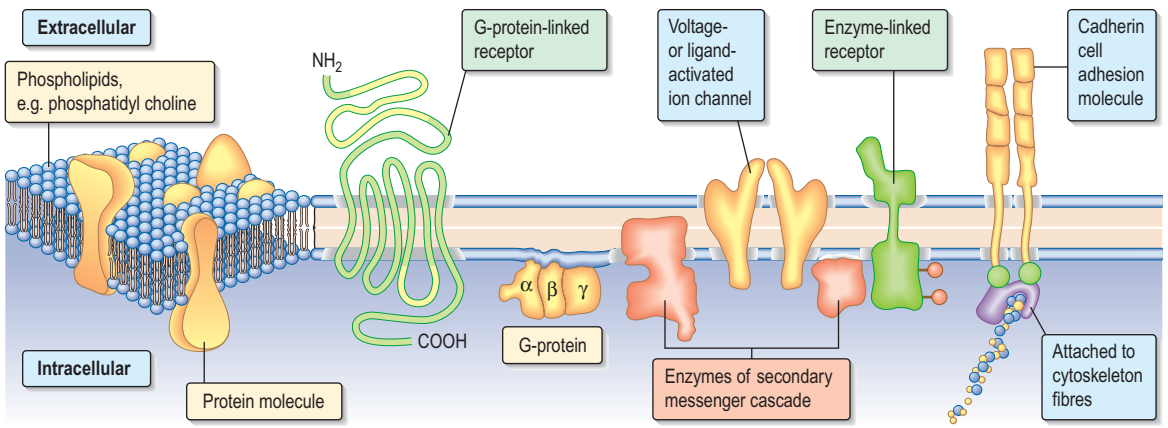


Fig. 2.2 **Cell membrane showing lipid structures and main types of integral proteins** such as receptors, G-proteins, channels, secondary messenger enzyme complexes and cell adhesion molecules.

ones (the catalysis of protein folding and unfolding) and the degradation of proteins (often by ubiquitination pathways). As well as heat, cytotoxic chemicals and free radicals can trigger HSP expression. The unifying feature, which leads to the activation of HSPs, is the accumulation of damaged intracellular protein. Tumours have an abnormal thermotolerance, which is the basis for the observation of the enhanced cytotoxic effect of chemotherapeutic agents in hyperthermic subjects. The HSPs are expressed in a wide range of human cancers and have been implicated in tumour cell proliferation, differentiation, invasion, metastasis, cell death and immune response.

Phagocytosis, pinocytosis and exocytosis

- *Phagocytosis*. Specialized cells such as macrophages and neutrophils can engulf about 20% of their surface area in the pursuit of large particles such as bacteria. Lysosomes rapidly fuse with phagosomes, giving equally rapid digestion of the contents, and recycling as much of the internalized membrane as possible. Phagocytosis is only triggered when specific cell surface receptors – such as the macrophage Fc receptor – are occupied by their ligand.
- *Pinocytosis* is a much smaller-scale model of phagocytosis and is continually occurring in all cells. In contrast to phagocytosis, receptors for smaller molecular complexes, such as low-density lipoprotein (LDL) (Fig. 2.3) result in surface clumping and the internal accumulation of a protein called *clathrin*. Clathrin-coated pits pinch inwards as clathrin-coated vesicles. Clathrin prevents fusion of lysosomes, and thus its removal will result in lysosomal fusion and degradation of the contents.
- *Exocytosis*. Maintenance of a clathrin coat can result in transcellular transit of the contents and their exocytosis at another side of the plasma membrane, i.e. apical to basal surface transcytosis. Similarly, cell organelles bud off vesicles coated in clathrin to prevent lysosomal fusion and degradation. Some of these vesicles rapidly fuse with the plasma membrane and exocytose their contents. Other vesicles do not immediately fuse with the plasma membrane (or indeed any other organelle). The clathrin-coated vesicles have additional lipid bilayer-embedded proteins called v-SNAREs (signal and response elements), which interact with target organelle membrane proteins called t-SNAREs. Vesicle fusion is therefore specific, comprising fusion in the correct place (at a particular organelle or part of the plasma membrane) and at the correct time (e.g. the fusion of neuronal transmitter vesicles and release of the transmitter at the synaptic membrane when stimulated).

Defects in every molecular process associated with phago/pinocytosis can lead to disease. For example, hypercholesterolaemia can arise not only because mutations render the LDL receptor unable to bind LDL but also due to failures of secondary proteins to trigger clathrin coating of the vesicle.

Membrane transport and ion channels

The plasma membrane is freely permeable to gases such as O_2, CO_2 and N_2, and to small uncharged molecules such as H_2O (not H^+ and OH^-) and urea. Whilst larger hydrophobic

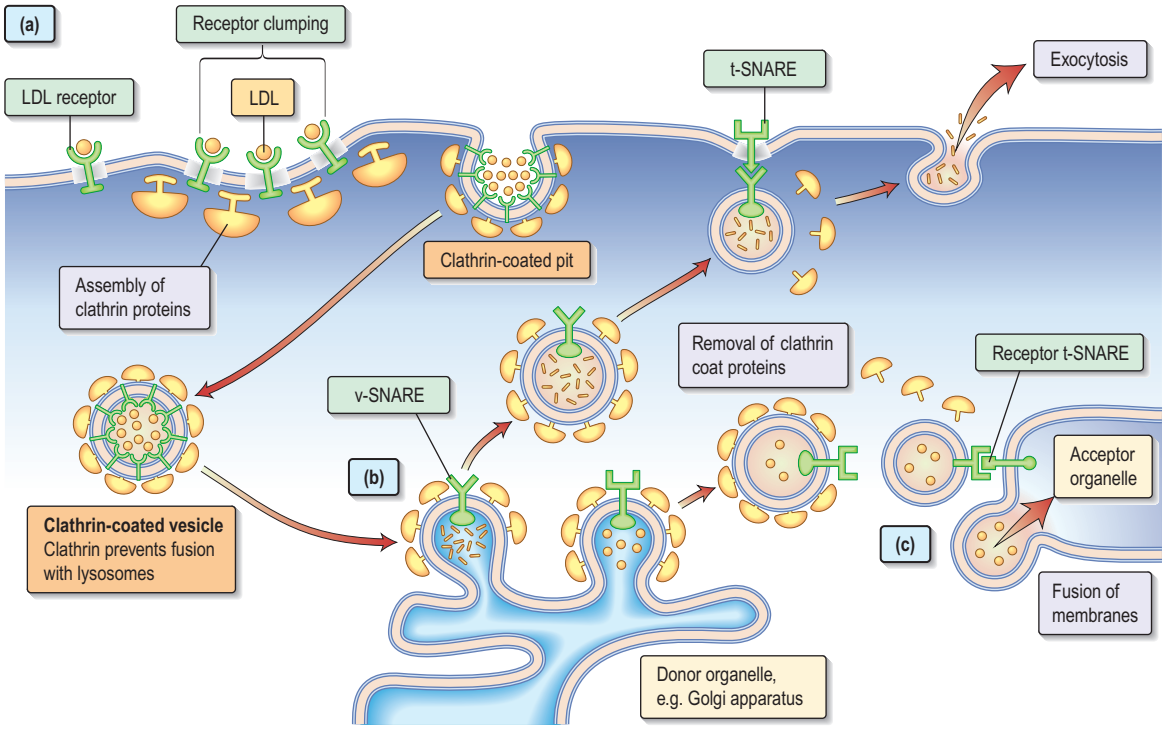


Fig. 2.3 **Intracellular transport. (a)** Receptor-mediated pinocytosis. **(b)** Trafficking of vesicles containing synthesized proteins to the cell surface (e.g. hormones). **(c)** Traffic between organelles is also mediated by v- and t-SNARE-containing organelles. v-SNARE, vesicle-specific SNARE; t-SNARE, target-specific SNARE.

lipid-soluble molecules – like steroids – also pass freely through the membrane, large uncharged molecules (glucose, amino acids and nucleotides) and small charged ions (K^+, Na^+, Ca^{2+}, Cl^-, Mg^{2+} and HCO_3^-) cannot pass unless via a specific transport protein embedded in the plasma membrane. Two structural types of transport molecules/complexes exist (Fig. 19.2):

- *Channel proteins* literally open a channel in the lipid membrane to allow a specific solute to pass through.
- *Carrier proteins* are slower in action, shuttling the solute across and either facilitating diffusion down a gradient across the membrane, or actively pumping solutes against the gradient using ATP as an energy source.

Active carrier pumps and gated ion channels work together in neural transmission (see Fig. 21.1). These carrier proteins pump Na^+ and K^+ across the neuronal cell membrane to create a differential gradient, but ion channels open in response to stimuli to cause a rapid depolarization, allowing the ions to flow back. At synaptic junctions these ion channels open in response to chemical signals such as the release of glutamate, epinephrine (adrenaline) or acetylcholine.

ATP-dependent transport molecules (ATPases) belong to a superfamily called the '*ABC transporter superfamily*'. These include the multidrug-resistance protein (MRP), which pumps out hydrophobic drugs and is overexpressed by tumour cells, and the chloride ion pump coded by the cystic fibrosis gene (see Fig. 2.19). All share a common structure of repeating six transmembrane domains interrupted by a cytoplasmic ATPase domain. The cystic fibrosis chloride ion channel is unusual in that it requires the binding and hydrolysis of both ATP and cAMP for activation.

Receptors

Membrane surface receptors pass their extracellular signal across the plasma membrane to cytoplasmic secondary signalling molecules. Examples of these receptors include the non-lipid-soluble ligands such as growth hormone (GH), insulin, insulin-like growth factor (IGF) and luteinizing hormone (LH). These membrane-bound receptors can be subclassified according to the mechanism by which they activate signalling molecules:

- ion channel linked (see above)
- G-protein linked
- enzyme linked.

Structurally these plasma membrane receptors can be:

- *serpentine* (seven transmembrane domains, e.g. the LH receptor)
- *transmembrane with large extra- and intracellular domains* (e.g. the epidermal growth factor (EGF) receptor)
- *transmembrane with a large extracellular domain only* (the macrophage scavenging receptors)
- *entirely linked onto the outer membrane* leaflet by a lipid moiety known as a GPI (glycan phosphatidylinositol) anchor (e.g. T cell receptor).

The function of these membrane receptors is to initiate a secondary message that ultimately results in activation of a specific enzyme or a DNA-binding protein. This may involve translocation to the nucleus and initiation of transcription of a specific set of genes.

G-protein-linked receptors

The G-protein-linked receptor, once activated by a ligand, binds a trimeric complex (α, β, γ) which is anchored to the inner surface of the plasma membrane. This complex is a GTP-binding protein, or G-protein. The G-protein binds GTP rather than GDP, and then interacts with enzyme complexes anchored into the inner leaflet of the membrane. These complexes in turn activate one or all three of the secondary messengers:

- cyclic AMP (cAMP)
- Ca^{2+} ions
- inositol 1,4,5-trisphosphate/diacylglycerol (IP_3/DAG).

Enzyme-linked surface receptors

These receptors usually have a single transmembrane-spanning region, and a cytoplasmic domain that has intrinsic enzyme activity or will bind and activate other membrane-bound or cytoplasmic enzyme complexes. Four classes of enzymes have been designated:

- *Guanylyl cyclase-linked receptors* (e.g. the atrial natriuretic peptide receptor), which produce cyclic GMP. This in turn activates a cGMP-dependent kinase (G-kinase), which binds to and phosphorylates serine and threonine residues of specific secondary messengers.
- *Tyrosine kinase receptors* (e.g. the platelet-derived growth factor (PDGF) receptor), which either specifically phosphorylate kinases on a small set of intracellular signalling proteins, or associate with proteins that have tyrosine kinase activity.
- *Tyrosine phosphatase receptors* (e.g. CD45), which remove phosphates from tyrosine residues of specific intracellular signalling proteins.
- *Serine/threonine kinase receptors* (e.g. the transforming growth factor-beta (TGF-β) receptor), which phosphorylate specific serine and threonine residues of intracellular signalling proteins.

There are many intracellular receptors that bind lipid-soluble ligands such as steroid hormones (e.g. progesterone, cortisol) and T_3 and T_4. These cytoplasmic receptors often change shape in response to binding their ligands, form dimers, enter the nucleus and interact directly with specific DNA sequences (see DNA-binding proteins, p. 30).

CYTOPLASM

This is the fluid component inside the cell membrane and contains many specialized organelles (see Fig. 2.1). It contains a scaffolding or cytoskeleton that regulates the passage and direction in which the interior solutes and storage granules flow. The cytoplasm contains:

- *Endoplasmic reticulum (ER)*. This consists of interconnecting tubules or flattened sacs (cisternae) of lipid bilayer membrane. It may contain ribosomes on the surface (termed rough endoplasmic reticulum (RER) when present, or smooth endoplasmic reticulum (SER) when absent). The ER is involved in the processing of proteins: the ribosomes translate mRNA into a primary sequence of amino acids of a protein peptide chain (see Fig. 2.4). This chain is synthesized into the ER where it

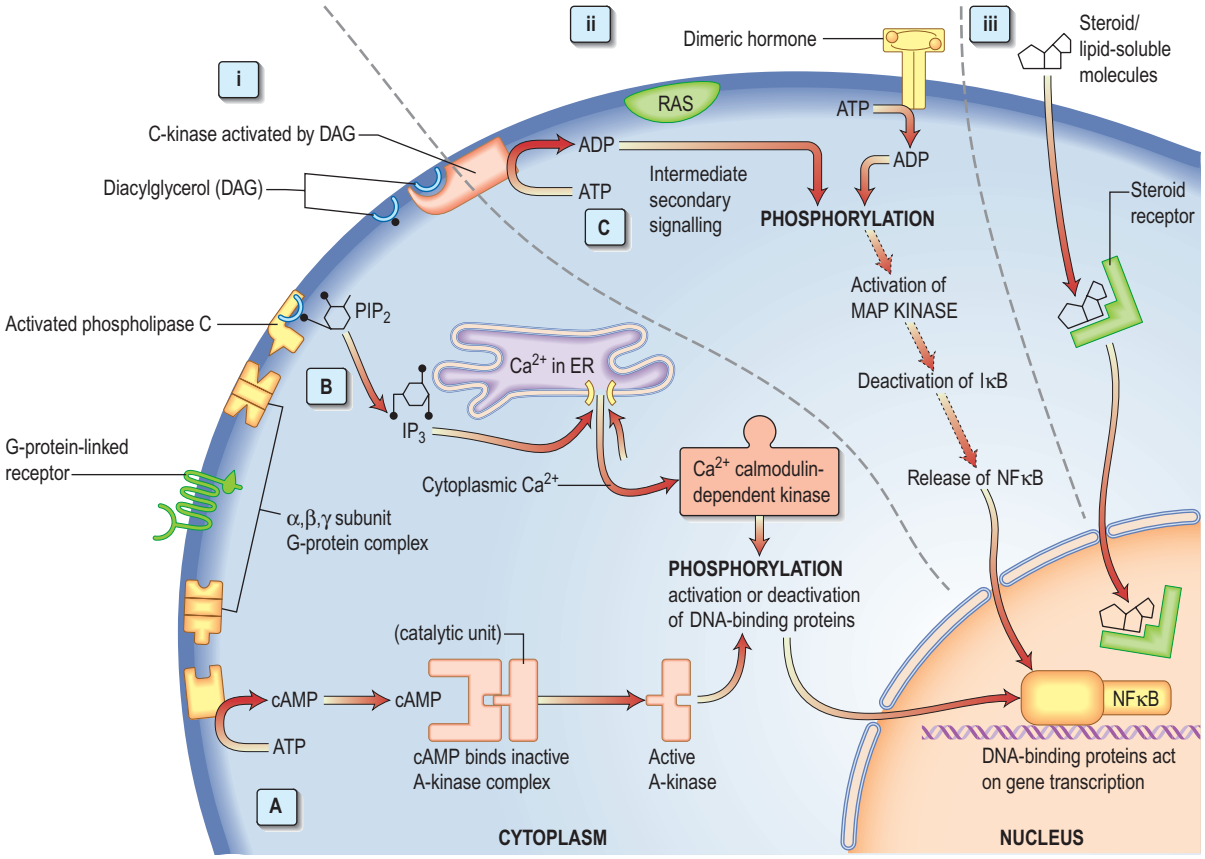


Fig. 2.4 **Receptor and secondary messengers. (i) G-protein receptor** binds ligand (e.g. hormone) and activates G-protein complex. The G-protein complex can activate three different secondary messengers: (a) cAMP generation; (b) inositol 1,4,5-trisphosphate (IP3) and release of Ca^{2+}; (c) diacylglycerol (DAG) activation of C-kinase and subsequent protein phosphorylation. **(ii) Dimeric hormone** binds receptor subunits bringing them into close association. Intracellular domains cross-phosphorylate and link to the phosphorylation cascades via molecules such as RAS (superfamily of GTPases). **(iii) Lipid-soluble molecules**, e.g. steroids, pass through the cell membrane and bind to cytoplasmic receptors, which enter the nucleus and bind directly to DNA. ER, endoplasmic reticulum; IκB, inhibitory factor kappa B; NFκB, nuclear factor kappa B.

is first folded and modified into mature peptides. The ER is the major site of drug metabolism.

- *Golgi apparatus*. This consists of flattened cisternae similar to the ER. It is characterized as a stack of cisternae from which vesicles bud off from the thickened ends. The primary processed peptides of the ER are exported to the Golgi apparatus for maturation into functional proteins (e.g. glycosylation of proteins which are to be excreted occurs here) before packaging into secretory granules and cellular vesicles that bud off the ends (see Fig. 2.3).
- *Lysosomes*. These are dense cellular vesicles containing acidic digestive enzymes. They fuse with phagocytotic vesicles from the outer cell membrane, digesting the contents into small biomolecules that can cross the lysosomal lipid bilayer into the cell cytoplasm. Lysosomal enzymes can also be released outside the cell by fusion of the lysosome with the plasma membrane. Lysosomal action is crucial to the function of macrophages and polymorphs in killing and digesting infective agents, tissue remodelling during development and osteoclast remodelling of bone. Not surprisingly, *many metabolic disorders* result from impaired lysosomal function (p. 1071).
- *Peroxisomes*. These are dense cellular vesicles so named because they contain enzymes that catalyse the breakdown of hydrogen peroxide. They are involved in the metabolism of bile and fatty acids, and are primarily concerned with detoxification, e.g. D-amino acid oxidase and H_2O_2 catalase. The inability of the peroxisomes to function correctly can lead to rare metabolic disorders such as *Zellweger's syndrome* and *rhizomelic dwarfism*.
- *Mitochondria*. These organelles are the powerhouse of the cell. Each mitochondrion comprises two lipid bilayer membranes and a central matrix. It also possesses several copies of its own DNA in a circular genome. The outer membrane contains many gated receptors responsible for the import of raw materials like pyruvate and ADP, and the export of products such as oxaloacetate (precursor of amino acids and sugars) and ATP. An interesting caveat to our symbiotic relationship is that proteins of the *Bcl-2/Bax* family are incorporated in this outer membrane and can release mitochondrial enzymes that trigger apoptosis. The *inner membrane* is often highly infolded to form cristae to increase its effective surface area. It contains transmembrane enzyme complexes of the electron transport chain, which generate an H^+ ion gradient. This gradient then drives the adjacent transmembrane ATPase complex to form ATP from ADP and P_i. The *inner matrix* contains the enzymes of the Krebs cycle that generate the substrates of both the electron transport chain ($FADH_2$

and NADH) and central metabolism (e.g. succinyl CoA, α-oxoglutarate, oxaloacetate).

SECONDARY MESSENGERS

Secondary messengers are molecules that transduce a signal from a bound receptor to its site of action (e.g. the nucleus). There are essentially four mechanisms by which secondary messengers act but they cross talk and are rarely activated independently of each other (Fig. 2.4). These mechanisms are cyclic AMP, IP_3/DAG, Ca^{2+} ions and protein phosphorylation.

Cyclic AMP, IP_3/DAG and Ca^{2+} ions

- The generation of cAMP by G-protein-linked receptors results in an increase in cellular cAMP (Fig. 2.4(ia)), which binds and activates specific cAMP-binding proteins. These dimerize and enter the cell nucleus to interact with set DNA sequences (the cAMP response elements). In addition, cofactors in the cAMP response element binding proteins (CREB) are co-activated and interact with the phosphorylation pathway.
- Other G-protein complexes activate inner membrane-bound phospholipase complexes. These in turn cleave membrane phospholipid-polyphosphoinositide (PIP_2) into two components (Fig. 2.4(ib)). The first is the water-soluble molecule inositol trisphosphate, IP_3. This floats off into the cytoplasm and interacts with gated ion channels in the endoplasmic reticulum (or sarcoplasmic reticulum in muscle cells), causing a rapid release of Ca^{2+}. The lipid-soluble component diacylglycerol (DAG) (Fig. 2.4(ic)) remains at the membrane, but activates a serine/threonine kinase, protein kinase C (see phosphorylation section below).
- The cellular calcium-binding proteins and ion pumps rapidly remove Ca^{2+} from the cytoplasm back into a storage compartment (such as the endoplasmic reticulum). Free Ca^{2+} interacts with target proteins in the cytoplasm, inducing a phosphorylation/dephosphorylation cascade, resulting in activated DNA-binding proteins entering the nucleus.

Protein phosphorylation

Although phosphorylation of the cytoplasmic secondary messengers is often a consequence of secondary activation of cAMP, Ca^{2+} and DAG, the principal route for the protein phosphorylation cascades is from the dimerization of surface protein kinase receptors, which have bound their ligands. The tyrosine kinase receptors phosphorylate each other when ligand binding brings the intracellular receptor components into close proximity (see Fig. 2.4(ii)). The inner membrane and cytoplasmic targets of these activated receptor complexes are *ras*, protein kinase C and ultimately the MAP (mitogen activated protein) kinase, Janus-Stat pathways or phosphorylation of IκB causing it to release its DNA-binding protein, nuclear factor kappa B (NFκB). These intracellular signalling proteins usually contain conserved non-catalytic regions called SH2 and SH3 (SRC homology regions 2 and 3). The SH2 region binds to phosphorylated tyrosine. The SH3 domain has been implicated in the recruitment of intermediates that activate *ras* proteins. Like G-proteins, *ras* (and its homologous family members *rho* and *rac*) switch between an inactive GDP-binding state and an active GTP-binding state. NFκB undergoes a conformational change, enters the nucleus and initiates transcription of specific genes.

Lipid-soluble ligands (e.g. steroids; Fig. 2.4(iii)) do not need secondary messengers; their cytoplasmic receptors, once activated, enter the nucleus as DNA-binding proteins and alter gene expression directly.

THE CYTOSKELETON

This is a complex network of structural proteins which regulates not only the shape of the cell, but also its ability to traffic internal cell organelles and even move in response to external stimuli (see Fig. 2.2). The major components are microtubules, intermediate filaments and microfilaments.

- *Microtubules* (Fig. 2.5). These are made up of two protein subunits, α and β tubulin (50 kDa), and are continuously changing length. They form a 'highway', transporting organelles through the cytoplasm. There are two motor microtubule-associated proteins – dynein and kinesin – allowing antegrade and retrograde movement. Dynein is also responsible for the beating of cilia. During interphase the microtubules are rearranged by the microtubule-organizing centre (MTOC), which consists of centrosomes containing tubulin and provides a structure on which the daughter chromosomes can separate. Another protein involved in the binding of organelles to microtubules is the cytoplasmic linker protein (CLIP). Drugs that disrupt the microtubule assembly (e.g. colchicine and vinblastine) affect the positioning and morphology of the organelles. The anticancer drug paclitaxel causes cell death by binding

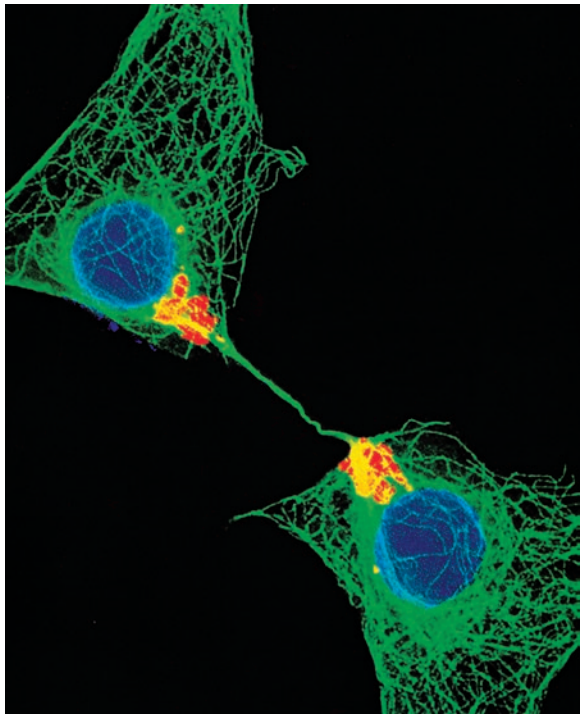


Fig. 2.5 **Immunofluorescent micrograph of dividing fibroblasts showing the cytoskeleton.** Microtubules are shown in green, Golgi apparatus in yellow and the nuclei in blue. Reproduced with permission of Dr Philip Huie, Stanford University, from *Biotechniques* July/August 1995.

to microtubules and stabilizing them so that organelles cannot move, and thus mitotic spindles cannot form.

- *Intermediate filaments*. These form a network around the nucleus and extend to the periphery of the cell. They make cell-to-cell contacts with the adjacent cells via desmosomes, and with basement matrix via hemidesmosomes (Fig. 2.6). Their function appears to be in structural integrity; they are prominent in cellular tissues under stress. The intermediate filament fibre proteins are specific to the embryonic lineage of the cell concerned, for example keratin intermediate fibres are only found in epithelial cells whilst vimentin is only found in mesothelial (fibroblastic) cells.
- *Microfilaments*. Muscle cells contain a highly ordered structure of actin (a globular protein, 42–44 kDa) and myosin filaments, which form the contractile system. These filaments are also present throughout the non-muscle cells as truncated myosins (e.g. myosin 1), in the cytosol (forming a contractile actomyosin gel), and beneath the plasma membrane. Cell movement is mediated by the anchorage of actin filaments to the plasma membrane at adherent junctions between cells (Fig. 2.6). This allows a non-stressed coordination of contraction between adjacent cells of a tissue. Similarly, vertical contraction of tissues is anchored across the cell membrane to the basement matrix at focal adhesion junctions where actin fibres converge (Fig. 2.6). Actin-binding proteins (e.g. fimbria) modulate the behaviour of microfilaments, and their effects are often calcium-dependent. The actin-associated proteins can be tissue type specific, for example actin-binding troponin is a complex of three subunits and two of these have

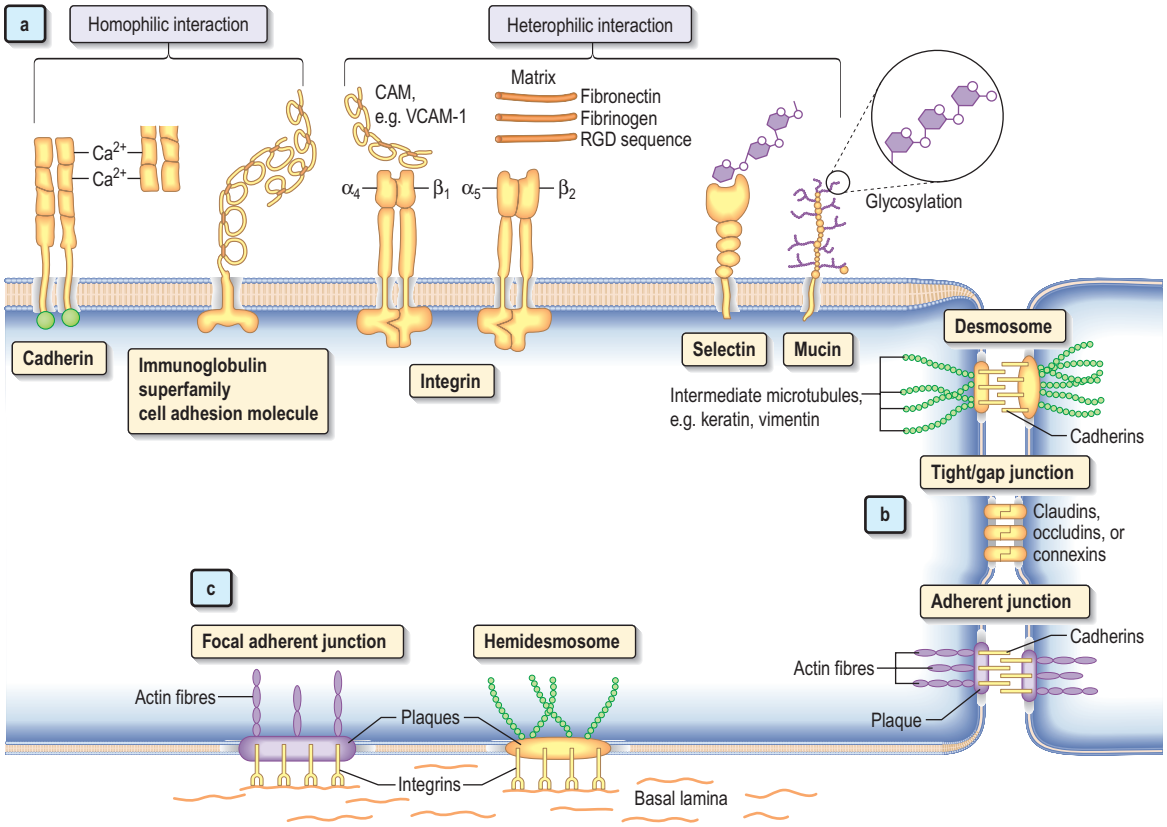


Fig. 2.6 **Cell adhesion molecules and cellular junctions.**

(a) Five main groups of adhesion molecules. The ***cadherins*** have a Ca^{2+}-dependent homodimer homophilic interaction and their intracellular domains link to the cytoskeleton. ***Immunoglobulin*** superfamily cell adhesion molecules (CAMs) have both homophilic and heterophilic (integrin) interactions. Binding is Ca^{2+} independent and the intracellular domains are cell signalling. ***Integrins*** have heterophilic interactions mostly with basement matrix components. They have an α–β chain structure and their intracellular domains are predominantly cell signalling but can directly interact with cytoskeletal complexes. ***Selectins*** have a weak binding affinity for specific sugar molecules found on mucins. ***Mucins*** are long tandem-repeat peptides which protrude from the cell membrane and are covered in glycosylation moieties.

(b) Adjacent cells form focal adhesion junctions. ***Desmosomes*** are where the membrane forms a proteinaceous plaque containing molecules like desmoglein, from which cadherins protrude and bind cadherins of the adjacent cell. Intracellularly the plaque binds loops of cytoskeleton intermediate filaments, e.g. keratin in epithelial cells and vimentin in fibroblasts. ***Tight junctions*** are mediated by integral membrane proteins, claudins and diclaudins, which associate to form subunits bridging the intercellular gap. ***Gap junctions*** consist of connexin subunits that form a regulated hollow tube. ***Adherent junctions*** are similar to desmosomes in that cell-to-cell adhesion is mediated by cadherins but the membrane proteinaceous plaque components are different and bind contractile cytoskeletal fibres like actin at the terminus.

(c) Basement membrane adhesion. This is similar to desmosomes and adherent junctions in that membrane plaques link intercellular intermediate filaments (e.g. keratin or vimentin) in ***hemidesmosomes*** and contractile cytoskeleton actin in ***focal adherent junctions*** to the basement matrix. However, integrins replace cadherins as the surface adhesion molecules.

RGD, Arg–Gly–Asp adhesion sequence.

isomers which are only found in cardiac muscle. Cardiac troponin I and T are released into the blood circulation after the onset of a myocardial infarction (p. 752).

Alterations in the cell's actin architecture are also controlled by the activation of small *ras*-like GTP-binding proteins *rho* and *rac*. These are involved in rearrangement of the cell during division, and thus dysfunctions of these proteins are associated with malignancy.

INTERCELLULAR CONNECTIONS

The cytoskeleton and plasma membrane interconnect, and extracellular domains form junctions between cells to form tissues. There are three types of junction between cells: tight junctions, adherent junctions and gap junctions (Fig. 2.6).

Tight junctions

Tight junctions (zonula occludens) hold cells together. They are situated at the ends of margins adjacent to epithelial cells (e.g. intestinal and renal cells) and form a barrier to the movement of ions and solutes across the epithelium, although they can be variably 'leaky' to certain solutes. The proteins responsible for the intercellular tight junction closure are called *claudins*. They show selective expression within tissue and regulate what small ions may pass through the gaps between cells. For example, the kidney displays a differential expression of these claudin proteins. Mutations of claudin-16 (which is expressed in the thick ascending limb in the loop of Henle in the kidney, where magnesium is reabsorbed) are responsible for some abnormalities of magnesium reabsorption of Gitelman's syndrome (p. 665). Since magnesium reabsorption is paracellular, tight junctions (which contain claudin-16) prevent these divalent ions rapidly diffusing back between the cells into renal tubules.

Adherent junctions

Adherent junctions (zonula adherens) are continuous on the basal side of cells. They contain cadherins and are the major site of attachment of intracellular microfilaments. Intermediate filaments attach to *desmosomes*, which are apposed areas of thickened membranes of two adjacent cells. *Hemidesmosomes* attach cells to the basal lamina and are also connected to intermediate filaments. Transmembrane *integrins* link the extracellular matrix to microfilaments at focal areas where cells also attach to their basal laminae. In blistering dermatological disorders autoantibodies cause damage by attacking tight junction desmosomal proteins such as desmoglein-3 in *pemphigus vulgaris* and desmoglein-1 in *pemphigus foliaceus* (p. 1255).

Gap junctions

Gap junctions allow substances to pass directly between cells without entering the extracellular fluids. Protein channels (connexins) are lined up between two adjacent cells and allow the passage of solutes up to molecular weight 1000 kDa (e.g. amino acids and sugars), as well as ions, chemical messengers and other factors. The diameter of these channels is regulated by intracellular Ca^{2+}, pH and voltage. Connexins are made up of six subunits surrounding a channel and their isoforms in tissues are encoded by different genes. Mutant connexins can cause disorders, such as the X-linked form of Charcot–Marie–Tooth disease (p. 1175).

CELL ADHESION MOLECULES

Adhesion molecules and adhesion receptors are essential for tissue structural organization. Differential expression of such molecules is implicit in the processes of cell growth and differentiation, such as wound repair and embryogenesis. There are four major families of cell adhesion molecules.

Cadherins

The cadherins establish molecular links between adjacent cells. They form zipper-like structures at 'adherens junctions', areas of the plasma membrane where cells make contact with other cells. Through these junctions, bundles of actin filaments run from cell to cell. Related molecules such as desmogleins form the main constituents of desmosomes, the intercellular contacts found abundantly between epithelial cells. Desmosomes serve as anchoring sites for intermediate filaments of the cytoskeleton. When dissociated embryonic cells are grown in a dish, they tend to cluster according to their tissue of origin. The homophilic (like with like) interaction of cadherins is the basis of this separation, and has a key role in segregating embryonic tissues. The expression of specific adhesion molecules in the embryo is crucial for the migration of cells and the differentiation of tissues. For example, when neural crest cells stop producing N-CAM (see below) and N-cadherin and start to display integrin receptors they can separate, and begin to migrate on the extracellular matrix. Changes in cadherin expression are often associated with tumour metastatic potential.

Integrins

These are membrane glycoproteins with α and β subunits which exist in active and inactive forms. The integrins principally bind to extracellular matrix components such as fibrinogen, elastase and laminin. The amino acid sequence arginine–glycine–aspartic acid (RGD) is a potent recognition sequence for integrin binding, and integrins replace cadherins in the focal membrane anchorage of hemidesmosomes and focal adhesion junctions (Fig. 2.6). A feature of integrins is that the active form can come about as a result of a cytoplasmic signal that causes a conformational change in the extracellular domain, increasing affinity for its ligand. This 'inside-out' signalling occurs when leucocytes are stimulated by bacterial peptides, rapidly increasing leucocyte integrin affinity for immunoglobulin superfamilies structures such as the Fc portion of immunoglobulin. The 'outside-in' signalling follows the binding of the ligand to the integrin and stimulates secondary signals resulting in diverse events such as endocytosis, proliferation and apoptosis. Defective integrins are associated with many immunological and clotting disorders such as *Bernard–Soulier syndrome* and *Glanzmann's thrombasthenia* (p. 437).

Immunoglobulin superfamily cell adhesion molecules (CAMs)

These molecules contain domain sequences which are immunoglobulin-like structures. The neural cell adhesion molecule (N-CAM) is found predominantly in the nervous system. It mediates a homophilic (like with like) adhesion. When bound to an identical molecule on another cell, N-CAM can also associate laterally with a fibroblast growth factor receptor and stimulate the tyrosine kinase activity of that receptor to induce the growth of neurites. Thus adhesion molecules can trigger cellular responses by indirect activa-

tion of other types of receptors. The placenta and gastrointestinal tract also express immunoglobulin superfamily members, but their function is not completely understood.

Selectins

Unlike most adhesion molecules (which bind to other proteins), the selectins interact with carbohydrate ligands or mucin complexes on leucocytes and endothelial cells (vascular and haematological systems). Selectins were named after the tissues in which they were first identified. *L-selectin* (CD62L) is found on leucocytes and mediates the homing of lymphocytes to lymph nodes. *E-selectin* (CD62E) appears on endothelial cells after they have been activated by inflammatory cytokines; the small basal amount of E-selectin in many vascular beds appears to be necessary for the migration of leucocytes. *P-selectin* (CD26P) is stored in the alpha granules of platelets and the Weibel–Palade bodies of endothelial cells, but it moves rapidly to the plasma membrane upon stimulation of these cells. All three selectins play a part in leucocyte rolling (p. 68).

THE NUCLEUS AND ITS RESPONSES

A nucleus is present in all eukaryotic cells that divide. It contains the human genome and is bound by two bilayer lipid membranes. The outer of the two is continuous with the endoplasmic reticulum. Nuclear pores are present in the membranes, allowing the passage of nucleotides and DNA interacting proteins in, and mRNA out (see Fig. 2.4). The genome consists of DNA and all the apparatus for replication and transcription into RNA (see p. 29). There are two types of cell division – meiosis and mitosis. In *meiosis*, which occurs only in germ cells, the chromosome complement is halved (haploid) and, at fertilization, the union of two cells restores the full complement of 46 chromosomes. *Mitosis* occurs in dividing cells after fertilization, and results in two identical daughter cells. It is only during cell division that chromosomes (see p. 36) become visible.

A nucleolus is a dense area within the nucleus. It is rich in proteins and RNA and is chiefly concerned with the synthesis of ribosomal RNA (rRNA) and ribosomes.

The cell cycle (Fig. 2.7)

Cells in the quiescent G0 phase (G, gap) of the cycle are stimulated by the receptor-mediated actions of growth factors (e.g. EGF, epithelial growth factor; PDGF, platelet-derived growth factor; IGF, insulin-like growth factor) via intracellular second messengers. Stimuli are transmitted to the nucleus (see below) where they activate transcription factors and lead to the initiation of DNA synthesis, followed by mitosis and cell division. Cell cycling is modified by the cyclin family of proteins (see below). Thus from G0 the cell moves on to G1 (gap 1) when the chromosomes are prepared for replication. This is followed by the synthetic (S) phase, when the 46 chromosomes are duplicated into chromatids, followed by another gap phase (G2), which eventually leads to mitosis (M).

Cyclin and cyclin-dependent kinases

As the cell progresses through the cell cycle a series of small serine/threonine kinases are produced. Indeed it is the coordinated cyclic expression of these cyclin-dependent kinases that drives the cell replication cycle. The progression of the cell cycle is catalysed by cyclin-dependent kinases (Cdk) which are activated by a class of proteins called cyclins (Cyc). After stimulation from a pro-mitotic extracellular signal (e.g. a growth factor), **G1** cyclin–Cdk complexes (CycB/Cdk4/6; CycE/Cdk2) become active to prepare the cell for **S phase**, promoting the expression of transcription factors that in turn promote the expression of S cyclins (CycB/Cdk2) and of enzymes required for DNA replication. The G1 cyclin–Cdk complexes also promote the degradation of molecules that function as S phase inhibitors by targeting them for ubiquitination. Active S cyclin–Cdk complexes phosphorylate proteins that make up the pre-replication complexes assembled during **G1 phase** on DNA replication origins. The phosphorylation serves two purposes: to activate each already assembled pre-replication complex, and to prevent new complexes from forming. This ensures that every portion of the cell's genome will be replicated once only. Mitotic cyclin–Cdk complexes (e.g. CycB/CdK2), which are synthesized but inactivated during **G2 phase**, promote the initiation of mitosis by stimulating downstream proteins involved in chromosome condensation and mitotic spindle assembly.

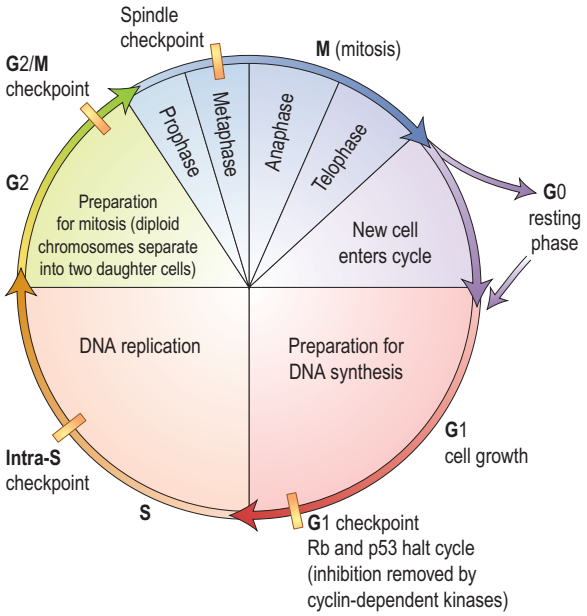


Fig. 2.7 **The cell cycle**. Cells are stimulated to leave non-cycle **G0** to enter **G1** phase by growth factors. During **G1**, transcription of the DNA synthesis molecules occurs. Rb is a 'checkpoint' (inhibition molecule) between G1 and S phases and must be removed for the cycle to continue. This is achieved by the action of the cyclin-dependent kinase produced during G1. During the **S** phase any DNA defects will be detected and p53 will halt the cycle (see p. 50). Following DNA synthesis (S phase), cells enter **G2**, a preparation phase for cell division. Mitosis takes place in the **M** phase. The new daughter cells can now either enter G0 and differentiate into specialized cells, or re-enter the cell cycle.

Apoptosis (programmed cell death)

Apoptosis or physiological cell death occurs through the deliberate activation of constituent genes whose function is to cause their own demise.

Necrotic cell death, in contrast, is where some external factor (e.g. hypoxia, chemical toxins) damages the cell's physiology and results in the disintegration of the cell. Char-

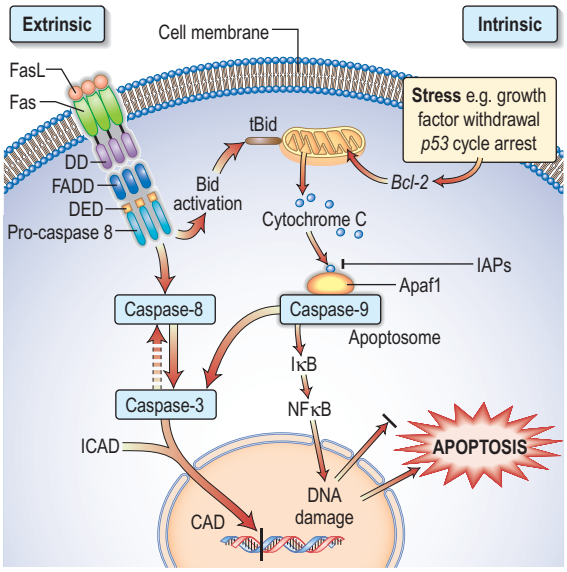


Fig. 2.8 **Extrinsic and intrinsic apoptotic signalling network**. The Fas protein and Fas ligand (FasL) are two proteins that interact to activate an apoptotic pathway.

Fas and **FasL** are both members of the TNF (tumour necrosis factor) family – Fas is part of the transmembrane receptor family and FasL is part of the membrane-associated cytokine family. When the homotrimer of FasL binds to Fas, it causes Fas to trimerize and brings together the **death domains (DD)** on the cytoplasmic tails of the protein. The adaptor protein, **FADD (Fas-associating protein with death domain)**, binds to these activated death domains and they bind to pro-caspase 8 through a set of **death effector domains (DED)**.

When pro-caspase 8s are brought together, they transactivate and cleave themselves to release **caspase 8**, a protease that cleaves protein chains after aspartic acid residues. Caspase 8 then cleaves and activates other caspases which eventually leads to activation of caspase 3.

Caspase 3 cleaves **ICAD**, the inhibitor of CAD (caspase activated DNase), which frees **CAD** to enter the nucleus and cleave DNA. Although caspase 3 is the pivotal execution caspase for apoptosis, the processes can be initiated by intrinsic signalling which always involves mitochondrial release of cytochrome C and activation of **caspase 9**.

The release of **cytochrome C** and mitochondrial inhibitor of IAPs is mediated via Bcl-2 family proteins (including Bax, Bak) forming pores in the mitochondrial membrane. Interestingly, the extrinsic apoptotic signal is aided and amplified by activation of **tBid** which recruits Bcl-2 family members and hence also activates the intrinsic pathway.

Apaf1, apoptotic protease activating factor 1; **Bid**, family member of Bcl-2 protein; **IAPs**, inhibitor of apoptosis proteins.

acteristically there is an influx of water and ions, after which cellular organelles swell and rupture. Cell lysis induces acute inflammatory responses in vivo owing to the release of lysosomal enzymes into the extracellular environment.

Apoptotic cell death has characteristic morphological features:

- chromatin aggregation, with nuclear and cytoplasmic condensation into distinct membrane-bound vesicles which are termed apoptotic bodies
- organelles remain intact
- cell 'blebs' (which are intact membrane vesicles)
- no inflammatory response
- cellular 'blebs' and remains are phagocytosed by adjacent cells and macrophages.

This process requires energy (ATP), and several Ca^{2+}- and Mg^{2+}-dependent nuclease systems are activated which specifically cleave nuclear DNA at the inter-histone residues.

An endonuclease destroys DNA following apoptosis. This involves the enzyme caspase (cysteine-containing aspartase-specific protease) which activates the CAD (caspase-activated DNase)/ICAD (inhibitor of CAD) system which can destroy DNA. Regulated apoptosis is essential for many life processes, from tissue structure formation in embryogenesis and wound healing to normal metabolic processes such as autodestruction of the thickened endometrium to cause menstruation in a non-conception cycle. In oncology, chemotherapy and radiotherapy regimens only work if they can trigger the tumour cells' own apoptotic pathways. Failure to do so in resistant tumours can result in the accumulation of further genetic damage to the surviving cells.

Several factors initiate apoptosis but in general there are two signalling pathways: the **extrinsic apoptotic pathway** triggered by death receptors on the cell surface and the **intrinsic pathway** initiated at the mitochondrial level. Death receptors are all members of the TNF receptor superfamily and include CD95 (APO-1/Fas), TRAIL (TNF-related apoptosis ligand)-R1, TRAIL-R2, TNF-R1, DR3 and DR6.

- The *extrinsic pathway* is involved in processes such as tissue remodelling and induction of immune self-tolerance. Activated receptors with internal death domain complexes multiply pro-caspase 8 molecules whose autocatalytic activity results in release of the initiator caspase 8 (Fig. 2.8). In turn caspase 8 cleaves pro-caspase 3 and caspase 3, in combination with the other effector caspases, activating DNA cleavage, cell condensation and fragmentation.
- The *intrinsic pathway* centres on the release of cytochrome C from the mitochondria. Cellular stress, such as growth factor withdrawal and p53 cell cycle arrest induces the expression of the pro-apoptotic Bcl-2 family of proteins, Bax and Bak. These form tetrameric complexes, which imbed into the outer mitochondrial membrane forming permissive pores. Cytochrome C, released from the mitochondria, binds Apaf1, forming a complex known as the apoptosome, which then activates an initiator caspase, in this case caspase 9. Caspase 9 then activates the effector caspase, caspase 3. Other proteins released from damaged mitochondria, Smac/DIABLO and Omi/HtrA2, counteract the effect of IAPs (inhibitor of apoptosis proteins), which normally bind and prevent activation of pro-caspase 3. Anti-apoptotic Bcl-2 protein, when incorporated as a member of the Bak/Bax pore complex, renders the mitochondrial pore non-permissive to release of cytochrome C and the anti-IAPs.

There is an amplification link between the extrinsic and intrinsic apoptotic pathways in that caspase 8 cleaves a Bcl-2 family member, tBid, which then aids formation of the Bcl-2/Bax/Bak pore complexes. If this complex is predominantly formed from pro-apoptotic members of the Bcl-2 family of

FURTHER READING

Choumerianou DM, Dimitriou H, Kalmanti M. Stem cells: promises versus limitations. *Tissue Engineering Part B Rev.* 2008; **14**(1): 53–60.

Friedman J, Xue D. To live or die by the sword: the regulation of apoptosis by the proteasome. *Developmental Cell* 2004; **6**(4): 460–461.

Wyllie AH (ed). Apoptosis: an overview. *British Medical Bulletin* 1997; **53**: 451–465.

proteins then apoptosome/caspase 9, along with mitochondrial anti-IAPs, amplifies the apoptotic activation of effector caspases 3 (Fig. 2.8). Conversely, overexpression of anti-apoptotic Bcl-2 will not only inhibit intrinsic, but also dampen down extrinsic apoptotic signalling.

Stem cells

The majority of our cells are terminally differentiated in that although they contain the blueprint to produce all the proteins ever expressed by the body, each tissue has permanently deactivated all but those required for the function of the specialized cells. Such differentiated cells can not replicate indefinitely and eventually shut down. Therefore we must have nests of cells within all our different tissues that have not shut down their genetic blueprint. These stem cells give rise to not only a daughter cell that is differentiated and of a limited ability to replicate but a daughter cell which will not differentiate and has the infinite ability to replicate. In mammals there are three broad source categories of stem cells:

- **embryonic stem cells**, derived from blastocysts
- **adult stem cells**, found in adult tissues
- **cord blood stem cells**, found in the umbilical cord.

The source of stem cells can also be subcategorized by 'potency'. This specifies the differentiation potential (the potential to differentiate into different cell types) of the stem cell.

Totipotent stem cells are produced from the fusion of an egg and sperm cell. Cells produced by the first few divisions of the fertilized egg are also totipotent. These cells can differentiate into embryonic and extraembryonic cell types.

Pluripotent stem cells are the descendants of totipotent cells and can differentiate into cells derived from any of the three germ layers.

Multipotent stem cells can produce only cells of a closely related family of cells (e.g. haematopoietic stem cells differentiate into red blood cells, white blood cells, platelets, etc.).

Unipotent cells can produce only one cell type, but have the property of self-renewal which distinguishes them from non-stem cells.

MOLECULAR BIOLOGY

In 2003 the Human Genome Project was completed. All 3.2×10^9 base pairs of DNA were sequenced. Over 99% of the DNA sequence is identical between individuals, but still millions of different base pair variations (polymorphisms) occur.

DNA STRUCTURE AND FUNCTION

Genetic information is stored in the form of double-stranded deoxyribonucleic acid (DNA). Each strand of DNA is made up of a deoxyribose–phosphate backbone and a series of purine (adenine (A) and guanine (G)) and pyrimidine (thymine (T) and cytosine (C)) bases of the nucleic acid. For practical purposes the length of DNA is generally measured in numbers of base-pairs (bp).

The monomeric unit in DNA (and in RNA) is the nucleotide, which is a base joined to a sugar–phosphate unit (Fig. 2.9a). The two strands of DNA are held together by hydrogen bonds between the bases. There are only four possible pairs of nucleotides – TA, AT, GC and CG (Fig. 2.9b). The two strands twist to form a double helix with major and minor grooves, and the large stretches of helical DNA are coiled around histone proteins to form nucleosomes and further condensed into the chromosomes that are seen at metaphase (Fig. 2.9c,d).

GENES

A gene is a portion of DNA that contains the codes for a polypeptide sequence. Three adjacent nucleotides (a codon) code for a particular amino acid, such as AGA for arginine, and TTC for phenylalanine. There are only 20 common amino acids, but 64 possible codon combinations that make up the genetic code. This means that most amino acids are encoded for by more than one triplet; other codons are used as signals for 'initiating' or 'terminating' polypeptide-chain synthesis.

Genes consist of lengths of DNA that contain sufficient nucleotide triplets to code for the appropriate number of amino acids in the polypeptide chains of a particular protein. Genes vary greatly in size: most extend over 20–40 kbp, but a few (such as the gene for the muscle protein dystrophin) can extend over millions of base-pairs. In bacteria the coding sequences are continuous, but in higher organisms these coding sequences (exons) are interrupted by intervening sequences that are non-coding (introns) at various positions (see Fig. 2.10). Some genes code for RNA molecules which will not be further translated into proteins. These code for functional ribosomal RNA (rRNA) and transfer RNA (tRNA), which play vital roles in polypeptide synthesis. Micro RNAs are single-stranded RNA molecules of about 22 nucleotides. They inactivate specific messenger RNAs and disrupt the expression of their proteins, thereby regulating cell proliferation and apoptosis. They in turn can be inactivated by DNA methylation.

TRANSCRIPTION AND TRANSLATION (Fig. 2.10)

The conversion of genetic information to polypeptides and proteins relies on the transcription of sequences of bases in DNA to messenger RNA molecules; mRNAs are found mainly in the nucleolus and the cytoplasm, and are polymers of nucleotides containing a ribose–phosphate unit attached to a base. The bases are adenine, guanine, cytosine and uracil (U) (which replaces the thymine found in DNA). RNA is a single-stranded molecule but it can hybridize with a complementary sequence of single-stranded DNA (ssDNA). Genetic information is carried from the nucleus to the cytoplasm by mRNA, which in turn acts as a template for protein synthesis.

Each base in the mRNA molecule is lined up opposite to the corresponding base in the DNA: C to G, G to C, U to A and A to T. A gene is always read in the 5′–3′ orientation and at 5′ promoter sites which specifically bind the enzyme

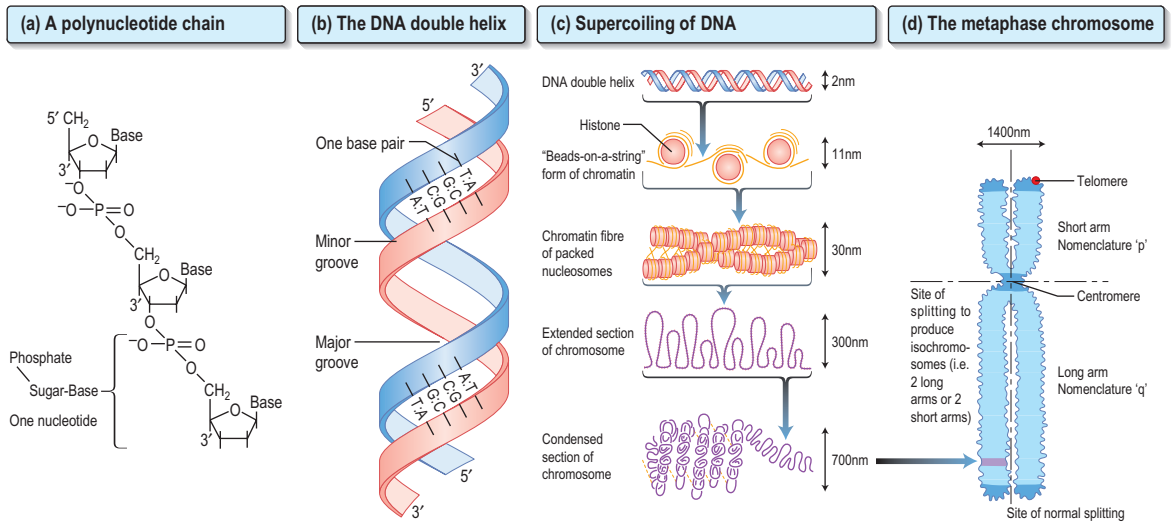


Fig. 2.9 **DNA and its structural relationship to human chromosomes.**

(a) A polynucleotide strand with the position of the nucleic bases indicated. Individual nucleotides form a polymer linked via the deoxyribose sugars. The 5′ carbon of the heterocyclic sugar structure links to the 3′ carbon of the next via a phosphate molecule forming the sugar–phosphate backbone of the nucleic acid. The 5′–3′ linkage gives an orientation to a sequence of DNA.

(b) Double-stranded DNA. The two strands of DNA are held together by hydrogen bonds between the bases. As T always pairs with A, and G with C, there are only four possible pairs of nucleotides – TA, AT, GC and CG. The orientation of the complementary single strands of DNA (ssDNA) is always opposite; i.e. one will be 5′–3′ whilst the partner will be 3′–5′. The helical 3D structure will have major and minor grooves and a complete turn of the helix will contain 12 base-pairs. These grooves are structurally important, as DNA-binding proteins predominantly interact with the major grooves.

(c) Supercoiling of DNA. In humans, and other higher organisms, the large stretches of helical DNA are coiled to form nucleosomes and further condensed into the chromosomes that can be seen at metaphase. DNA is first packaged by winding around nuclear proteins – histones – every 180 bp. This can then be coiled and supercoiled to compact nucleosomes and eventually visible chromosomes.

(d) At the end of the metaphase DNA replication will result in a twin chromosome joined at the centromere. This picture shows the chromosome, its relationship to supercoiling, and the positions of structural regions: centromeres, telomeres and sites where the double chromosome can split.

Nomenclature of chromosomes. This is the assigned number or X or Y, plus short arm (p) or long arm (q). The region or subregion is defined by the transverse light and dark bands observed when staining with Giemsa (hence G-banding) or quinacrine and numbered from the centromere outwards. **Chromosome constitution** = chromosome number + sex chromosomes + abnormality; e.g. 46XX = normal female 47XX+21 = Down's syndrome (trisomy 21) 46XYt (2;19) (p21;p12) = male with a normal number of chromosomes but a translocation between chromosome 2 and 19 with breakages at short-arm bands 21 and 12 of the respective chromosomes.

RNA polymerase and so indicate where transcription is to commence.

Two AT-rich promoter sites are present in eurokaryotic genes. The first, the *TATA box*, is located about 25 bp upstream of (or before) the transcription start site, while the second, the *CAAT box*, is 75 bp upstream of the start site.

The initial or primary mRNA is a complete copy of one strand of DNA and therefore contains both introns and exons. While still in the nucleus, the mRNA undergoes post-transcriptional modification whereby the 5′ and 3′ ends are protected by the addition of an inverted guanidine nucleotide (CAP) and a chain of adenine nucleotides (Poly A) (see Fig. 2.10). In higher organisms, the primary transcript mRNA is further processed inside the nucleus whereby the introns are spliced out. Splicing is achieved by small nuclear RNA in association with specific proteins. Furthermore, alternative splicing is possible whereby an entire exon can be omitted. Thus more than one protein can be coded from the same gene. The processed mRNA then migrates out of the nucleus into the cytoplasm. Polysomes (groups of ribosomes) become attached to the mRNA; the ribosomes consist of subunits composed of small RNA molecules (rRNA) and proteins. The rRNA components are key to the binding and translation of the genetic code. Held by the ribosomes, triplets of adjacent bases on the mRNA called codons are exposed and recognized by complementary sequences, or anti-codons, in transfer RNA (tRNA) molecules. Each tRNA molecule carries an amino acid that is specific to the anti-codon. As the ribosome passes along the mRNA in the 5′–3′ direction, amino acids are transferred from tRNA molecules and sequentially linked by the ribosome in the order dictated by the order of codons. The ribosome, in effect, moves along the mRNA like a 'zipper', linking the assembled amino acids to form a polypeptide chain. The first 20 or more nucleotides are recognition and regulatory sequences and are untranslated but necessary for translation and possibly earlier transcription.

Translation begins when the triplet AUG (methionine) is encountered. All proteins start with methionine but this is often lost as the leading sequence of amino acids of the native peptides is removed during protein folding and post-translational modification into a mature protein. Similarly the Poly A tail is not translated (3′ untranslated region) and is preceded by a stop codon, UAA, UAG or UGA.

THE CONTROL OF GENE EXPRESSION

Gene expression can be controlled at many points in the steps between the translation of DNA to proteins. Proteins

Ask the authors
www.studentconsult.com

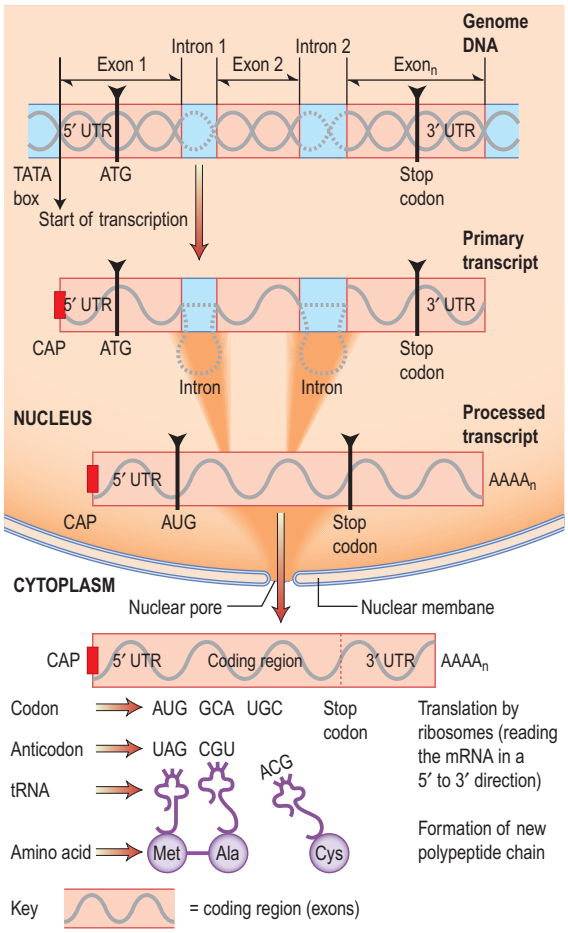


Fig. 2.10 **Transcription and translation (DNA to RNA to protein).** RNA polymerase creates an **RNA copy** of the DNA gene sequence. This primary transcript is processed: capping of the 5′ free end of the mRNA precursor involves the addition of an inverted guanine residue to the 5′ terminal which is subsequently methylated, forming a 7-methylguanosine residue. (The corresponding position on the gene is thus called the CAP site.) The 3′ end of an mRNA defined by the sequence AAUAAA acts as a cleavage signal for an endonuclease, which cleaves the growing transcript about 20 bp downstream from the signal. **The 3′ end** is further processed by a Poly A polymerase which adds about 250 adenosine residues to the 3′ end, forming a Poly A tail (polyadenylation). Without these additions, the mRNA sequence will be rapidly degraded 5′–3′ but the inverted cap nucleotide prevents nuclease attachment. The activity of specific 5′ mRNA nucleases to remove the cap is further regulated by the Poly A tail which must first be removed by other degradation enzymes. **Splicing** out of the introns then produces the mature mRNA (prokaryote genes do not contain introns). This then moves out of the nucleus via nuclear pores and aligns on endoplasmic reticulum. **Ribosomal subunits** assemble on the mRNA moving along 5′ to 3′. With the transport of amino acids to their active sites by specific tRNAs, the complex translates the code, producing the peptide sequence. Once formed, the peptide is released into the cytoplasmic reticulum for post-translational modification into a mature protein.

and RNA molecules are in a constant state of turnover; as soon as they are produced, processes for their destruction are at work. For many genes, transcriptional control is the key point of regulation. Deleterious, even oncogenic, changes to a cell's biology may arise through no fault in the expression of a particular gene. Apparent overexpression may be due to non-breakdown of mRNA or protein product. A pathway that stops gene expression by promoting degradation of RNA (RNA interference, RNAi) is also found.

Transcriptional control

Gene transcription (DNA to mRNA) is not a spontaneous event and is possible only as a result of the interaction of a number of DNA-binding proteins with genomic DNA. Regulation of a gene's expression must first start with the opening up of the double helix of DNA in the correct region of the chromosome. In order to do this, a class of protein molecules that recognize the outside of the DNA helix has evolved. These DNA-binding proteins preferentially interact with the major groove of the DNA double helix (see Fig. 2.9). The base-pair composition of the DNA sequence can change the geometry of a DNA helix to facilitate the fit of a DNA-binding protein with its target region: CG-rich areas form the Z-structure DNA helix; sequences such as AAAANNN cause a slight bend, and if this is repeated every 10 nucleotides it produces pronounced curves. DNA-binding proteins that recognize these distorted helices result either in the opening up of the helix so that the gene may be transcribed, or in the prevention of the helix being opened.

Structural classes of DNA-binding proteins

There are four basic classes of DNA-binding protein, classified according to their structural motifs (see Fig. 2.11 and Table 2.1).

Control regions and proteins

DNA-binding proteins act as regulators of gene expression in three different ways. They are the promoters, the operators and the enhancers. The primary gene expression regulators are the *promoters*. The RNA polymerases bind to a promoter region, normally adjacent to the transcribed sequence of DNA. In eukaryotes active transcription is possible only when a number of DNA-binding and associated proteins come together and interact. Known as 'general transcription factors', these proteins are thought to assemble at promoter sites used by RNA polymerases (e.g. Pol II) that are characterized by specific motifs such as the TATA sequence.

Other DNA regulator proteins operate in close proximity to the site of promoter binding. These are called *operator proteins/regions* and act either as repressors by binding to DNA sequences within the promoter site or as positive regulators facilitating RNA polymerase binding. The third class of regulator proteins operate as enhancer sequences which are at least 200 bp away from the site of transcription initiation. Binding of regulator proteins to enhancer regions, several hundred bases from the promoter site, upregulates the expression. This turns out to be a distance favourable for DNA to loop back on itself without straining the backbone bonds of the DNA double helix.

The GAL4 *enhancer* of yeast physically aids the binding of transcription factors to the TATA region of the promoter, and thus acts like a catalyst for general transcription factor assembly, and consequently also the rate of RNA polymerase activity. In mammals, it is frequently a region termed the 'cyclic AMP response element' (CRE) that acts in this manner. Increasing intracellular cAMP levels cause activation and release of CRE-binding protein (CREB). This binds to the CRE sequence and enhances the transcription rate.

These relatively remote regulatory regions need not just enhance gene expression but may also repress transcription.

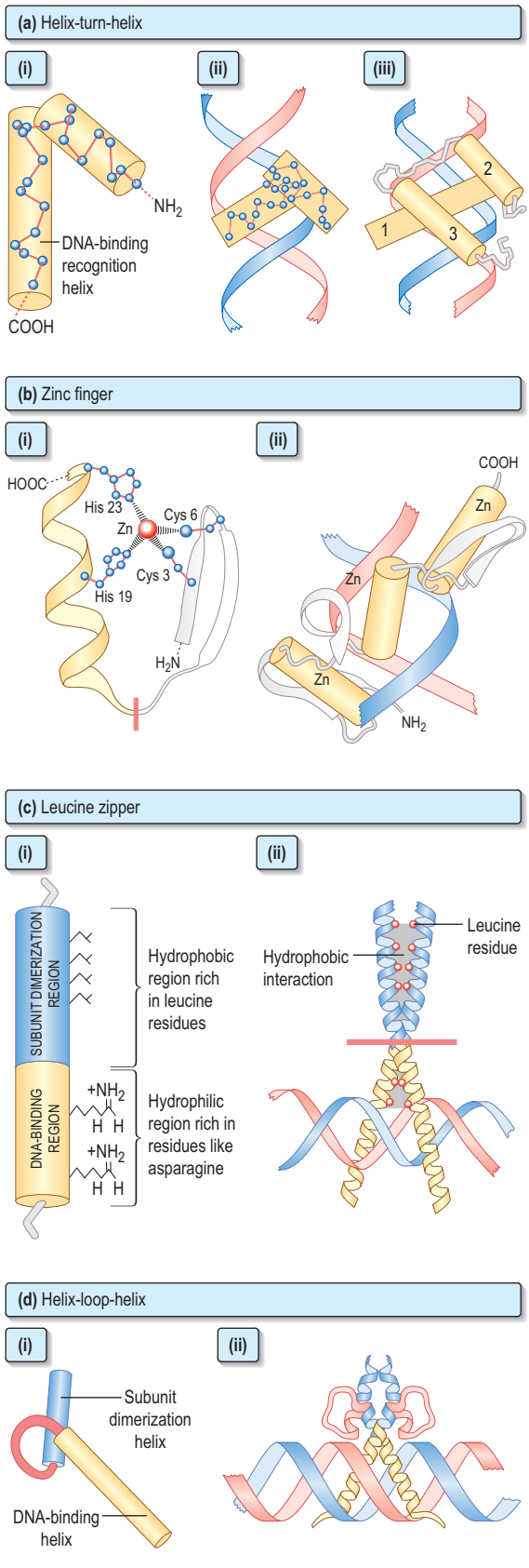


Fig. 2.11 **The four classes of DNA-binding proteins.**

(a) Helix–turn–helix (HTH) motifs. The simplest and most common, it consists of a helix connected by a fixed angle to a second helix. **(i)** This represents the 'core' motif of HTH DNA-binding proteins but considerable modifications occur. **(ii)** Like all DNA-binding proteins they interact as dimers, binding at exactly one turn of the double helix (3.4 nm or 12 bp) apart. The homeodomain HLH DNA-binding proteins are a special class of HTH proteins. Discovered in the 1980s during investigation of the gene controlling *Drosophila* development, they contain an almost identical stretch of 60 amino acids. They have a third helix which holds the DNA-binding region in a fixed orientation. Indeed, all homeodomain proteins appear to have specific conserved amino acid positions and residues which interact with DNA. **(iii)** Furthermore, random orientation extension chains at the N-terminal can interact with the minor groove.

(b) Zinc finger motifs, characterized by the incorporation of zinc, in some form, into the protein's quartenary structure. **(i)** The term originated from the structural model of a *Xenopus laevis* protein which uses a zinc ion to hold a loop in a 'finger'-shape by cross-linking two histidine residues with two cysteine residues. The Cys–Cys–His–His family of zinc finger DNA regulator proteins typically consists of an anti-parallel sheet forming a tight tertiary association with its own helix by zinc interaction with Cys and His residues within each of the two secondary structures. **(ii)** Several clusters of zinc fingers are found together and form a repeating structure which can interact with repetitive sequences in DNA.

(c) Leucine zipper motif. (i) This consists of a long α-helix which has many hydrophobic leucine residues at one end, responsible for dimer formation, whilst the opposite hydrophilic ends interact with DNA across the major groove of the double helix. **(ii)** The quaternary structure of the leucine zippers need not be homodimeric, and indeed heterodimers are extremely common. Thus multiple sequences may be recognized by two or three DNA leucine zipper proteins depending on the type of dimer formed.

(d) Helix–loop–helix (HLH) motifs. (i) These consist of a DNA-binding α-helix joined to a protein dimerization secondary α-helix via a loop. **(ii)** These combine leucine zipper properties with those of helix motif DNA-binding proteins. A helix interacting with DNA is linked via a loop to a large second helix which non-covalently binds to a similar HLH protein. Homo- and heterodimers can form, but the loop gives flexibility in the orientation of the given DNA-binding α-helix domain.

Table 2.1 Examples of DNA-binding proteins

Class of DNA-binding protein	Examples
Helix–turn–helix	CREB (cAMP response element binding protein)
Zinc finger	Steroid and thyroid hormone receptors Retinoic acid and vitamin D receptors *Bcl-6* oncogene product (lymphoma) *WT1* oncogene product (Wilms' tumour) GATA-1 erythrocyte differentiation and Hb expression factor *BRCA 1* (familial breast cancer)
Leucine zippers	*c-jun* cell replication oncogene *c-fos* cell replication oncogene
Helix–loop–helix	*myc* oncogene *mad* oncogene *max* oncogene

Indeed, loops can be formed from regulatory regions downstream as well as upstream of the gene's coding sequence. Repressors can inhibit the transcription of a given gene by binding to the regulatory sequence and blocking positive regulators, binding and thus inactivating the positive regulator, or by interfering with the promoter protein assembly. Multiple regulatory regions and DNA-binding proteins can surround a given gene and precisely control its expression at a basal level and in response to a cellular stimulus (Table 2.1).

CHROMOSOMES, INTRONS AND THE SIZE OF THE HUMAN GENOME

Human genetics differs from that of bacteria in structural components of the genetic code and in the way DNA is packaged. The fact that humans have chromosomes and introns affects how much DNA we require to code for all our proteins in an unusual way. We have far more DNA (3.2×10^9 bp) than protein-coding genes (32 000), whereas simple bacteria appear to have a much more economical DNA to gene ratio. Simple organisms like bacteria that have circular genomes that are not contained within a membrane-bound organelle (the nucleus) are termed 'prokaryotes'. Higher eukaryotic organisms have their linear genomic packages – chromosomes – separated from the general cytoplasm by the nuclear envelope. In eukaryotes, genomic DNA is associated with nuclear proteins called histones.

Coiling around histones and structural regions such as the centromeres and telomeres requires regions of DNA devoted specifically to the purpose of packaging (see Fig. 2.9c). Ten per cent of human DNA is highly repetitive (or 'satellite' DNA), consisting of long arrays of tandem repeats. These regions tend to be supercoiled around histones in condensed regions termed heterochromatin, even when the cell is not undergoing division. In contrast, most other DNA regions are relatively uncondensed and constitute the euchromatin. This remaining DNA is either moderately repetitive, accounting for about 30% of the genome, or codes for unique genes and gene families, occupying some 2% of the genome.

Although supercoiling of DNA around histones, intronic spacing (all giving tighter control of specific gene expression) and packaging into functional chromosomes accounts for some 40% of our total DNA, at least 60% is random and has no apparent function. However, life outside of the sea is exposed to increased DNA damage by radiation and chemicals. This additional genomic mass of terrestrial animals, compared to deep-sea creatures, may be there to absorb the increased risk of DNA damage.

FURTHER READING

Hurwitz J. The discovery of RNA polymerase. *Journal of Biological Chemistry* 2005; **280**: 42477–42485.

Lehman IR. Discovery of DNA polymerase. *Journal of Biological Chemistry* 2003; **278**: 34733–34738.

National Center for Biotechnology Information (NCBI). *Genes and Disease (eBook)*. Bethesda, MD: National Library of Medicine. http://www.ncbi.nlm.nih.gov/books/bv.fcgi

Stevenson M. Therapeutic potential of RNA interference. *New England Journal of Medicine* 2004; **351**: 1772–1777.

TOOLS FOR MOLECULAR BIOLOGY

Preparation of genomic DNA

The first step in studying the DNA of an individual involves preparation of genomic DNA. This is a simple procedure in which any cellular tissue including blood (the nucleated cells are isolated from the erythrocytes) can be used. The cells are lysed in order to open their cell and nuclear membranes, releasing chromosomal DNA. Following digestion of all cellular protein by the addition of proteolytic enzymes, the genomic DNA is isolated by chemical extraction with phenol. DNA is stable and can be stored for years.

Restriction enzymes and gel electrophoresis

Restriction enzymes recognize specific DNA sequences and cut double-stranded DNA at these sites. For example, the enzyme EcoRI will cut DNA wherever it reads the sequence GAATTC, and so human genomic DNA is cut into hundreds of thousands of fragments. Whenever the genomic DNA from an individual is cut with EcoRI, the same 'restriction fragments' (restriction fragment length polymorphisms, RFLPs) are produced. As DNA is a negatively charged molecule, genomic DNA fragments digested with a restriction enzyme can be separated according to their size and charge, by electrophoresing the DNA through a gel matrix. The DNA sample is loaded at one end of the gel, a voltage is applied across the gel, and the DNA migrates towards the positive anode. The small fragments move more quickly than the large fragments, and so the DNA fragments separate out. Fragment size can be determined by running fragments of known size on the same gel.

Pulsed-field gel electrophoresis (PFGE) can be used to separate very long pieces of DNA (hundreds of kilobases) which have been cut by restriction enzymes that cut at rare sites in the genome.

Hybridization techniques

A fundamental property of DNA is that when two strands are separated, for example by heating, they will always reassoci-

ate and stick together again because of their complementary base sequences. Therefore the presence or position of a particular gene can be identified using a gene 'probe' consisting of DNA or RNA with a base sequence that is complementary to that of the sequence of interest. A DNA probe is thus a piece of single-stranded DNA that can be labelled with a radioactive isotope (usually ^{32}P) or a fluorescent signal. The probe will locate and bind to its complementary sequence.

Hybridization is exploited in a number of techniques including:

- *Southern blotting* (DNA fragments separated by gel electrophoresis and transferred onto a membrane sheet)
- *Northern blotting* (RNA separated by gel electrophoresis and transferred onto a membrane sheet)
- *in situ hybridization* (localization of native nucleic acid sequences within the cell and its component organelles, including chromosomes).

The polymerase chain reaction (PCR)

Minute amounts of DNA can be amplified over a million times within a few hours using this in vitro technique (Fig. 2.12). The exact DNA sequence to be amplified needs to be known because the DNA is amplified between two short (generally 17–25 bp) single-stranded DNA fragments ('oligonucleotide primers') which are complementary to the sequences at each end of the DNA of interest.

The technique has three steps.

- The double-stranded genomic DNA is denatured by heat into single-stranded DNA.
- The reaction is then cooled to favour DNA annealing, and the primers bind to their target DNA.
- Finally, a DNA polymerase is used to extend the primers in opposite directions using the target DNA as a template.

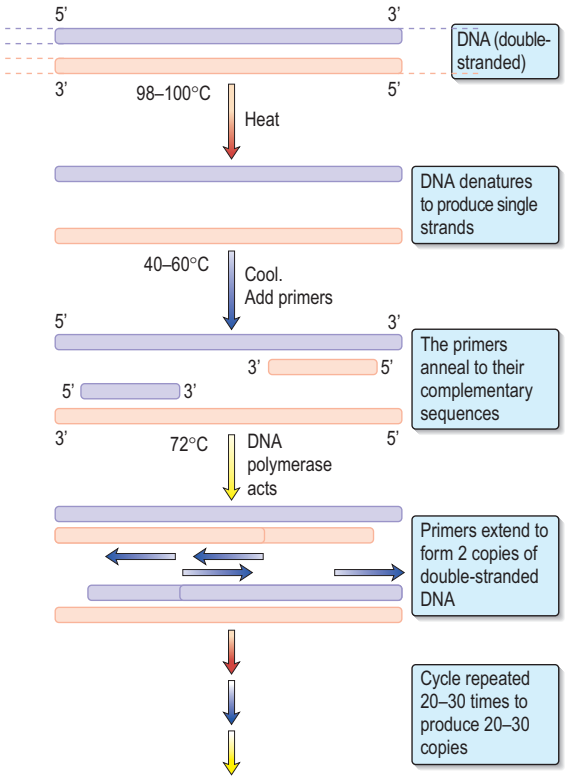


Fig. 2.12 **Polymerase chain reaction.**

After one cycle there are two copies of double-stranded DNA, after two cycles there are four copies, and this number rises exponentially with the number of cycles. Typically a polymerase chain reaction is set for 25–30 cycles, allowing millions of amplifications, i.e. 2^n where n = number of cycles.

This technique has revolutionized genetic research because minute amounts of DNA not previously amenable to analysis can be amplified, such as from buccal cell scrapings, blood spots, or single embryonic cells.

Real-time PCR

Real-time polymerase chain reaction, also called quantitative real-time polymerase chain reaction (QRT-PCR), is the simultaneous quantification and amplification of a given DNA sequence. It can be used to determine whether or not a specific sequence is present in the sample (e.g. a viral genome) and if it is present, the number of copies in the sample. Frequently, real-time polymerase chain reaction is combined with reverse transcription polymerase chain reaction to quantify low abundance messenger RNA (mRNA), enabling a researcher to quantify relative gene expression at a particular time in a particular cell/tissue.

The procedure follows the general pattern of the polymerase chain reaction, but the DNA is quantified after each round of amplification; this is the 'real-time' aspect of it. Two common methods of quantification are the use of fluorescent dyes that intercalate with double-stranded DNA, and modified DNA oligonucleotide probes that fluoresce when hybridized with a complementary DNA.

Expression microarrays/gene chips

This is a methodology developed to examine the relative abundance of mRNA for thousands of genes present in cells/tissue of different types or conditions, for example, to examine the changes in gene expression from normal colonic tissue to that of malignant colonic polyps. The basic technology is the ability to immobilize sequences of DNA complementary to specific genes or different regions of known genes, onto a solid surface in precise microdot arrays (Fig. 2.13). Total mRNA is extracted from one tissue and labelled with fluorescent tag Cy3-green, and the mRNA from the second tissue with fluorescent tag Cy5-red. The two fluorescent-tagged total mRNA samples are mixed in a 1:1 ratio and washed over the DNA gene chips. The mRNA for specific genes will bind to their complementary microdot and can be detected by laser-induced excitation of the fluorescent tag and the position and light wavelength and intensity recorded by a scanning confocal microscope. The relative intensity of Cy5-red:Cy3-green is a reliable measure of the relative abundance of specific mRNAs in each sample. Yellow results from equal binding of both fluorescent-tagged mRNAs. If no hybridization occurs on a dot then the area is black. The power of the system is that many thousands of genes can be screened not only for their expression but relative expression in normal and diseased tissue. Microarray analysis is now widely used for gene expression profiling and mutation detection; it is rapid and many samples can be analysed.

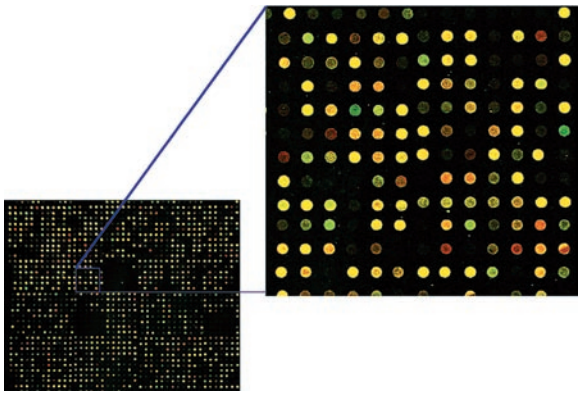


Fig. 2.13 **Expression microarray**. Example of part of a gene chip with hundreds of DNA sequences in microspots. The enlarged area shows the colour coding of no expression (black), overexpression (red), underexpression (green) and equal expression (yellow) of detected mRNA in two related samples.

DNA cloning

A particular DNA fragment of interest can be isolated and inserted into the genome of simple self-replicating organisms or organelles such as viruses and plasmids. When used for this purpose they are referred to as vectors. Replication by the million of the vector results in multiple copies or clones of the inserted sequence. Thus, after removal from the host vector, cloned gene sequences can be prepared in large quantities independently of other sequences. Vectors include: bacteriophage viruses; plasmids, which are self-replicating episomal circular DNA molecules found in bacteria that carry antibody resistance genes; and 'yeast artificial chromosomes' (YACs), which are derived from centromeric and telomeric DNA sequences found in yeast. Each vector takes an optimum size of cloned DNA insert. Typically, viruses can accommodate only small sequences up to a maximum of a kilobase, larger fragments of 2–10 kilobases can be inserted into a plasmid, and sequences of several hundred kilobases can be inserted into a YAC. Each has its relative merits, viruses being very efficient, but the vectors which take large clones being considerably less so. A hybrid between a plasmid and a bacteriophage (called a cosmid) has been constructed artificially. This has the ability to clone reasonably large sequences as plasmids within host bacteria. However, cosmids trick bacteriophages into packaging them into a viral body, and this viral body is then able to infect the target bacteria, giving efficient transfection rates.

The DNA fragment of interest is inserted into the vector DNA sequence using an enzyme called a ligase. This takes place in vitro. The next step, cloning, creates many copies of the 'recombinant DNA molecule' and takes place in vivo when the plasmid or other vector is placed back into the bacterial (or yeast) host. Bacteria that have successfully taken up the recombinant plasmid can be selected if the plasmid also carries an antibiotic resistance gene (so bacteria without the plasmid die in the presence of antibiotic).

The DNA fragment of interest to be cloned may be a restriction fragment. Alternatively it could be cDNA which has been copied from an mRNA sequence. mRNA provides the template from which a viral enzyme called reverse transcriptase (RT) can synthesize a complementary single-stranded DNA copy (cDNA). A DNA polymerase may then be used to produce a double-stranded copy by PCR. A cDNA molecule contains all the sequences necessary for a functional gene, but unlike genomic DNA it lacks introns.

DNA sequencing

A chemical process known as dideoxy-sequencing allows the identification of the exact nucleotide sequence of a piece of DNA. As in PCR, an oligonucleotide primer is annealed adjacent to the region of interest. This primer acts as the starting point for a DNA polymerase to build a new DNA chain that is complementary to the sequence under investigation. Chain extension can be prematurely interrupted when a dideoxynucleotide becomes incorporated (because they lack the necessary 3′ -hydroxyl group). As the dideoxynucleotides are present at a low concentration, not all the chains in a reaction tube will incorporate a dideoxynucleotide in the same place; so the tubes contain sequences of different lengths but which all terminate with a particular dideoxynucleotide. Each base dideoxynucleotide (G, C, T, A) has a different fluorochrome attached, and thus each termination base can be identified by its fluorescent colour. As each strand can be separated efficiently by capillary electrophoresis according to its size/length, simply monitoring the fluorescence as the reaction products elute from the capillary will give the gene sequence (see Fig. 2.14).

Genetic polymorphisms

Techniques have been developed to identify and quantitate genetic polymorphisms such as single nucleotide morphisms (SNPs). SNPs (pronounced 'snips') consist usually of two nucleotides at a particular site and vary between populations

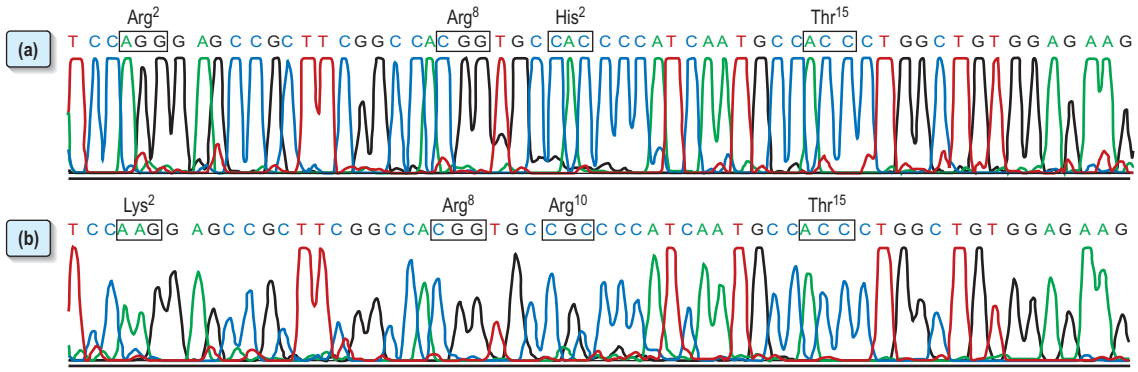


Fig. 2.14 **Gene sequencing.** Sequence profile of: **(a)** part of the normal luteinizing hormone beta chain gene; **(b)** that from a polymorphic variant. The base changes and consequential alterations in amino acid residues are indicated.

and ethnic groups. They must occur in at least 1% of the population to be a SNP. SNPs can be in coding or non-coding regions of the genes or be between genes and thus may not change the amino acid sequence of the protein. SNP databases can be developed by microarray techniques. Thus polymorphisms can be mapped by using SNPs or different patterns of DNA fragments with a restriction enzyme technique. An inherited disease can be attributed to a specific gene by comparing a 'map' of an affected with an unaffected person and identifying the 'locus' where the abnormal gene is placed.

Linkage disequilibrium. SNPs that are closer together (50 kb or less) are more likely to have alleles that move together in a block than those further apart. This phenomenon is called 'linkage disequilibrium' and enables us to use one SNP variant in this block to act as a marker for the presence of other SNP variants.

The 'lod score'. The likelihood of recombination between the marker under study and the disease allele must be taken into account. This measure of likelihood is known as the 'lod score' (the logarithm of the odds) and is a measure of the statistical significance of the observed co-segregation of the marker and the disease gene, compared with what would be expected by chance alone. Positive lod scores make linkage more likely; negative lod scores make it less likely. By convention a lod score of +3 is taken to be definite evidence of linkage because this indicates 1000 to 1 odds that the co-segregation of the DNA marker and the disease did not occur by chance alone.

Linkage analysis has provided many breakthroughs in mapping the positions of genes that cause genetic diseases, such as the gene for cystic fibrosis which was found to be tightly linked to a marker on chromosome 7, or the gene for Friedreich's ataxia which is tightly linked to a marker on chromosome 9.

International Hapmap Project. A haplotype map ('hapmap') is being developed by collaborating international scientists. SNPs close together are inherited in blocks (haplotypes). Each haplotype has a few SNPs that can be identified (tag SNPs), which reduces the number of SNPs in the genome that have to be identified and typed and correlated with a phenotype. A comprehensive hapmap will enable studies on the role of genetic factors in different diseases and is being used to decide which therapy is likely to have a beneficial effect (p. 467).

The Wellcome Trust Case Control Consortium was set up to analyse thousands of DNA samples from patients with different diseases in which there is thought to be a genetic component. Genetic markers have been found in many diseases including diabetes, hypertension and Crohn's disease. Many polymorphic markers are also being studied in breast cancer, multiple sclerosis and autoimmune thyroid disease.

In silico cloning

Information arising from the human genome and similar animal programmes is now being used for bioinformatic (semi-automated) identification of genes. In silico cloning is largely based on the abundance of databases of *expressed sequence tags* (ESTs). ESTs represent portions of expressed genes and are produced by one-shot sequencing of cDNA clones taken from a cDNA library. These are rapidly generated and result in sequences whose length is limited by the prevailing technology – usually only 500–800 nucleotides in length. The significance is that these ESTs represent markers of active genes present in the specific tissue from which the cDNA library was generated. Bioinformatics allows the searching and matching of ESTs to complementary sequences (spliced or not) in the huge databases of the genome project. The position of the gene can be defined and other ESTs in libraries throughout the world can be matched to the same location, mapping out the expressed gene. The genomic limits of the expressed gene, predicting its protein product, shape and function and tissue specific expression can all be achieved by searching and manipulating computer-based databases.

FURTHER READING

Ding C, Cantor CR. Quantitative analysis of nucleic acids – the last few years of progress. *Journal of Biochemistry and Molecular Biology* 2004; **37**: 1–10.

Griffiths AJF, Wessler SR, Lewontin RC et al. *Introduction to Genetic Analysis,* 8th edn. New York: WH Freeman, 2004.

Nolan T, Hands RE, Bustin SA. Quantification of mRNA using real-time RT-PCR. *Nature Protocols* 2006; **1**: 1559–1582.

THE BIOLOGY OF CHROMOSOMES

HUMAN CHROMOSOMES

The nucleus of each diploid cell contains 6×10^9 bp of DNA in long molecules called *chromosomes*. Chromosomes are massive structures containing one linear molecule of DNA that is wound around histone proteins into small units called nucleosomes, and these are further wound to make up the structure of the chromosome itself. Diploid human cells have 46 chromosomes, 23 inherited from each parent; thus there are 23 'homologous' pairs of chromosomes (22 pairs of 'autosomes' and two 'sex chromosomes'). The sex chromosomes, called X and Y, are not homologous but are different in size and shape. Males have an X and a Y chromosome; females have two X chromosomes. (Primary male sexual characteristics are determined by the *SRY* gene – sex determining region, Y chromosome.)

The chromosomes can be classified according to their size and shape, the largest being chromosome 1. The constriction in the chromosome is the centromere, which can be in the middle of the chromosome (metacentric) or at one extreme end (acrocentric). The centromere divides the chromosome into a short arm and a long arm, referred to as the p arm and the q arm respectively (see Fig. 2.9d). In addition, chromosomes can be stained when they are in the metaphase stage of the cell cycle and are very condensed. The stain gives a different pattern of light and dark bands that is diagnostic for each chromosome. Each band is given a number, and gene mapping techniques allow genes to be positioned within a band within an arm of a chromosome. For example, the *CFTR* gene (in which a defect gives rise to cystic fibrosis) maps to 7q21; that is, on chromosome 7 in the long arm in band 21.

During cell division (mitosis), each chromosome divides into two so that each daughter nucleus has the same number of chromosomes as its parent cell. During gametogenesis, however, the number of chromosomes is halved by meiosis, so that after conception the number of chromosomes remains the same and is not doubled. In the female, each ovum contains one or other X chromosome but, in the male, the sperm bears either an X or a Y chromosome.

Chromosomes can only be seen easily in actively dividing cells. Typically, lymphocytes from the peripheral blood are stimulated to divide and are processed to allow the chromosomes to be examined. Cells from other tissues can also be used – for example amniotic fluid, placental cells from chorionic villus sampling, bone marrow and skin (see Box 2.1).

The biology of chromosomes
Human chromosomes
The X chromosome and inactivation
Telomeres and immortality
The mitochondrial chromosome
GENETIC DISORDERS
Chromosomal disorders

Box 2.1 Indications for chromosomal analysis

Chromosome studies may be indicated in the following circumstances.

Antenatal

- Pregnancies in women over 35 years
- Positive maternal serum screening test for aneuploid pregnancy
- Ultrasound features consistent with an aneuploid fetus
- Severe fetal growth retardation
- Sexing of fetus in X-linked disorders

In the neonate

- Congenital malformations
- Suspicion of trisomy or monosomy
- Ambiguous genitalia

In the adolescent

- Primary amenorrhoea or failure of pubertal development
- Growth retardation

In the adult

- Screening parents of a child with a chromosomal abnormality for further genetic counselling
- Infertility or recurrent miscarriages
- Learning difficulties
- Certain malignant disorders (e.g. leukaemias and Wilms' tumour)

THE X CHROMOSOME AND INACTIVATION

Although female chromosomes are XX, females do not have two doses of X-linked genes (compared with just one dose for a male XY) because of the phenomenon of X inactivation or Lyonization (after its discoverer, Dr Mary Lyon). In this process, one of the two X chromosomes in the cells of females becomes transcriptionally inactive, so the cell has only one dose of the X-linked genes. Inactivation is random and can affect either X chromosome.

TELOMERES AND IMMORTALITY

The ends of chromosomes, telomeres (see Fig. 2.9d), do not contain genes but many repeats of a hexameric sequence TTAGGG. Replication of linear chromosomes starts at coding sites (origins of replication) within the main body of chromosomes and not at the two extreme ends. The extreme ends are therefore susceptible to single-stranded DNA degradation back to double-stranded DNA. Thus cellular ageing can be measured as a genetic consequence of multiple rounds of replication with consequential telomere shortening. This leads to chromosome instability and cell death.

Stem cells have longer telomeres than their terminally differentiated daughters. However, germ cells replicate without shortening of their telomeres. This is because they express an enzyme called telomerase, which protects against telomere shortening by acting as a template primer at the extreme ends of the chromosomes. Most somatic cells (unlike germ and embryonic cells) switch off the activity of telomerase after birth and die as a result of apoptosis. Many cancer cells, however, reactivate telomerase, contributing to their immortality. Conversely, cells from patients with progeria (premature ageing syndrome) have extremely short telomeres. Transient expression of telomerase in various stem and daughter cells is part of their normal biology. The inability to activate telomerase (in addition to overexpression) in cells such as those of the immune system can produce disease pathologies.

THE MITOCHONDRIAL CHROMOSOME

In addition to the 23 pairs of chromosomes in the nucleus of every diploid cell, the mitochondria in the cytoplasm of the cell also have their own genome. The mitochondrial chromosome is a circular DNA (mtDNA) molecule of approximately 16 500 bp, and every base-pair makes up part of the coding sequence. These genes principally encode proteins or RNA molecules involved in mitochondrial function. These proteins are components of the mitochondrial respiratory chain involved in oxidative phosphorylation (OXPHOS) producing ATP. They also have a critical role in apoptotic cell death. Every cell contains several hundred mitochondria, and therefore several hundred mitochondrial chromosomes. Virtually all mitochondria are inherited from the mother as the sperm head contains no (or very few) mitochondria. Disorders are shown in Figure 2.15.

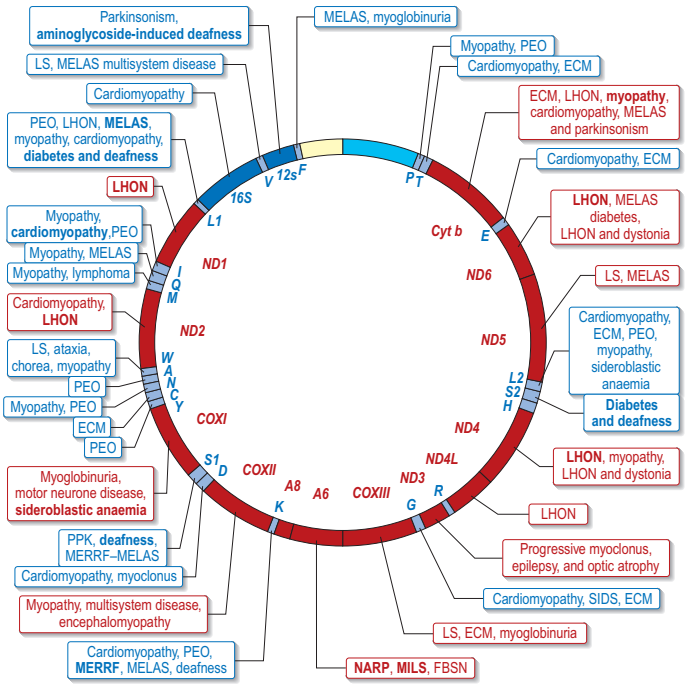


Fig. 2.15 **Mitochondrial chromosome abnormalities**. Disorders that are frequently or prominently associated with mutations in a particular gene are shown in bold. Diseases due to mutations that impair mitochondrial protein synthesis are shown in blue. Diseases due to mutations in protein-coding genes are shown in red.

ECM, encephalomyopathy; FBSN familial bilateral striatal necrosis; LHON, Leber's hereditary optic neuropathy; LS, Leigh's syndrome; MELAS, mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes; MERRF, myoclonic epilepsy with ragged-red fibres; MILS, maternally inherited Leigh's syndrome; NARP, neuropathy, ataxia and retinitis pigmentosa; PEO progressive external ophthalmoplegia; PPK, palmoplantar keratoderma; SIDS, sudden infant death syndrome.

FURTHER READING

Nosek J, Tomaska L, Bolotin-Fukuhara M et al. Mitochondrial chromosome structure: an insight from analysis of complete yeast genomes. *FEMS Yeast Research* 2006; **6**(3): 356–370.

Wong JMY, Collins K. Telomere maintenance and disease. *Lancet* 2003; **362**: 983–988.

GENETIC DISORDERS

The spectrum of inherited or congenital genetic disorders can be classified as the chromosomal disorders, including mitochondrial chromosome disorders, the Mendelian and sex-linked single-gene disorders, a variety of non-Mendelian disorders, and the multifactorial and polygenic disorders (Table 2.2 and Box 2.2). All are a result of a mutation in the genetic code. This may be a change of a single base-pair of a gene, resulting in functional change in the product protein (e.g. thalassaemia) or gross rearrangement of the gene within a genome (e.g. chronic myeloid leukaemia). These mutations can be congenital (inherited at birth) or somatic (arising during a person's life). The latter are responsible for the collective disease known as cancer, and the principles underlying Mendelian inheritance act in a similar manner to produce dominant and recessive traits. Both gross chromosomal and point mutations occur in somatic genetic disease.

CHROMOSOMAL DISORDERS

Chromosomal abnormalities are much more common than is generally appreciated. Over half of spontaneous abortions have chromosomal abnormalities, compared with only 4–6 abnormalities per 1000 live births. Specific chromosomal abnormalities can lead to well-recognized and severe clinical syndromes, although autosomal aneuploidy (a differing from the normal diploid number) is usually more severe than the sex-chromosome aneuploidies. Abnormalities may occur in either the number or the structure of the chromosomes.

Abnormal chromosome numbers

If a chromosome or chromatids fail to separate ('non-disjunction') either in meiosis or mitosis, one daughter cell will receive two copies of that chromosome and one daughter cell will receive no copies of the chromosome. If this non-disjunction occurs during meiosis it can lead to an ovum or sperm having either (i) an extra chromosome, so resulting in a fetus that is 'trisomic' and has three instead of two copies of the chromosome; or (ii) no chromosome, so the fetus is 'monosomic' and has one instead of two copies of the chromosome. Non-disjunction can occur with autosomes or sex chromosomes. However, only individuals with trisomy 13, 18 and 21 survive to birth, and most children with trisomy 13 and trisomy 18 die in early childhood. Trisomy 21 (Down's syndrome) is observed with a frequency of 1 in 650 live births regardless of geography or ethnic background. This should be reduced with widespread screening (p. 45). Full autosomal monosomies are extremely rare and very deleterious. Sex-

Table 2.2 Prevalence of genetic disease

Genetic disease	Congenital malformation
0.5% of all newborns have a chromosomal abnormality	3–5% of all births result in congenital malformations
7% of all stillborns have a chromosomal abnormality	30–50% of post-neonatal deaths are due to congenital malformations
20–30% of all infant deaths are due to genetic disorders	18% of paediatric hospital admissions are for children with congenital malformations
11% of paediatric hospital admissions are for children with genetic disorders	
12% of adult hospital admissions are for genetic causes	
15% of all cancers have an inherited susceptibility	
10% of chronic diseases of the adult population (heart, diabetes, arthritis) have a significant genetic component	

European incidences per 1000 births

Single gene disorders:	**10**	**Congenital malformations:**	**31**
Dominant	7.0	Genetically determined	0.6
Recessive	1.66	Multifactorial	30
X-linked	1.33	Non-genetic	~0.4
Chromosomal disorders:	**3.5**		
Autosomes	1.69		
Sex chromosomes	1.80		

Although individual genetic diseases are rare, regional variation is enormous – the incidence of Down's syndrome varies from 1/1000 to 1/100 world-wide. Single gene diseases collectively comprise over 15 500 recognized genetic disorders. The global prevalence of all single gene diseases at birth is approximately 10/1000.

Box 2.2 Genetic disorders

Mendelian

- Inherited or new mutation
- Mutant allele or pair of mutant alleles at single locus
- Clear pattern of inheritance (autosomal or sex-linked) dominant or recessive
- High risk to relatives

Chromosomal

- Loss, gain or abnormal rearrangement of one or more of 46 chromosomes in diploid cell
- No clear pattern of inheritance
- Low risk to relatives

Multifactorial

- Common
- Interaction between genes and environmental factors
- Low risk to relatives

Mitochondrial

- Due to mutations in mitochondrial genome
- Transmitted through maternal line
- Different pattern of inheritance from Mendelian disorders

Somatic cell

- Mutations in somatic cells
- Somatic event is not inherited
- Often give rise to tumours

chromosome trisomies (e.g. Klinefelter's syndrome, XXY) are relatively common. The sex-chromosome monosomy in which the individual has an X chromosome only and no second X or Y chromosome is known as *Turner's syndrome* and is estimated to occur in 1 in 2500 live-born girls.

Occasionally, non-disjunction can occur during mitosis shortly after two gametes have fused. It will then result in the formation of two cell lines, each with a different chromosome complement. This occurs more often with the sex chromosome, and results in a 'mosaic' individual.

Very rarely the entire chromosome set will be present in more than two copies, so the individual may be triploid rather than diploid and have a chromosome number of 69. Triploidy and tetraploidy (four sets) result in spontaneous abortion.

Abnormal chromosome structures

As well as abnormal numbers of chromosomes, chromosomes can have abnormal structures, and the disruption to the DNA and gene sequences may give rise to a genetic disease.

- *Deletions*. Deletions of a portion of a chromosome may give rise to a disease syndrome if two copies of the genes in the deleted region are necessary, and the individual will not be normal with just the one copy remaining on the non-deleted homologous chromosome. Many deletion syndromes have been well described. For example, *Prader–Willi syndrome* (p. 228) is the result of cytogenetic events resulting in deletion of part of the long arm of chromosome 15, *Wilms' tumour* is characterized by deletion of part of the short arm of chromosome 11, and microdeletions in the long arm of chromosome 22 give rise to the *DiGeorge syndrome*.
- *Duplications*. Duplications occur when a portion of the chromosome is present on the chromosome in two copies, so the genes in that chromosome portion are present in an extra dose. A form of neuropathy, *Charcot–Marie–Tooth disease* (p. 1175), is due to a small duplication of a region of chromosome 17.
- *Inversion*. Inversions involve an end-to-end reversal of a segment within a chromosome; e.g. abcdefgh becomes abcfedgh, for example, haemophilia (p. 438).
- *Translocations*. Translocations occur when two chromosome regions join together, when they would not normally. Chromosome translocations in somatic cells may be associated with tumorigenesis (see p. 451).

Translocations can be very complex, involving more than two chromosomes, but most are simple and fall into one of two categories:

- *Reciprocal translocations* occur when any two non-homologous chromosomes break simultaneously and rejoin, swapping ends. In this case the cell still has 46 chromosomes but two of them are rearranged. Someone with a balanced translocation is likely to be normal (unless a translocation breakpoint interrupts a gene); but at meiosis, when the chromosomes separate into different daughter cells, the translocated chromosomes will enter the gametes and any resulting fetus may inherit one abnormal chromosome and have an unbalanced translocation, with physical manifestations.
- *Robertsonian translocations* occur when two acrocentric chromosomes join and the short arm is lost, leaving only 45 chromosomes. This translocation is balanced as no genetic material is lost and the individual is healthy. However, any offspring have a risk of inheriting an unbalanced arrangement. This risk depends on which acrocentric chromosome is involved. Clinically relevant is the 14/21 Robertsonian translocation. A woman with this karyotype has a 1 in 8 risk of having a baby with Down's syndrome (a male carrier has a 1 in 50 risk). However, they have a 50% risk of producing a carrier like themselves, hence the necessity for genetic family studies. Relatives should be alerted to the increased risk of Down's syndrome in their offspring, and should have their chromosomes checked.

Table 2.3 shows some of the syndromes resulting from chromosomal abnormalities.

Mitochondrial chromosome disorders (see Fig. 2.15)

The mitochondrial chromosome (p. 37) carries its genetic information in a very compact form; for example there are no introns in the genes. Therefore any mutation has a high chance of having an effect. However, as every cell contains hundreds of mitochondria, a single altered mitochondrial genome will not be noticed. As mitochondria divide there is a statistical likelihood that there will be more mutated mitochondria, and at some point this will give rise to a mitochondrial disease.

Most mitochondrial diseases are myopathies and neuropathies with a maternal pattern of inheritance. Other abnormalities include retinal degeneration, diabetes mellitus and hearing loss. Many syndromes have been described.

Myopathies include chronic progressive external ophthalmoplegia (CPEO); encephalomyopathies include myoclonic epilepsy with ragged red fibres (MERRF) and mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes (MELAS) (see p. 1182).

Kearns–Sayre syndrome includes ophthalmoplegia, heart block, cerebellar ataxia, deafness and mental deficiency due to long deletions and rearrangements. *Leber's hereditary optic neuropathy* (LHON) is the commonest cause of blindness in young men, with bilateral loss of central vision and cardiac arrhythmias, and is an example of a mitochondrial disease caused by a point mutation in one gene.

Multisystem disorders include *Pearson's syndrome* (sideroblastic anaemia, pancytopenia, exocrine pancreatic failure, subtotal villous atrophy, diabetes mellitus and renal tubular dysfunction). In some families, hearing loss is the only symptom, and one of the mitochondrial genes implicated may predispose patients to aminoglycoside ototoxicity.

Analysis of chromosome disorders

The cell cycle can be arrested at mitosis with colchicines and, following staining, the chromosomes with their characteristic banding can be seen and any abnormalities identified (Fig. 2.16).

YAC-cloned probes are also available and cover large genetic regions of individual chromosomes. These probes can be labelled with fluorescently tagged nucleotides and used in in situ hybridization of the nucleus of isolated tissue

Table 2.3 Chromosomal abnormalities: examples of a few syndromes

Syndrome	Chromosome karyotype	Incidence and risks	Clinical features	Mortality
Autosomal abnormalities				
Trisomy 21 (Down's syndrome)	47, +21 (95%) Mosaicism Translocation 5%	1:650 (overall) (risk with 20- to 29-year-old mother 1:1000; > 45-year-old mother 1:30)	Flat face, slanting eyes, epicanthic folds, small ears, simian crease, short stubby fingers, hypotonia, variable learning difficulties, congenital heart disease up to 50%)	High in first year, but many survive to adulthood
Trisomy 13 (Patau's syndrome)	47, +13	1:5000	Low-set ears, cleft lip and palate, polydactyly, micro-ophthalmia, learning difficulties	Rarely survive for more than a few weeks
Trisomy 18 (Edwards' syndrome)	47, +18	1:3000	Low-set ears, micrognathia, rocker-bottom feet, learning difficulties	Rarely survive for more than a few weeks
Sex chromosome abnormalities				
Fragile X syndrome	46, XX, fra (X) 46, XY, fra (X)	1:2000	Most common inherited cause of learning difficulties predominantly in male Macro-orchidism	
Female				
Turner's syndrome	45, XO	1:2500	Infantilism, primary amenorrhoea, short stature, webbed neck, cubitus valgus, normal IQ	
Triple X syndrome	47, XXX	1:1000	No distinctive somatic features, learning difficulties	
Others	48, XXXX 49, XXXXX	Rare	Amenorrhoea, infertility, learning difficulties	
Male				
Klinefelter's syndrome	47, XXY (or XXYY)	1:1000 (more in sons of older mothers)	Decreased crown–pubis:pubis–heel ratio, eunuchoid, testicular atrophy, infertility, gynaecomastia, learning difficulties (20%; related to number of X chromosomes)	
Double Y syndrome	47, XYY	1:800	Tall, fertile, minor mental and psychiatric illness, high incidence in tall criminals	
Others	48, XXXY 49, XXXXY		Learning difficulties, testicular atrophy	

from patients. These tagged probes allow rapid and relatively unskilled identification of metaphase chromosomes, and allow the identification of chromosomes dispersed within the nucleus. Furthermore, tagging two chromosome regions with different fluorescent tags allows easy identification of chromosomal translocations (Fig. 2.17).

GENE DEFECTS

Mendelian and sex-linked single-gene disorders are the result of mutations in coding sequences and their control elements. These mutations can have various effects on the expression of the gene, as explained below, but all cause a dysfunction of the protein product.

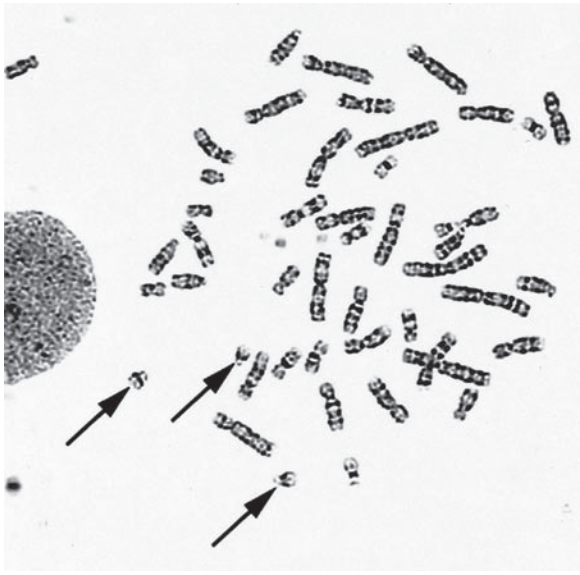
Mutations

Although DNA replication is a very accurate process, occasionally mistakes occur to produce changes or mutations. These changes can also occur owing to other factors such as radiation, ultraviolet light or chemicals. Mutations in gene sequences or in the sequences which regulate gene expression (transcription and translation) may alter the amino acid sequence in the protein encoded by that gene. In some cases protein function will be maintained; in other cases it will change or cease, perhaps producing a clinical disorder.

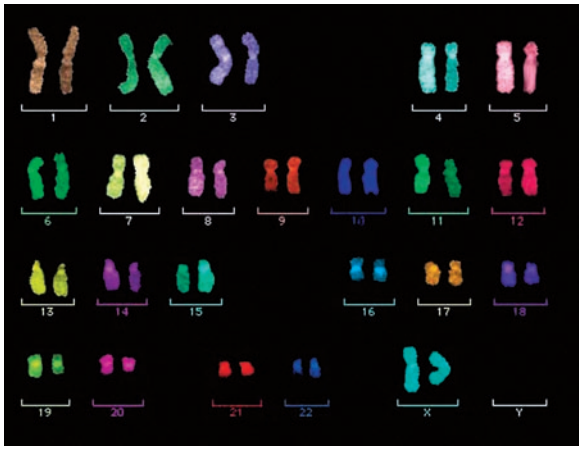
Many different types of mutation occur.

Point mutation

This is the simplest type of change and involves the substitution of one nucleotide for another, so changing the codon in



(a)



(b)

Fig. 2.16 **Karyotyping**. **(a)** G-banded spread of metaphase chromosomes, showing trisomy 21 (arrowed) Down's syndrome. **(b)** Whole chromosome painting probes labelled with different fluorochromes or fluorochrome combinations on 24 colour or mFISH analysis is based on a DNA probe kit. Digital image analysis software analyses the colour information and identifies the chromosomal origin of each individual pixel within the image. Inter-chromosomal rearrangements will show up as colour changes within the aberrant chromosome. Courtesy of D Lillington, Medical Oncology Unit, St Bartholomew's Hospital.

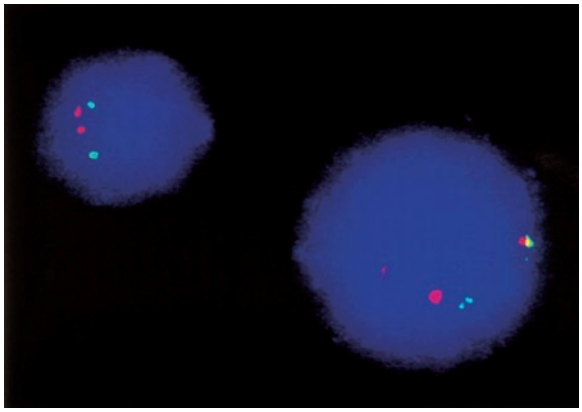


Fig. 2.17 **Fluorescence in situ hybridization (FISH).** Two-coloured FISH using a *red* paint for the *ABL* gene on chromosome 9 and a *green* paint for the *BCR* gene on chromosome 22 in the anaphase cell nucleus. Where a translocation has occurred the two genes become juxtaposed on the Philadelphia chromosome and a hybrid *yellow* fluorescence can be seen only in the affected cell nucleus on the right. Courtesy of D Lillington, Medical Oncology Unit, St Bartholomew's Hospital.

a coding sequence. For example, the triplet AAA, which codes for lysine, may be mutated to AGA, which codes for arginine. Whether a substitution produces a clinical disorder depends on whether it changes a critical part of the protein molecule produced. Fortunately, many substitutions have no effect on the function or stability of the proteins produced as several codons code for the same amino acid. However, some mutations may have a severe effect; for example, in sickle cell disease a mutation within the globin gene changes one codon from GAG to GTG, so that instead of glutamic acid, valine is incorporated into the polypeptide chain, which radically alters its properties.

Insertion or deletion

Insertion or deletion of one or more bases is a more serious change, as it results in the alteration of the rest of the following sequence to give a frame-shift mutation. For example, if the original code was:

TAA GGA GAG TTT

and an extra nucleotide (A) is inserted, the sequence becomes:

TAA AGG AGA GTT T

Alternatively, if the third nucleotide (A) is deleted, the sequence becomes:

TAG GAG AGT TT

In both cases, different amino acids are incorporated into the polypeptide chain. This type of change is responsible for some forms of *thalassaemia* (p. 407). Insertions and deletions can involve many hundreds of base-pairs of DNA. For example, some large deletions in the dystrophin gene remove coding sequences and this results in *Duchenne muscular dystrophy*. Insertion/deletion (ID) polymorphism in the angiotensin-converting enzyme (ACE) gene has been shown to result in the genotypes II, ID and DD. The deletion is of a 287 bp repeat sequence, and DD is associated with higher concentrations of circulating ACE and possibly *cardiac disease* (see p. 747).

Splicing mutations

If the DNA sequences which direct the splicing of introns from mRNA are mutated, then abnormal splicing may occur. In this case the processed mRNA which is translated into protein by the ribosomes may carry intron sequences, so altering which amino acids are incorporated into the polypeptide chain.

Termination mutations

Normal polypeptide chain termination occurs when the ribosomes processing the mRNA reach one of the chain termination or 'stop' codons (see above). Mutations involving these codons will result in either late or premature termination. For example, Haemoglobin Constant Spring is a haemoglobin variant where instead of the 'stop' sequence, a single base change allows the insertion of an extra amino acid (see p. 409).

The Copy Number Variation Project has been set up to investigate large pieces of DNA between 10 thousand and 5 million letters (copy number variation) in the hope of identifying diseases that may be associated with this type of mutation.

More online
www.studentconsult.com

Single-gene disease

Monogenetic disorders involving single genes can be inherited as dominant, recessive or sex-linked characteristics. Although classically divided into autosomal dominant, recessive or X-linked disorders, many syndromes show multiple forms of inheritance pattern. This is predominantly because multiple defects can occur within a given disease-associated gene or in separate genes which all contribute to a particular molecular/cellular pathway and thus give rise to the same phenotype. For example in Ehlers–Danlos syndrome we find autosomal dominant, recessive and X-linked inheritance. In addition there is a spectrum between autosomal recessive and autosomal dominance in that having just one defective allele gives a mild form of the disease whilst having both alleles with the mutation results in a more severe form of the syndrome. In some cases, such as factor V Leiden disease, the boundary between dominant and recessive forms is very blurred.

A list of some autosomal dominant and autosomal recessive genetic diseases with their chromosomal localization can be found on the internet (http://www.genetests.org). Some diseases show a racial or geographical prevalence. Thalassaemia (see p. 407) is seen mainly in Greeks, South East Asians and Italians, porphyria variegata occurs more frequently in the South African white population, and Tay–Sachs (p. 1072) disease particularly occurs in Ashkenazi Jewish people. The most common recessive disease in the UK is cystic fibrosis (see pp. 46 and 844). Thus the prevalence of a disease may be low world-wide but much higher in specific populations. Some disorders are rare to the extent that the prevalence is counted in the total number of cases ever reported. Molecular biology has enabled the subclassification of some syndromes according to which gene is giving rise to the disease (e.g. polycystic kidney disease) and the proportion of cases arising from defects in that particular gene can be estimated from linkage analysis. However, this may not be reflected in the detection rate of a given clinical genetic test, since a single test cannot detect all the possible mutations arising at a particular locus.

Autosomal dominant disorders (Fig. 2.18a)

Each diploid cell contains two copies of all the autosomes. An autosomal dominant disorder occurs when one of the two copies has a mutation and the protein produced by the normal form of the gene cannot compensate. In this case a heterozygous individual who has two different forms (or alleles) of the same gene will manifest the disease. The offspring of heterozygotes have a 50% chance of inheriting

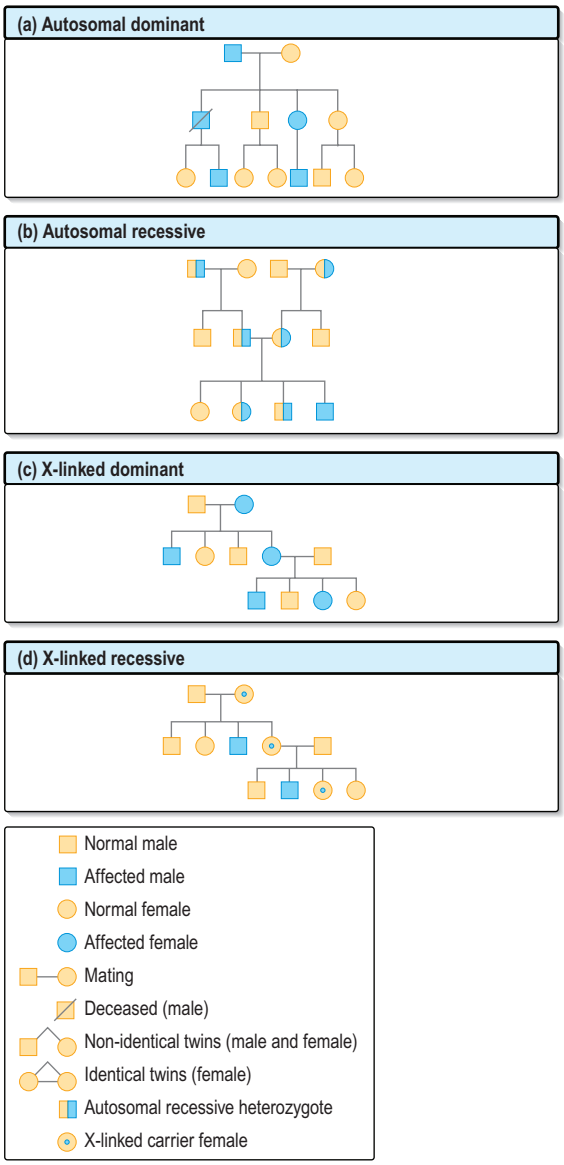


Fig. 2.18 **Modes of inheritance of simple gene disorders**, with a key to the standard pedigree symbols.

the chromosome carrying the disease allele, and therefore also of having the disease. However, estimation of risk to offspring for counselling families can be difficult because of three factors:

- These disorders have a great variability in their manifestation. 'Incomplete penetrance' may occur if patients have a dominant disorder but it does not manifest itself clinically in them. This gives the appearance of the gene having 'skipped' a generation.
- Dominant traits are extremely variable in severity (variable expression) and a mildly affected parent may have a severely affected child.
- New cases in a previously unaffected family may be the result of a new mutation. In this case, the risk of a further affected child is negligible. Most cases of achondroplasia are due to new mutations.

The overall incidence of autosomal dominant disorders is 7 per 1000 live births.

Autosomal recessive disorders (Fig. 2.18b)

These disorders manifest themselves only when an individual is homozygous for the disease allele, i.e. both chromosomes carry the mutated gene. In this case the parents are generally unaffected, healthy carriers (heterozygous for the disease allele). There is usually no family history, although the defective gene is passed from generation to generation. The offspring of an affected person will be healthy heterozygotes unless the other parent is also a carrier. If carriers marry, the offspring have a 1 in 4 chance of being homozygous and affected, a 1 in 2 chance of being a carrier, and a 1 in 4 chance of being genetically normal. Consanguinity increases the risk. The clinical features of autosomal recessive disorders are usually severe; patients often present in the first few years of life and have a high mortality. The overall incidence of autosomal recessive disorders is about 2.5 per 1000 live births.

Sex-linked disorders (Fig. 2.18c,d)

Genes carried on the X chromosome are said to be 'X-linked', and can be dominant or recessive in the same way as autosomal genes. As females have two X chromosomes they will be unaffected carriers of X-linked recessive diseases. However, since males have just one X chromosome, any deleterious mutation in an X-linked gene will manifest itself because no second copy of the gene is present.

X-linked dominant disorders

These are rare. Females who are heterozygous for the mutant gene and males who have one copy of the mutant gene on their single X chromosome will manifest the disease. Half the male or female offspring of an affected mother and all the female offspring of an affected man will have the disease. Affected males tend to have the disease more severely than the heterozygous female.

X-linked recessive disorders

These disorders present in males and present only in homozygous females (usually rare). X-linked recessive diseases are transmitted by healthy female carriers or affected males if they survive to reproduce. An example of an X-linked recessive disorder is haemophilia A (see p. 438), which is caused by a mutation in the X-linked gene for factor VIII. It has been shown that in 50% of cases there is an intrachromosomal rearrangement (inversion) of the tip of the long arm of the X chromosome (one break point being within intron 22 of the factor VIII gene).

Of the offspring from a carrier female and a normal male:

- 50% of the girls will be carriers as they inherit a mutant allele from their mother and the normal allele from their father; the other 50% of the girls inherit two normal alleles and are themselves normal
- 50% of the boys will have haemophilia as they inherit the mutant allele from their mother (and the Y chromosome from their father); the other 50% of the boys will be normal as they inherit the normal allele from their mother (and the Y chromosome from their father).

The male offspring of a male with haemophilia and a normal female will not have the disease as they do not inherit his X chromosome. However, all the female offspring will be carriers as they all inherit his X chromosome.

Y-linked genes

Genes carried on the Y chromosome are said to be Y-linked and only males can be affected. However, there are no known examples of Y-linked single-gene disorders which are transmitted.

More online
www.studentconsult.com

Sex-limited inheritance

Occasionally a gene can be carried on an autosome but manifests itself only in one sex. For example, frontal baldness is an autosomal dominant disorder in males but behaves as a recessive disorder in females.

More online
www.studentconsult.com

Other single-gene disorders

These are disorders which may be due to mutations in single genes but which do not manifest as simple monogenic disorders. They can arise from a variety of mechanisms, including the following.

Triplet repeat mutations (Table 2.4)

In the gene responsible for dystrophia myotonica (p. 1181), the mutated allele was found to have an expanded 3′UTR region in which three nucleotides, CTG, were repeated up to about 200 times. In families with dystrophia myotonica, people with the late-onset form of the disease had 20–40 copies of the repeat, but their children and grandchildren who presented with the disease from birth had vast increases

Table 2.4 Examples of trinucleotide repeat genetic disorders

Syndrome inheritance pattern	Disease prevalence	Gene, product, location and disorder	Genetic test detection rate
Friedreich's ataxia – AR	2–4/100 000	*FRDA* (Frataxin) 9q13 – GAA trinucleotide repeat expansion disorder in intron 1 of FRDA	96%
Fragile X syndrome – X-linked	16–25/100 000	*FMR1* (Fragile X mental retardation 1 protein) Xq27.3 – CGG trinucleotide repeat expansion and methylation changes in the 5′ untranslated region of *FMR1* exon 1	99%
Huntington's disease – AD	3–15/100 000	*HD* (Huntingtin protein) 4p16.3 – CAG trinucleotide repeat expansion within the translated protein giving rise to long tracts of repeat glutamine residues in HD	98%
Dystrophia myotonica – AD	1/20 000	*DMPK* (Myotonin-protein kinase) 19q13.2–13.3 – CTG trinucleotide repeat expansion in the 3′ untranslated region of the *DMPK* gene (dystrophia myotonica 1–DM1). Less common form – expanded. CCTG repeat in zinc finger protein 9 (*ZNF9*) gene (DM2)	100%

AD, autosomal dominant; AR, autosomal recessive.

in the number of repeats, up to 2000 copies. It is thought that some mechanism during meiosis causes this 'triplet repeat expansion' so that the offspring inherit an increased number of triplets. The number of triplets affects mRNA and protein function. See also page 45 for the phenomenon of 'anticipation'.

Mitochondrial disease (see Fig. 2.15)

As discussed on pages 39 and 40, various mitochondrial gene mutations can give rise to complex disease syndromes with incomplete penetrance maternal inheritance.

Imprinting

It is known that normal humans need a diploid number of chromosomes of 46. However, the maternal and paternal contributions are different and, in some way which is not yet clear, the fetus can distinguish between the chromosomes inherited from the mother and the chromosomes inherited from the father, although both give 23 chromosomes. In some way the chromosomes are 'imprinted' so that the maternal and paternal contributions are different. Imprinting is relevant to human genetic disease because different phenotypes may result depending on whether the mutant chromosome is maternally or paternally inherited. A deletion of part of the long arm of chromosome 15 (15q11–q13) will give rise to the *Prader–Willi syndrome* (PWS) if it is paternally inherited. A deletion of a similar region of the chromosome gives rise to *Angelman's syndrome* (AS) if it is maternally inherited. The affected gene has been identified as ubiquitin (*UBE3A*). Significantly, maternal chromosome 15 *UBE3A* is expressed in the brain and hypothalamus. Defective maternal ubiquitin in Angelman's syndrome is thus responsible for accumulation of undegraded protein, and hence neuronal damage.

Complex traits: multifactorial and polygenic inheritance

Characteristics resulting from a combination of genetic and environmental factors are said to be multifactorial; those involving multiple genes can also be said to be polygenic.

Measurements of most biological traits (e.g. height) show a variation between individuals in a population, and a unimodal, symmetrical (Gaussian) frequency distribution curve can be drawn. This variability is due to variation in genetic factors and environmental factors. Environmental factors may play a part in determining some characteristics, such as weight, whilst other characteristics such as height may be largely genetically determined. This genetic component is thought to be due to the additive effects of a number of alleles at a number of loci, many of which can be individually identified using molecular biological techniques, for example studying identical twins in different environments.

A common genetic variation in the 3′ untranslated region of the prothrombin gene is associated with elevated prothrombin (up to fourfold) and an increased risk of venous thrombosis (thrombophilia, see p. 442). Presumably the 3′ mutation induces overexpression of the prothrombin gene by increasing mRNA stability. In some pedigree studies, those homozygous for this mutation and heterozygous for factor V Leiden mutation have an even greater risk of thrombotic events (p. 442). Conversely, downregulation of the α_1-antitrypsin gene (due to similar mutation in the 3′ and 5′ untranslated regions) is associated with emphysema and cirrhosis (p. 358).

Table 2.5 Examples of disorders that may have a polygenic inheritance

Disorder	Frequency (%)	Heritability (%)*
Hypertension	5	62
Asthma	4	80
Schizophrenia	1	85
Congenital heart disease	0.5	35
Neural tube defects	0.5	60
Congenital pyloric stenosis	0.3	75
Ankylosing spondylitis	0.2	70
Cleft palate	0.1	76

*Percentage of the total variation of a trait which can be attributed to genetic factors.

There are sex differences. Congenital pyloric stenosis is most common in boys, but if it occurs in girls the latter have a larger number of affected relatives. This difference suggests that a larger number of the relevant genes are required to produce the disease in girls than in boys. Most human diseases, such as heart disease, diabetes and common mental disorders, are multifactorial traits (Table 2.5).

POPULATION GENETICS

The genetic constitution of a population depends on many factors. The Hardy–Weinberg equilibrium is a concept, based on a mathematical equation, that describes the outcome of random mating within populations. It states that 'in the absence of mutation, non-random mating, selection and genetic drift, the genetic constitution of the population remains the same from one generation to the next'.

This genetic principle has clinical significance in terms of the number of abnormal genes in the total gene pool of a population. The Hardy–Weinberg equation states that:

$$p^2 + 2pq + q^2 = 1$$

where p is the frequency of the normal gene in the population, q is the frequency of the abnormal gene, p^2 is the frequency of the normal homozygote, q^2 is the frequency of the affected abnormal homozygote, $2pq$ is the carrier frequency, and $p + q = 1$.

Example. The equation can be used, for example, to find the frequency of heterozygous carriers in cystic fibrosis. The incidence of cystic fibrosis is 1 in 2000 live births. Thus q^2 = 1/2000, and therefore q = 1/44. Since $p = 1 - q$, then p = 43/44. The carrier frequency is represented by $2pq$, which in this case is 1/22. Thus 1 in 22 individuals in the whole population is a heterozygous carrier for cystic fibrosis.

CLINICAL GENETICS AND GENETIC COUNSELLING

Genetic disorders pose considerable health and economic problems because often there is no effective therapy. In any pregnancy the risk of a serious developmental abnormality is approximately 1 in 30 pregnancies; approximately 15% of paediatric inpatients have a multifactorial disorder with a predominantly genetic element.

FURTHER READING

Brown LY, Brown SA. Alanine tracts: the expanding story of human illness and trinucleotide repeats. *Trends in Genetics* 2004; **20**: 51–58.

Leonard JV, Shapira AHV. Mitochondrial respiratory chain disorders I and II. *Lancet* 2000; **355**: 299–304, 389-394.

Sybert VP, McCauley E. Turner's syndrome. *New England Journal of Medicine* 2004; **351**: 1227–1238.

People with a history of a congenital abnormality in a member of their family often seek advice as to why it happened and about the risks of producing further abnormal offspring. Interviews must be conducted with great sensitivity and psychological insight, as parents may feel a sense of guilt and blame themselves for the abnormality in their child.

Genetic counselling should have the following aims:

- *Obtaining a full history*. The pregnancy history, drug and alcohol ingestion during pregnancy and maternal illnesses (e.g. diabetes) should be detailed.
- *Establishing an accurate diagnosis*. Examination of the child may help in diagnosing a genetically abnormal child with characteristic features (e.g. trisomy 21) or whether a genetically normal fetus was damaged in utero.
- *Drawing a family tree is essential*. Questions should be asked about abortions, stillbirths, deaths, marriages, consanguinity and medical history of family members. Diagnoses may need verification from other hospital reports.
- *Estimating the risk of a future pregnancy being affected or carrying a disorder.* Estimation of risk should be based on the pattern of inheritance. Mendelian disorders (see earlier) carry a high risk; chromosomal abnormalities carry a low risk. Empirical risks may be obtained from population or family studies.
- *Information giving* on prognosis and management with adequate time given so that all information is discussed openly, freely and repeated as necessary.
- *Continued support and follow-up*. Explanation of the implications for other siblings and family members.
- *Genetic screening*. This includes prenatal diagnosis if requested, carrier detection and data storage in genetic registers. A growing number of molecular genetic tests are now available (http://www.geneclinics.org).

Genetic counselling should be non-directive, with the couple making their own decisions on the basis of an accurate presentation of the facts and risks in a way they can understand.

GENETIC ANTICIPATION

It has been noted that successive generations of patients with, for example, dystrophia myotonica and Huntington's chorea, present earlier and with progressively worse symptoms. This 'anticipation' is due to unstable mutations occurring within the disease gene. Trinucleotide repeats such as CTG (dystrophia myotonica) and CAG (Huntington's chorea) expand within the disease gene with each generation, and somatic expansion with cellular replication is also observed. This novel type of genetic mutation can occur within the translated region or untranslated (and presumably regulatory) regions of the target genes. This genetic distinction has been used to subclassify a number of genetic diseases which have now been shown to be caused by trinucleotide repeat expansion and display phenotypic 'anticipation' (Table 2.4).

PRENATAL DIAGNOSIS

This should be offered to *all* pregnant women in the UK (NICE Clinical Guideline 62, 2008). Practice and uptake varies in different maternity units, with some offering screening only to high-risk mothers. The risks of Down's syndrome increase disproportionately and rapidly for children born to mothers older than 35 years. Infants born to mothers with a history or family history of other conditions due to chromosomal abnormalities may be at increased risk.

Personal choice

There should be a detailed discussion with all mothers as to the possible consequences of each screening test before they are offered it.

In particular they should have an understanding of the failure rates, the detection rates, the false positive and most importantly the false negative rates of each test so that they can properly exercise choice.

Investigations depend on gestation

7–11 Weeks (vaginal ultrasound)

Confirm viability, fetal number and gestation by crown–rump measurement.

11–13 Weeks and 6 days (combined test)

Ultrasound for nuchal translucency measurement (normal fold less than 6 mm) to attempt to detect major chromosomal abnormalities (e.g. trisomies and Turner's syndrome).

Maternal serum is tested for pregnancy-associated plasma protein-A (PAPP-A from the syncytial trophoblast) and β-human chorionic gonadotrophin for trisomy 21. The combined test is more accurate than the triple test alone at 16 weeks.

NB. All serum marker measurements are corrected for gestational ages, a multiple of the mean (MOM) value for the appropriate week of gestation. If abnormalities are detected, it is necessary to continue to discuss whether further investigation is desired or not. Chorionic villus sampling (CVS) at 11–13 weeks under ultrasound control to sample the placental site, or amniocentesis at 15 weeks to sample amniotic fluid and obtain the fetal cells necessary for cytogenetic testing are the next options.

14–20 Weeks (serum triple or quadruple test) if too late for earlier test or previous option not offered)

The triple test for chromosomal abnormalities consists of testing maternal serum for:

- α-fetoprotein (low)
- unconjugated oestradiol (low)
- human chorionic gonadotrophin (high) for Down's syndrome and for neural tube defects.

The α-fetoprotein is high for neural tube defects.

The **quadruple test** also measures inhibin A – high in Down's syndrome.

14–22 Weeks

Ultrasound for structural abnormalities (e.g. neural tube defects; the gestation period for detection depends on severity). The best time to detect congenital heart defects is 18–22 weeks.

Reported detection rates for all congenital defects vary, e.g. from 14 to 61% for hypoplastic ventricle to 97 to 100% for anencephaly.

In time some of these tests are likely to be superseded by the salvage of fetal cells from the maternal blood, from cervical secretions or by retrieving maternal plasma cell free fetal DNA. Other conditions such as myotonic dystrophy and Huntington's chorea may be detected from fetal circulating nucleic acids.

FURTHER READING

Finsterer J. Mitochondriopathies. *European Journal of Neurology* 2004; **11**: 163–186.

Guetta E, Simchen MJ, Mammon-Daviko K et al. Analysis of fetal blood cells in the maternal circulation: challenges, ongoing efforts, and potential solutions. *Stem Cells and Development* 2004; **13**: 93–99.

Kumar S, O'Brien A. Recent developments in fetal medicine. *British Medical Journal* 2004; **328**: 1002–1006.

NICE. *NICE Clinical Guideline 62 – Antenatal Care.* NICE, 2008. www.nice.org.uk/CG062

APPLICATIONS OF MOLECULAR GENETICS

The use of molecular biological techniques in genetics is having a massive impact on the investigation, diagnosis, treatment and control of genetic disorders.

The avoidance and control of genetic disease

Some genetic disorders, such as phenylketonuria or haemophilia, can be managed by diet or replacement therapy, but most have no effective treatment. By understanding what causes genetic damage, potential mutagens such as radiation, environmental chemicals, viruses or drugs (e.g. thalidomide) can be avoided.

GENE THERAPY

There are many technical problems to overcome in gene therapy, particularly in finding delivery systems to introduce DNA into a mammalian cell. Very careful control and supervision of gene manipulation will be necessary because of its potential hazards and the ethical issues.

Gene therapy entails placing a normal copy of a gene into the cells of a patient who has a defective copy of the gene.

Experiments have concentrated on recessive disorders, such as cystic fibrosis where the disease is due to the absence of a normal gene product. In dominant disorders the pathogenic potential of the mutant allele is normally expressed in the presence of a normal allele. This requires gene correction, whereby the mutant sequence is replaced by an equivalent sequence from a normal allele, or the mutant allele is inactivated. Such procedures are more difficult.

Two major factors are involved in gene therapy:

- the introduction of the functional gene sequence into target cells
- the expression and permanent integration of the transfected gene into the host cell genome.

Suitable diseases for current gene therapy experiments include cystic fibrosis, adenosine deaminase deficiency, and familial hypercholesterolaemia.

Cystic fibrosis (see also p. 361 and p. 844)

The gene responsible for cystic fibrosis (the cystic fibrosis transmembrane regulator gene (*CFTR*)) was first localized to chromosome 7 by linkage analysis. The *CFTR* gene spans about 250 kbp and contains 27 exons. The DNA sequence analysis predicts a polypeptide sequence of 1480 amino acids. The *CFTR* gene also encodes a simple chloride ion channel within the cystic fibrosis transmembrane regulator (Fig. 2.19). The commonest is a single mutation with a 3 bp deletion in exon 10 resulting in the removal of a codon specifying phenylalanine (F508del). There are also over 1000 different minor mutations of the *CFTR* gene with most mapping to the ATP-binding domains.

Gene therapy experiments are still under way in trying to restore *CFTR* function by transfection of cells with wild type receptor. Two different routes have been tried, either placing the *CFTR* gene in an adenovirus vector (Fig. 14.30) or into a liposome. The latter can be conveyed to the lung using an aerosol spray, and the fatty surface of the liposome fuses with the cell membrane to deliver the *CFTR* DNA into the cell, where the gene should function normally. However, neither is yet a treatment option. An alternative method is to suppress premature termination codons and thus permit translation to continue; topical nasal gentamicin (an aminoglycoside antibiotic) has been shown to result in the expression of functional *CFTR* channels.

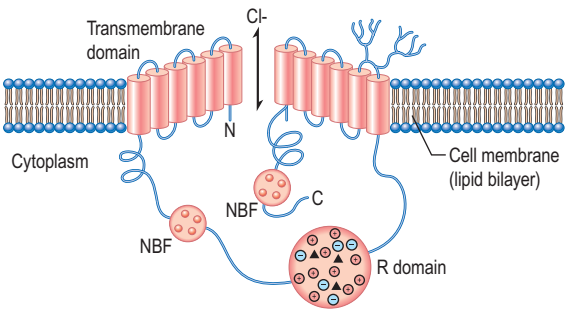


Fig. 2.19 **Model of cystic fibrosis transmembrane regulator (CFTR).** This is an integral membrane glycoprotein, consisting of two repeated elements. The cylindrical structures represent six membrane-spanning helices in each half of the molecule. The nucleotide-binding folds (NBFs) are in the cytoplasm, and the dots in these shaded areas represent the means of entry by the nucleotide. The regulatory (R) domain links the two halves and contains charged individual amino acids and protein kinase phosphorylation sites (black triangles). N and C are the N and C terminals. The branched structure on the right half represents potential glycosylation sites. The chloride channel is shown.

Adenosine deaminase (ADA) deficiency

Gene therapy for this rare immunodeficiency disease entails introducing a normal human ADA gene into the patient's lymphocytes to reconstitute the function of the cellular and humoral immune system in severe combined immunodeficiency (SCID). Lymphocytes have been tried for short-term therapy, but for longer-term treatment bone marrow transplantation would be the definitive approach (see p. 462).

Familial hypercholesterolaemia

This disorder is a result of a defective low-density lipoprotein (LDL) receptor gene. In therapy, a receptor gene is inserted into hepatocytes, removed by liver biopsy from the patient. Gene-corrected hepatocytes are then reinjected into the portal circulation of the patient. These cells migrate back to the liver where they are reincorporated and should start to produce LDL receptor protein, which would dramatically lower the patient's cholesterol level.

Muscle-cell-mediated gene therapy

Much of the problem associated with gene therapy of chronic genetic disease is how long the transfected cells will survive. Isolated myoblasts transfected with a retrovirus have been shown to function and live for the lifetime of mouse models (2 years). The myoblasts fuse with the animal's muscle fibres and express the transfected protein. The long life of the muscle fibres and their rich blood supply make them an ideal site for treatment of diseases in which functional serum-born factors are missing (e.g. human growth hormone, coagulation factors and erythropoietin).

Obviously, myoblast transfection would appear to be the best option for the treatment of Duchenne muscular dystrophy. However, the Duchenne gene is far too large to fit into

www.studentconsult.com
Ask the authors

any current viral vector, and reimplanted myoblasts do not colonize muscles fibres distant from the site of injection.

Treatment of somatic disease

Gene therapy which requires only a transient expression of the transfected genetic material circumvents the problems currently plaguing gene therapy of inherited disorders. This may also prove to be the front-line of gene therapy.

Vascular disease

The ideal therapy in heart disease would be neovascularization to increase blood flow and the repair of cardiac tissue after a myocardial infarction. In fact, any vascular disease might require regeneration or new blood vessel growth. Temporary expression of angiogenic factors at the site of a blockage would induce new blood vessels. Alternatively, local temporary expression of clot-disintegrating enzymes such as streptokinase and lipases may repair damaged and diseased arteries. It is quite possible to deliver liposomes loaded with DNA or, in fact, directly inject DNA plasmids into the tissue, and the protein will be expressed by the cells which take it up. Only 1–3% will do so, but this is sufficient for the local effect required and it is a transient expression. This gives a controllable gene therapy.

Neuronal disease

Neurotrophic factors can be transiently expressed as described above for vascular diseases. The local expression of neurotrophins is essential for nerve cell regeneration and maintenance. It is possible to extend the expression period of the neurotrophin by injecting transfected myocytes into the damaged area. They will fuse with any adjacent muscle tissue to give a prolonged expression of the factor gene.

Cancer

Cancer is a genetic disease and many genes are deregulated. *p53* is a tumour suppressor gene; the reintroduction and overexpression of a functional *p53* in tumours is being investigated with some success. Since *p53* will induce apoptosis in cells with damaged genetic material, the transient expression of high levels in a tumour cell with *p53* pathway defects should induce its own (apoptotic) demise. Since this is only likely to occur in rapidly dividing cells, this is a perfect target for cancer gene therapy. Transient expression induced by repeat exposure to vectors such as retroviruses, liposomes and naked DNA plasmids is all relatively straightforward. Trials using aerosols of these vectors in patients with lung cancer are being performed.

Tumour growth depends on the development of new blood vessels (angiogenesis), and inhibitors of this process are also being used in trials.

Stem cell therapy

It is claimed that stem cell therapy has the potential to radically change the treatment of human disease. A number of adult stem cell therapies already exist, particularly bone marrow transplants. It is currently anticipated that technologies derived from stem cell research can be used to treat a wider variety of diseases in which replacement of destroyed specialist tissues is required, such as in Parkinson's disease, spinal cord injuries and muscle damage.

The potency of the stem cells used in therapy is therefore of great importance. In a developing embryo, embryonic stem cells can differentiate into all of the specialized embryonic tissues. In adult organisms, adult stem cells (and progenitor cells) act as a repair system for the body, replenishing specialized cells. In 2001 several studies reported on the plasticity of human embryonic stem cells (HESCs). With respect to repair of nerves, HESCs can be induced to differentiate into the three neural lineages (astrocytes, oligodendrocytes and mature neurons) in vitro. Furthermore, when transplanted into new born (mouse) brains they migrated along established neuronal tracts and differentiated to the appropriate cell type in a region-specific manner (i.e. responded to local paracrine signals and cues – *totipotentiality*).

The blood within the umbilical cord is a readily available and rich source of haemopoietic stem cells (i.e. CD34-positive and CD38-negative). These cord blood stem cells readily colonize the bone marrow, rapidly populating the marrow with all the various cells (erythrocytes and leukocytes) required of a fully functional blood system – *pluripotentiality*.

Umbilical cord blood is proving an invaluable source of stem cells for bone marrow transplants. Cord blood banks have been set up where collected cord blood or indeed the clamped and early cut umbilical cord itself, is cryogenically preserved for long-term (and costly) storage should the child or a tissue-type-matched individual ever require the stem cells. In 2005 researchers claimed to have discovered a category of cord stem cell, dubbed *cord-blood-derived embryonic like stem cells* (CBEs), which were able to differentiate into more types of tissue than simply haemopoietic cells (super pluripotentiality). Similarly, although peripheral blood monocytes are believed to be the unipotent precursors of phagocytes, in 2007 it was claimed that a primitive *monocyte-derived multipotential cell* (MOMC) could be isolated from adult peripheral circulating monocytes. Furthermore, these MOMCs can be induced (given the correct paracrine, environmental and adhesion signals) into endothelia, neurones, cardiomyocytes and mesenchymal lineages. Similar reports have appeared concerning adult stem cells isolated from skin.

There is a great deal of social, as well as scientific controversy surrounding stem cell research. Human embryonic stem cell research is particularly controversial as the establishment of a stem cell line requires the destruction of a human embryo and/or therapeutic cloning. However, recently, it has been shown that embryonic stem cell lines can be generated without destroying embryos using single-cell biopsy, similar to the approach used in pre-implantation genetic diagnosis. This, in combination with recent results discussed above concerning the isolation of cord blood and adult totipotent (or super pluiripotent) stem cells, may circumvent the ethical blockade on using HESCs in stem cell therapy.

Creating and using animal models

There is a need for model systems in which to test gene therapies prior to use in patients. To some extent gene 'knock-out' experiments in mice are providing new animal models of disease. In these experiments the normal gene of interest in a mouse is targeted using recombinant DNA techniques so that gene function is impaired, in analogy to the situation in the particular gene of a human patient. However, mice and humans are different and it will not be possible to

mimic some human genetic diseases in mice. Transgenic mice are created when exogenous DNA carrying a gene of interest is injected into a mouse egg. If this egg is fertilized then all the cells of the resulting animal will carry the extra gene sequences. Transgenic mice have also been used as animal models of human diseases in which new therapies may be tested.

Other animal models have paved the way for gene therapy techniques. For example, some of the first experiments in the transfer of globin genes (which will be useful for gene therapy for sickle cell disease and thalassaemia) have taken place in mice. These include transplanting normal donor cells with normal genes into lethally irradiated mice, which results in engraftment of the donor cells.

The human proteome project

A more direct route to understanding genetic and somatic disease is by studying the protein expression characteristics of normal and diseased cells – the proteome. This relies on the separation of proteins expressed by a given tissue by molecular size and charge on a simple two-dimensional display and is achieved by using two-dimensional gel electrophoresis. The pattern of dots corresponds to the different proteins expressed. With the improvement in technology, the patterns are reproducible and can be stored as electronic images. Non-, over- and underexpression of a given protein can be detected by a corresponding change on the proteome two-dimensional electrophoresis image. As can be seen in Figure 2.20, the protein profile of synovial fluid from a diseased joint shows a new protein arising with disease. Such proteins have been identified as inflammatory cytokines. Furthermore, post-translational modifications of the protein show up as a change in either size or charge on the proteome picture. In order to positively identify the altered protein and the post-translational modification it may contain, these protein spots are eluted and subjected to modern mass spectrometry techniques such as *matrix-assisted laser desorption ionization* (MALDI) and *electrospray ionization* (ESI) time of flight (TOF), which not only give the precise mass of the protein, up to ~500 000 Da, but can also sequence its amino acid, phosphorylation and glycosylation structure. This cannot be detected by genome analysis. Looking for such changes has already led to the discovery of new protein markers for the diagnosis of Creutzfeldt–Jakob disease, multiple sclerosis, schizophrenia, Parkinson's disease (spinal fluid protein) and Alzheimer's disease (blood and brain proteins). In 2000, the Swiss Institute of Bioinformatics (SIB) and the European Bioinformatics Institute (EBI) announced a major effort to annotate, describe and provide highly accurate information concerning human protein sequences, effectively launching the Human Proteomics Initiative (HPI).

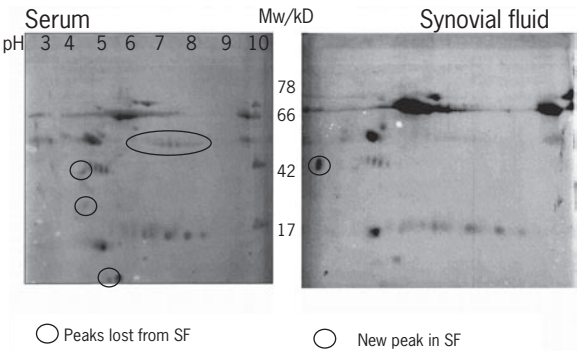


Fig. 2.20 **Two-dimensional gel electrophoresis comparing paired serum and synovial fluid in a patient with rheumatoid arthritis.** Biofluids were first separated according to isoelectric point on a linear immobilized pH gradient followed by SDS-PAGE in a 10% gel. The circled proteins indicate major proteins which differ between the two biofluids. Although serum contained many proteins not found in synovial fluid, one major protein was found to be present in the synovial fluid electropherogram but not in the serum electropherogram. This indicates that synovial fluid is not a simple transudate of serum. Courtesy of Prof. D Perrett and Dr R Bevan, Department of Medicine, Barts and The London School of Medicine.

Metabolomics

In the post-genomic era, computing power, statistical software, separation science and modern mass spectrometry have allowed the analysis of a complex mixture as a complete entity, and not merely the fluctuation in concentration of one analyte within it. Metabolomics is the study of the repertoire of non-proteinaceous, endogenously synthesized small molecules present in an organism. Such small molecules include well-known compounds like glucose, cholesterol, ATP and lipid signalling molecules. These molecules are the ultimate product of cellular metabolism, and the metabolome refers to the catalogue of those molecules in a specific organism, e.g. the human metabolome. In terms of clinical biochemistry, the analysis of the pattern of change of such molecules in urine samples of individuals with and without a particular disease and those treated with specific drugs represents a change in the metabolome. It is very likely that, in the future, medicine-regulating authorities will require metabolomic studies on all new drugs.

Ethical considerations

Ethical considerations must be taken into account in any discussion of clinical genetics. For example, prenatal diagnosis with the option of termination may be unacceptable on moral or religious grounds. With diseases for which there is no cure and currently no treatment (e.g. Huntington's), genetic tests can predict accurately which family members will be affected; however, many people would rather not know this information. One very serious outcome of the new genetic information is that disease susceptibility may be predictable, for example in Alzheimer's disease, so the medical insurance companies can decline to give policies for individuals at high risk.

Society has not yet decided who should have access to an individual's genetic information and to what extent privacy should be preserved.

FURTHER READING

Cyranoski D. Simple switch turns cells embryonic. *Nature* 2007; **447**: 618–619.

Goldman S, Windrem M. Cell replacement therapy in neurological disease. *Philosophical Transactions of the Royal Society of London: B, Biological Sciences* 2006; **361**: 1463–1475.

Holt I (ed). *Genetics of Mitochondrial Diseases.* Oxford: Oxford University Press, 2003.

Kiechle FL, Holland-Staley CA. Genomics, transcriptomics, proteomics, and numbers. *Archives of Pathology and Laboratory Medicine* 2003; **127**: 1089–1097.

Onyango P. The role of emerging genomics and proteomics technologies in cancer drug target discovery. *Current Cancer Drug Targets* 2004; **4**: 111–124.

Super M. CFTR and disease: implications for drug development. *Lancet* 2000; **355**: 1840–1842.

Wang S, El-Deiry WS. The p53 pathway: targets for the development of novel cancer therapeutics. *Cancer Treatment and Research* 2004; **119**: 175–187.

THE GENETIC BASIS OF CANCER

Cancers are genetic diseases and involve changes to the normal function of cellular genes. However, multiple genes interact during oncogenesis and an almost stepwise progression of defects leads from an overproliferation of a particular cell to the breakdown of control mechanisms such as apoptosis (programmed cell death). This would be triggered if a cell were to attempt to survive in an organ other than its tissue of origin. For the vast majority of cancer cases (especially those in older people) the multiple genetic changes which occur are somatic. However, it is clear that susceptibility to the development of a particular form of cancer can be inherited. Indeed for some rare cancers a dominant single-gene defect can give rise to an almost Mendelian trend. In some other cancers (e.g. forms of breast cancer) the mode of inheritance is much more complex. In these cases close relatives may have an increased susceptibility to cancer (Table 9.3), and the genetics are clearly those of a multifactorial trait.

- Cancer tissues are clonal, and tumours arise from changes in only one cell, which then proliferates in the body.
- The genes that are primarily damaged by the genetic changes which lead to cancer fall into two categories: oncogenes and tumour suppressor genes.
- Oncogenesis is a multistep process in that a number of mutations or alterations to key genes are required before a malignant phenotype is expressed.
- Once mutations have begun to cause unchecked clonal expansion of the primary tumour cells, further mutations occur within the subsequent generations of daughter cells that give rise to clones which are invasive and/or form metastases.

ONCOGENES (see p. 451)

The genes coding for growth factors, growth factor receptors, secondary messengers or even DNA-binding proteins would act as promoters of abnormal cell growth if mutated. This concept was verified when viruses were found to carry genes which, when integrated into the host cell, promoted oncogenesis. These were originally termed viral or 'v-oncogenes', and later their normal cellular counterparts, c-oncogenes, were found. Thus, oncogenes encode proteins that are known to participate in the regulation of normal cellular proliferation e.g. *erb-A* on chromosome 17q11–q12 encodes for the thyroid hormone receptor. See also Table 9.4.

Activation of oncogenes

Non-activated oncogenes which are functioning normally have been referred to as 'proto-oncogenes'. Their transformation to oncogenes can occur by three routes.

Mutation

Carcinogens such as those found in cigarette smoke, ionizing radiation and ultraviolet light can cause point mutations in genomic DNA. By chance some of these point mutations will occur in regions of the oncogene which lead to activation of that gene. Not all bases in an oncogene cause cancer if mutated, but some (e.g. those in the coding region) do.

Chromosomal translocation

If during cell division an error occurs and two chromosomes translocate, so that a portion swaps over, the translocation breakpoint may occur in the middle of two genes. If this happens then the end of one gene is translocated on to the beginning of another gene, giving rise to a 'fusion gene'. Therefore sequences of one part of the fusion gene are inappropriately expressed because they are under the control of the other part of the gene.

An example of such a fusion gene (the Philadelphia chromosome) occurs in *chronic myeloid leukaemia* (CML, see p. 472). Similarly in *Burkitt's lymphoma* a translocation causes the regulatory segment of the *myc* oncogene to be replaced by a regulatory segment of an unrelated immunoglobulin.

Viral stimulation

When viral RNA is transcribed by reverse transcriptase into viral cDNA and in turn is spliced into the cellular DNA, the viral DNA may integrate within an oncogene and activate it. Alternatively the virus may pick up cellular oncogene DNA and incorporate it into its own viral genome. Subsequent infection of another host cell might result in expression of this viral oncogene. For example, the Rous sarcoma virus of chickens was found to induce cancer because it carried the *ras* oncogene.

After the initial activation event other changes occur within the DNA. A striking example of this is amplification of gene sequences, which can affect the *myc* gene for example. Instead of the normal two copies of a gene, multiple copies of the gene appear either within the chromosomes (these can be seen on stained chromosomes as homogeneously staining regions) or as extrachromosomal particles (double minutes). *N-myc* sequences are amplified in neuroblastomas, as are *N-myc* or *L-myc* in some lung small-cell carcinomas.

TUMOUR SUPPRESSOR GENES

These genes restrict undue cell proliferation (in contrast to oncogenes), and induce the repair or self-destruction (apoptosis) of cells containing damaged DNA. Therefore mutations in these genes which disable their function lead to uncontrolled cell growth in cells with active oncogenes. An example is the germline mutations in genes found in non-polyposis colorectal cancer which are responsible for repairing DNA mismatches (p. 302).

The first tumour suppressor gene to be described was the *RB* gene. Mutations in *RB* lead to retinoblastoma which occurs in 1 in 20 000 young children and can be sporadic or familial. In the familial variety, the first mutation is inherited

The genetic basis of cancer
Oncogenes
Tumour suppressor genes

and by chance a second somatic mutation occurs with the formation of a tumour. In the sporadic variety, by chance both mutations occur in both the *RB* genes in a single cell. Since the finding of *RB*, other tumour suppressor genes have been described, including the gene *p53*. Mutations in *p53* have been found in almost all human tumours, including sporadic colorectal carcinomas, carcinomas of breast and lung, brain tumours, osteosarcomas and leukaemias. The protein encoded by *p53* is a cellular 53 kDa nuclear phosphoprotein that plays a role in DNA repair and synthesis, in the control of the cell cycle and cell differentiation and programmed cell death – apoptosis. p53 is a DNA-binding protein which activates many gene expression pathways but it is normally only short-lived. In many tumours, mutations that disable *p53* function also prevent its cellular catabolism. Although in some cancers there is a loss of *p53* from both chromosomes, in most cancers (particularly colorectal carcinomas; see Fig. 6.45) such long-lived mutant *p53* alleles can disrupt the normal alleles' protein. As a DNA-binding protein, p53 is likely to act as a tetramer. Thus, a mutation in a single copy of the gene can promote tumour formation because a hetero-tetramer of mutated and normal p53 subunits would still be dysfunctional.

How tumour suppressor genes work

Tumour suppressor gene products are intimately involved in control of the cell cycle (see Fig. 2.7). Progression through the cell cycle is controlled by many molecular gateways, which are opened or blocked by the cyclin group of proteins that are specifically expressed at various stages of the cycle. The RB and p53 proteins control the cell cycle and interact specifically within many cyclin proteins. The latter are affected by INK 4α acting on p16 proteins. The general principle is that being held at one of these gateways will ultimately lead to programmed cell death. p53 is a DNA-binding protein which induces the expression of other genes and is a major player in the induction of cell death. Its own expression is induced by broken DNA. The induction of *p53* by damage initially causes the expression of DNA repair enzymes. If DNA repair is too slow or cannot be effected, then other proteins that are induced by *p53* will effect programmed cell death.

One gateway event that has been largely elucidated is that between the G1 and the S phase of the cell cycle. The transcription factor dimer E2F-DP1 causes progression from the G1 to the S phase. However, the RB protein binds to the E2F transcription factor, preventing its induction of DNA synthesis. Other, cyclin D-related molecules inactivate the RB protein, thus allowing DNA synthesis to proceed. This period of rapid DNA synthesis is susceptible to mutation events and will propagate a pre-existing DNA mistake. Damaged DNA-induced p53 expression rapidly results in the expression of a variety of closely related (and possibly tissue-specific) proteins WAF-1/p21, p16, p27. These inhibit the inactivation of RB by cyclin D-related molecules. As a result RB, the normal gate which stops the cell cycle, binds to the E2F-DP1 transcription factor complex, halting S phase DNA synthesis. If the DNA damage is not repaired apoptosis ensues.

Viral inactivation of tumour suppressors

The suppression of normal tumour suppressor gene function can be achieved by disabling the normal protein once it has been transcribed, rather than by mutating the gene. Viruses have developed their own genes which produce proteins to do precisely this. The main targets of these proteins are RB and p53 to which they bind and thus disable. The best understood are the adenovirus E1A and human papillomavirus (HPV) E7 gene products which bind RB, whilst the adenovirus E1B and HPV E6 gene products bind p53. The SV40 virus large T antigen binds both RB and p53.

Microsatellite instability

Microsatellites are short (50–300 bp) sequences composed of tandemly repeated segments of DNA two to five nucleotides in length (dinucleotide/trinucleotide/tetranucleotide repeats). Scattered throughout the genome in the non-coding regions between genes or within genes (introns), many of these microsatellites are highly polymorphic. Often used as markers for linkage analysis because of high variability in repeat number between individuals, these regions are inherently unstable and susceptible to mutations. *Somatic microsatellite instability* (MSI) has been detected in a number of tumours. Detecting MSI involves comparing the length of microsatellite alleles amplified from tumour DNA with the corresponding allele in normal tissue from the same individual. Recent studies indicate that MSI can be detected in approximately 90% of tumours from individuals with hereditary non-polyposis colorectal cancer (HNPCC). The presence of these additional microsatellite alleles (repeated segments) in tumour cells results from the inherent susceptibility of these areas to such alterations and from mutations in the DNA mismatch repair mechanism that would normally correct these errors.

Tumour angiogenesis

Once a nest of cancer cells reaches 1–2 mm in diameter, it must develop a blood supply in order to survive and grow larger, as diffusion is no longer adequate to supply the cells with oxygen and nutrients. As with all tissues, solid tumour cancer cells secrete substances that promote the formation of new blood vessels – a process called angiogenesis. Over a dozen substances have been identified that promote angiogenesis (e.g. angiopoietin-1, basic fibroblast growth factor (bFGF) and vascular endothelial growth factor (VEGF)). This has led to the discovery of a number of inhibitors of angiogenesis and some have already advanced to clinical trials as part of a cancer treatment strategy such as:

- angiostatin – a polypeptide of approximately 200 amino acids, produced by the cleavage of plasminogen; it binds to subunits of ATP synthase exposed at the surface of the cell embedded in the plasma membrane
- endostatin – a polypeptide of 184 amino acids which is derived from the globular domain found at the C-terminal of type XVIII collagen (a specific collagen of blood vessels) cleaved from the parent molecule.

Several therapeutic vaccine preparations are under development to produce a range of host immunity responses (humoral, cellular) against pro-angiogenic factors and their receptors in tumours. One approach has been directed at cell adhesion molecules found in tumour blood vessels. It turns out that the new blood vessels in tumours express a vascular integrin – designated alpha-v/beta-3 – that is not found on the old blood vessels of normal tissues. Vitaxin, a monoclonal antibody directed against the alpha-v/beta-3 vascular integrin, shrinks tumours in mice without harming them. In phase II clinical trials in humans, vitaxin has shown some promise in shrinking solid tumours without harmful side-effects.

CANCER AETIOLOGY: INHERITANCE OR ENVIRONMENT?

It is clear that a mutation causing the dysfunction of a single oncogene or tumour suppressor gene is not sufficient to induce unregulated clonal expansion. The *Knudson multi-hit hypothesis* elegantly unites the genetics of familial and sporadic tumour development: an inherited mutation in one gene allele may be insufficient to cause a tumour but will cause a significant susceptibility to the development of a particular cancer. Subsequent lifetime exposure to environmental carcinogens (viral, chemical, radiation), along with simple mistakes during cell division, may deregulate the normal allele. Other mutations which accumulate in a similar manner then lead to tumour development. Research has clearly shown that germline mutations in particular genes, such as *p53* and *RB*, have a much stronger influence on the chance of subsequent tumour development than others.

Epigenetics and cancer

The term 'epigenetics' is used to explain changes in gene expression that do not involve changes in the underlying DNA sequence. Despite not altering the decoding sequence, the effects of epigenetic changes are stable over rounds of cell division, and sometimes between generations. There are two principal molecular epigenetic mechanisms implicated in cancer:

- Modifications to DNA's surface structure, but not its base pair sequence – DNA methylation resulting in non-recognition of gene transcription DNA-binding domains.
- Modification of chromatin proteins (in particular histones), which will not only support DNA but bind it so tight as to regulate gene expression – at the extreme, such binding can permanently prevent the DNA sequences being exposed to, let alone acted on by, gene transcription (DNA-binding) proteins.

Hypermethylation of gene regulatory regions (denoted by CpG islands, cytosine, guanine nucleotides joined by a phosphodiester bond (p. 31)) whereby cytosine is converted to 5-methyl cytosine has been identified at a number of gene loci at the 5′ regulatory region and is associated with oncogenesis. The *Hic-1* gene is noticeably silenced by hypermethylation in numerous tumours. The enzymes responsible for methylation, DNA methyltransferases (e.g. DNMT1), are attracted to replication foci and preferentially modify hemimethylated DNA CpG regions. The net result is not only maintenance of the methylated status in daughter strands but also its expansion within the loci.

Methylation is not limited to CpG-rich regions of DNA – the associated histones are also modified. The lysine-rich tail of histone protein 3 (H3) interacts with DNA by attracting and compensating for the strong polar negative charge–negative charge repulsion of the phosphate backbone of DNA. Effectively it buffers this repulsion, allowing DNA to twist and coil tighter. Acetylation of the H3 tail lysines removes their buffering positive charge, causing the DNA to spring away from the histone packaging. In this histone-influenced free form, the DNA helix is able to assume conformational shapes driven by its own internal bond length strains. Thus, where the double helix is under stress from distortion by CpG-rich regions, this springing away from the nucleosomal histones allows access to DNA-binding proteins that initiate transcription. Deacetylation of the H3 tail lysines results in DNA being pulled onto the histone body and effectively burying the CpG-rich transcription recognition regions. However, methylation of the H3 tail lysines results in even tighter coiling and like-with-like attraction of similar modified histones of other nucleosomes. The net result is the tight packing of large regions of DNA in what is termed heterochromation. Thus DNA methylation can recruit DNMTs to methylate daughter strand loci during replication and their associated histones to effectively shut a gene region. If the region includes tumour repressor gene loci, oncogenic progression can proceed without direct mutation of the gene in question.

More epigenetic mechanisms are being discovered that further help explain the phenomenon of 'book marking' which has been observed in several biological species. The name refers to the observation that some genes appear to remain active long after the initiating signals for transcription have been removed – much like a book falling open at a much-read page once dropped. Recent studies have focused on small RNA species that are the product (or spliced byproduct) of transcriptional activated genes. These RNAs tend to form double-stranded species that keep the gene loci open for continued transcription. This is not an unusual phenomenon in that the protein product of genes such as *Hnf4* and *MyoD* is known to feedback and maintain autotranscription. However, small RNA species can freely diffuse through nuclear pores and into other cells via gap junctions. This unhindered passage makes such RNA species pervasive, upregulating and locking the on switch of oncogenic genes in transformed and neighbouring cells without DNA mutation. In several species the so-called 'maternal effect phenotype' is believed to be due to such small RNA species being laid down in the oocyte by nurse cells and then acting on genes of the resulting offspring as it develops.

DNA methylation is low in tumour cells compared to normal tissue.

CHAPTER BIBLIOGRAPHY

Lodish H, Berk A, Kaiser CA et al. *Molecular Cell Biology,* 6th edn. New York: WH Freeman, 2007.

McKusick VA. *Mendelian Inheritance in Man,* 12th edn. London: Johns Hopkins Press, 1998.

Nabel GJ. Genetic, cellular and immune approaches to disease therapy: past and future. *Nature Medicine* 2004; **10**: 135–141.

Weatherall DJ. *Basic Molecular and Cell Biology,* 3rd edn. London: BMJ Publishing, 1996.

SIGNIFICANT WEBSITES

http://www.genetichealth.com
US e-health company

FURTHER READING

Croce CM. Molecular origins of cancer: oncogenes and cancer. *New England Journal of Medicine* 2008; **358**: 502–511.

Esteller M. Epigenetics in cancer. *New England Journal of Medicine* 2008; **358**: 1148–1159.

Fryer A, Newbury-Ecob R. Genetics. *Clinical Medicine* 2008; **8**(6).

Gearhart J, Pashos EE, Prasad MK. Pluripotency redux – advances in stem-cell research. *New England Journal of Medicine* 2007; **357**: 1469–1472.

Slee EA, O'Connor DJ, Lu X. To die or not to die: how does p53 decide? *Oncogene* 2004; **23**: 2809–2818.

This page intentionally left blank

3 The immune system and disease

Anatomy and principles of the immune system 53
Cells and molecules of the innate immune system 55
HLA molecules and antigen presentation 59
Cells and molecules of the adaptive immune system 62
Cell migration 66
The immune system in concert 68
Laboratory investigations of the immune system 69
Clinical immunodeficiency 69
Type I hypersensitivity and allergic disease 72
Autoimmune disease 73
Organ rejection in clinical transplantation 75
Immune-based therapies 75

ANATOMY AND PRINCIPLES OF THE IMMUNE SYSTEM

Immunity can be defined as protection from infection, whether it be due to bacteria, viruses, fungi or multicellular parasites. Like other organs involved in human physiology, the immune system is composed of cells and molecules organized into specialized tissues (Fig. 3.1).

The primary lymphoid organs are where the cells originate and develop into immature forms. Cells and molecules of the immune system circulate in the blood, but immune responses do not take place there. Rather, an immune response is initiated at the site of infection (typically the mucosa or skin) and then propagated and refined in the *secondary lymphoid organs* (e.g. lymph nodes). After resolution of the infection, immunological memory specific for the pathogen resides in cells (lymphocytes) in the spleen and lymph nodes as well as being widely secreted in a molecular form (antibodies).

Development of cells involved in immune responses

Cells destined to be myeloid cells or lymphocytes are derived from a common pluripotent stem cell in the bone marrow.

Myeloid cells. Neutrophils, eosinophils, basophils and dendritic cells are derived from precursor cells generated in the bone marrow (see Fig. 8.1).

T lymphocytes. Immature cells leave the bone marrow to enter the thymus as thymocytes. Gene rearrangement takes place to form the antigen-specific T cell receptor. Positive and negative selection occur to ensure that functional cells that are non-auto-reactive survive and differentiate to $CD4^+$ or $CD8^+$ cells. These antigen naïve T cells leave the thymus to populate peripheral lymphoid tissue. On encountering their specific antigen they proliferate and activation occurs to give them either helper ($CD4^+$) or cytocytoxic ($CD8^+$) cell function (see p. 65).

B lymphocytes undergo develoment to maturity within the bone marrow, including gene rearrangement for production of antigen-specific antibody molecules. The mature cells are released and populate the follicles of lymph nodes and lymphoid tisue.

Lymphoid tissue is frequently found in mucosal surfaces in non-encapsulated patches. This is termed mucosa associated lympohoid tissue (MALT), consisting of genitourinary and gut-associated lymphoid tissue (GALT, mainly Peyer's patches), bronchus-associated lymphoid tissue (BALT, found in the lobes of the lungs along the main bronchi) and skin-associated lymphoid tissue (SALT).

On meeting their specific antigen, B cells proliferate and differentiate to antibody-secreting plasma cells.

The immune system

Cells and molecules involved in immune responses are classified into **innate** and **adaptive** systems. There are also non-immunological barriers that are involved in host protection, and very often it is the lowering of these that allows a pathogen to take a foothold (Table 3.1).

The immune system is immensely powerful, in terms of its ability to inflame, damage and kill, and it has a capacity to recognize a myriad of molecular patterns in the microbial world. However immune responses are not always beneficial. They can give rise to a range of autoimmune and inflammatory diseases, known as immunopathologies. In addition, the immune system may fail, giving rise to immune deficiency states. These conditions are grouped under the umbrella of clinical immunology.

A major feature of the immune system is the complexity of the surface-bound, intracellular and soluble structures that mediate its functions. In particular it is necessary to be aware of the CD (clusters of differentiation)

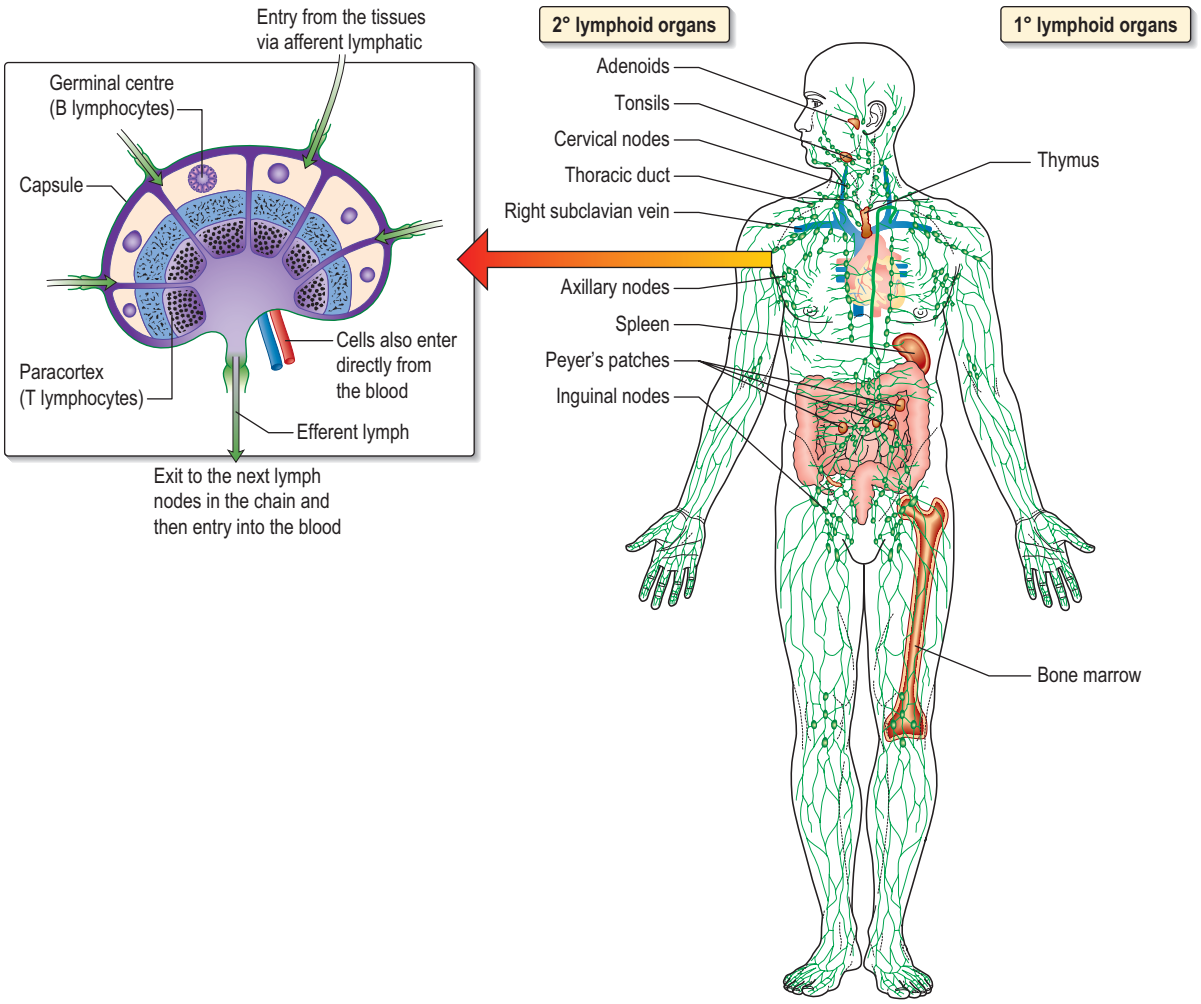


Fig. 3.1 **Primary and secondary lymphoid tissue.** Lymphocytes are generated as precursors in the bone marrow and differentiate into T (thymus) or B (bone marrow) lymphocytes in the primary lymphoid tissue. Once differentiated, 98% of lymphocytes reside in the secondary lymphoid tissue where the adaptive immune response takes place.

Table 3.1 Non-immunological host defence mechanisms

Normal barriers	Events that may compromise barrier function
Physical barriers	
Skin and mucous membranes	Trauma, burns, i.v. cannula
Cough reflex	Suppression, e.g. by opiates, neurological disease
Mucosal function	Ciliary paralysis (e.g. smoking) Increased mucus production (e.g. asthma) Abnormally viscid secretions (e.g. cystic fibrosis) Decreased secretions (e.g. sicca syndrome)
Urine flow	Stasis (e.g. prostatic hypertrophy)
Chemical barriers	Low gastric pH (gastric acid secretion inhibitors)
Resistance to pathogens provided by commensal skin and gut organisms	Changes in flora (e.g. broad-spectrum antibiotics)

classification (Box 3.1) and the functions of cytokines and chemokines.

Cytokines

These are small polypeptides released by a cell in order to change the function of the same or another cell. These chemical messengers are found in many organ systems, but especially the immune system. Cytokines have become markers in the investigation of disease pathogenesis; therapeutic agents in their own right; and the targets of therapeutic agents (see p. 75). The key features of a cytokine are *pleiotropy* (different effects on different cells); *autocrine* function (modulates the cell secreting it); *paracrine* function (modulates adjacent cells); *endocrine* effects (modulates cells and organs at remote sites); and *synergy* (acting in concert to achieve effects greater than the summation of their individual actions). The main immune cytokines are the interferons (IFNs) and the interleukins (IL). The IFNs are limited to a few major types (α, β and γ) whilst IL-35 is the latest cytokine to be defined (for listings and functions see http://www.copewithcytokines.de/).

Box 3.1 The CD classification

This is the ultimate way of defining a cell.

- Immune cells are distinguished by the surface receptors and molecules that they express in order to mediate their particular range of immunological functions, e.g. cell–cell signalling, cell activation
- The surfaces are covered with such proteins and indicate the cell lineage or differentiation pathway. The discovery of monoclonal antibodies (proteins tailormade to bind to a specific target) made this feasible
- Surface molecules defining the origin and function of selected groups of cells are known as clusters of differentiation (CD). Over 300 CD numbers exist
- In clinical practice, e.g. the number of peripheral blood lymphocytes expressing CD4 ('the CD4 count') are used to monitor HIV infection (see p. 196)
- An updated listing is available at http://www.hcdm.org/

Chemokines

The defining feature of chemokines is their function as chemotactic molecules – that is, they attract cells along a gradient of low to high chemical concentration. Chemokines are involved in the trafficking of cells of the immune system – particularly from the blood into the tissues and tissues into lymphatics – and also have the ability to activate immune cells. Like cytokines, chemokines are pleiotropic. All chemokines have a similar structure relating to the configuration of cysteine residues that gives rise to four families:

- Two cysteines separated by any other amino acid residue (X) is CXC.
- Two cysteines next to each other is CC.
- One cysteine is C.
- Two cysteines separated by any three amino acids are CX3C.

Receptors on the surface are denoted by 'R', and are all distinctive G-coupled protein receptors with seven transmembrane spanning domains. Some chemokines are released by many nucleated cells in response to inflammation and have a role in chemotaxis, attracting immune cells to inflammatory foci (e.g. CXCL8 (interleukin 8)); others are released in a programmed fashion by specialized cells in the immune system and have specific functions. CXCL12, for example, is necessary for the development of *secondary lymphoid organs*.

Anti-inflammatory drugs that block chemokine functions, or drugs that promote cell activation and migration, for example into tumours, are being developed (for listings and functions see http://www.copewithcytokines.de/).

CELLS AND MOLECULES OF THE INNATE IMMUNE SYSTEM

Innate immunity provides immediate, first line, host defence. The key features of this system are shown in Table 3.2. It is present at birth and remains operative at comparable intensity into old age. Innate immunity is mediated by a variety of cells and molecules (Table 3.3). Activation of innate immune responses is mediated through interaction between the:

Table 3.2 Features of the innate and adaptive immune responses

Innate	Adaptive
No memory: quality and intensity of response invariant	Memory: response adapts with each exposure
Recognizes limited number of non-varying, generic molecular patterns on, or made by, pathogens	Recognizes vast array of specific antigens* on, or made by, pathogens
Pattern recognition mediated by a limited array of receptors	Antigen recognition mediated by a vast array of antigen-specific receptors
Response immediate on first encounter	Response on first encounter takes 1–2 weeks; on second encounter 3–7 days

*Antigen is a molecular structure (protein, peptide, lipid, carbohydrate) that generates an immune response

- pathogen side comprising a relatively limited array of molecules (pathogen-associated molecular patterns (PAMPs))
- host side – a limited portfolio of receptors (pattern recognition receptors (PRRs)).

Activation of certain cells in the innate immune system leads to activation of the *adaptive immune response*.

The *dendritic cell* is especially involved in this process, and forms a bridge between innate and adaptive systems.

Complement

The complement cascade is activated through three distinct pathways, alternative, classical and mannan-binding lectin. These pathways are composed of three distinct enzyme cascades that culminate in the cleavage of C3 and C5. Cleavage of C3 produces a number of biological consequences, while breakdown of C5 achieves the same and, in addition, provides the triggering stimulus to the final common ('membrane attack') pathway, which provides most of the biological activity (Fig. 3.2).

The main functions of complement activation are to:

- promote inflammation (e.g. through the actions of the anaphylatoxins)
- recruit cells (e.g. through chemoattractants)
- kill targeted cells, such as bacteria
- to solubilize and remove from the circulation antigen–antibody ('immune') complexes.

During an immune response, removal of immune complexes protects unaffected tissues from the deposition of these large, insoluble composites which could result in unwanted inflammation. Failure of this protective mechanism can result in immunopathology, for example in the joints, kidney and eye.

Neutrophils (p. 430)

Neutrophils phagocytose and kill microorganisms. They have a distinctive multilobed nucleus (hence the alternative name 'polymorph') and neutral staining granules. Neutrophils come from the bone marrow, which can produce between

Table 3.3 Components of innate immunity

Type	Name (examples)	Function	Provide immunity against:
Molecules			
Complement	Cascade of >40 proteins	Lyse bacteria; opsonize* bacteria; promote inflammation; recruit and activate immune cells	Bacteria, viruses
Collectins	Mannan binding lectin	Binds bacteria; activates complement	Bacteria
Pentraxins	C-reactive protein†	Opsonize bacteria	Bacteria
Enzymes	Lysozyme	Present in secretions; cleaves bacterial cell wall	Bacteria
Cells			
Granulocytes	Neutrophil	Phagocytose and kill bacteria; release antibacterials (eg defensins).	Bacteria, fungi
	Eosinophils	Release proinflammatory mediators	Parasites (especially multi-cellular)
	Basophils	Release proinflammatory mediators	Unknown
	Mast cells	Release soluble proinflammatory molecules (e.g. histamine); tissue resident	Parasites
Mononuclear	Monocyte (present in blood) and macrophage (present in tissues)	Ingest and kill bacteria; release soluble proinflammatory molecules (e.g. IL-1, TNF-α, IFN-γ); present antigen to T lymphocytes	Intracellular pathogens
Dendritic	Myeloid dendritic cell	Ingest pathogens; migrate to lymph nodes and present antigen to T lymphocytes	All pathogens via interaction with adaptive immune system
	Plasmacytoid dendritic cell	Migrate to sites of inflammation including lymph nodes; release interferons and present antigen to T lymphocytes	Pathogens (especially viruses) via interaction with adaptive immune system
Lymphoid	Natural killer cells	Lyse infected cells	Viruses

*Opsonize = coat bacteria to enhance phagocytosis by granulocytes and monocytes/macrophages.
†C-reactive protein (CRP) is an acute phase protein. Blood levels rise 10–100 fold within hours of the start of an infective or inflammatory process, making it extremely useful in monitoring infective or inflammatory diseases, and their response to treatment.

10^{11} (healthy state) and 10^{12} (during infection) new cells per day. In health, neutrophils are rarely seen in the tissues. There are approximately 100 different molecules in neutrophil granules (azurophilic and specific), the key ones and their functions being shown in Table 3.4.

Neutrophils are phagocytic cells activated by interaction with bacteria, either directly or when these are coated to make them more ingestible (termed opsonization (Fig. 3.3)). Granule contents are released both intracellularly (predominantly azurophilic granules) and extracellularly (specific granules) following fusion with the plasma membrane. Granule contents kill and digest microorganisms.

- *Myeloperoxidase and cytochrome b_{558}* are each key components of major oxygen-dependent bactericidal systems (called the hydrogen peroxide–myeloperoxidase –halide system and the respiratory burst, respectively).
- *Cathepsins, proteinase-3 and elastase* are deadly to Gram-positive and Gram-negative organisms, as well as some *Candida* species.
- *Defensins* are naturally occurring cysteine-rich antibacterial and antifungal polypeptides (29–35 amino acids).
- *Collagenase and elastase* break down fibrous structures in the extracellular matrix, facilitating progress of the neutrophil through the tissues.

Granule release is initiated by the products of bacterial cell walls, complement proteins (e.g. inactive complement 3b, iC3b), leukotrienes (LTB_4) and chemokines (e.g. CXCL8, also known as interleukin 8 (IL-8)) and cytokines such as tumour necrosis factor α (TNF-α).

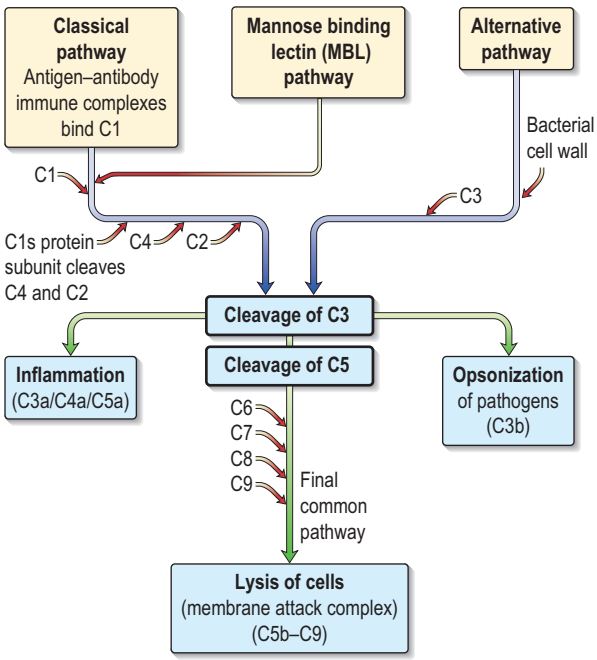


Fig. 3.2 **The complement pathway and its effector functions.**

Eosinophils

In contrast to neutrophils, several hundred times more eosinophils are present in the tissues than in the blood. They collect preferentially at epithelial surfaces and survive for

Table 3.4 Contents and function of neutrophil granules

Function	Primary or azurophilic granules	Secondary or specific granules
Antibacterial	Lysozyme Defensins Myeloperoxidase (MPO) Proteinase-3 Elastase Cathepsins Bactericidal/ permeability increasing protein (BPI)	Respiratory burst components (e.g. cytochrome b_{558}) producing reactive oxygen metabolites, such as hydrogen peroxide, hydroxyl radicals and singlet oxygen Lysozyme Lactoferrin
Cell movement		Collagenase CD11b/CD18 (adhesion molecule) *N*-formyl-methionyl-leucylphenylalanine receptor (FMLP-R)

several weeks. The main role of eosinophils is protection against multicellular parasites such as worms (helminths). This is achieved by the release of toxic, cationic proteins. In populations and societies in which such parasites are rare, eosinophils contribute mainly to allergic disease, particularly asthma (see p. 850). Eosinophils have specific (95%) and primary (5%) granules.

The *specific granules* contain the cationic proteins, of which there are four main types. *Major basic protein (MBP), eosinophil cationic protein (ECP)* and *eosinophil neurotoxin* are all potently and exquisitely toxic to helminths, while ECP also has some bactericidal properties. *Eosinophil peroxidase* has similar activity to neutrophil myeloperoxidase.

Primary granules synthesize and release *leukotrienes C4 and D4* and *platelet activating factor (PAF)* which alter airway smooth muscle and vasculature (see p. 850).

Eosinophils are activated and recruited by a variety of mediators via specific surface receptors, including complement factors and leukotriene (LT) B_4. In addition, the CC chemokines eotaxin-1 (CCL11) and eotaxin-2 (CCL24) are highly selective in eosinophil recruitment. Receptors are also present for the cytokines IL-3 and IL-5, which promote the development and differentiation of eosinophils.

Mast cells and basophils

Mast cells and basophils share features in common, especially the presence of histamine-containing granules and the possession of high-affinity receptors for immunoglobulin E (IgE; an antibody type that is involved in allergic disease, see p. 72). Mast cells are tissue fixed and found at regular intervals in the skin and mucosae. Mast cell and basophil mediators are categorized as pre-formed and those synthesized de novo (Table 3.5).

Histamine is a low-molecular-weight amine (111 Da) with a blood half-life of less than 5 minutes; it constitutes 10% of the mast cell's weight. When injected into the skin, histamine induces the typical 'weal and flare' or 'triple' response. Initially there is reddening (erythema) of the skin at the site as arterioles dilate and post-capillary venules contract. This is followed by increased vascular permeability, with leakage of plasma fluid into the tissues causing swelling (weal). Finally, histamine acts directly on local axons to induce more widespread vascular changes distant from the injection site (flare).

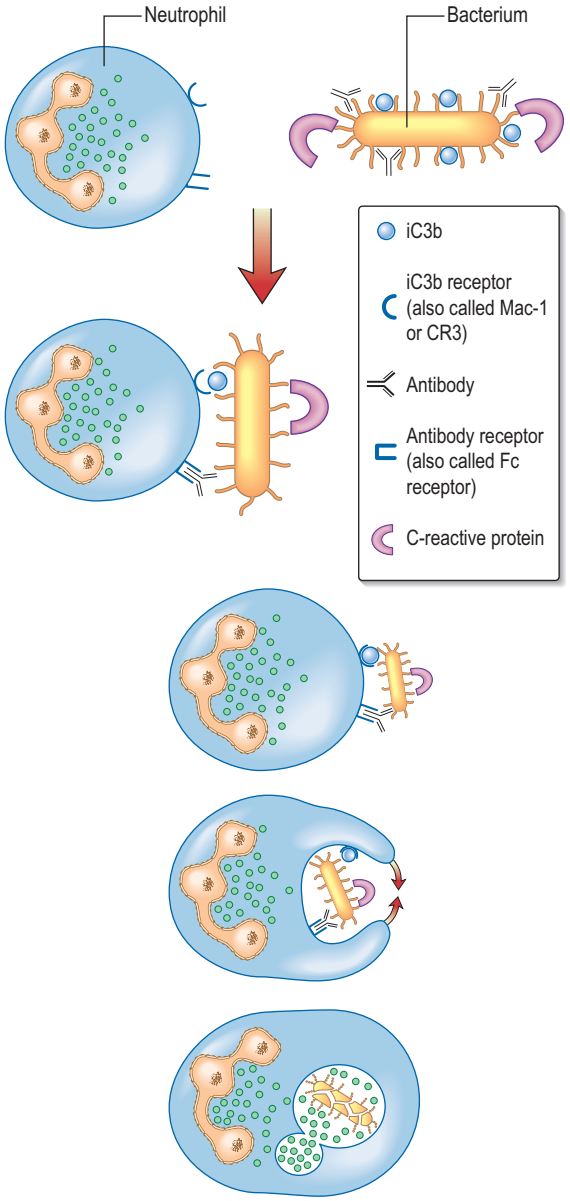


Fig. 3.3 **Opsonization.** Bacteria are coated with a variety of soluble factors from the innate immune system (opsonins) which enhance phagocytosis. This leads to engulfment of the bacteria into phagosomes and fusion with granules to release antibacterial agents. iC3b, inactive complement 3b.

The *complement-derived anaphylatoxins* C3a, C4a and C5a activate basophils and mast cells in response to a range of stimuli, whilst activation through binding of antigen to surface-bound antibody molecules of the IgE class is a feature of allergic disease. New evidence emphasizes that the mast cell also has a role in the early response to bacteria through release of TNF-α, which recruits and activates neutrophils as well as promoting readiness in nearby lymph nodes. In animal models mast cells can also be involved in recruitment to inflammatory sites such as arthritic joints, in promotion of tumour growth by enhancing neovascularization and in allograft tolerance.

Table 3.5 Mast cell and basophil mediators

Mediators	Effects
Pre-formed:	
Histamine	Vasodilatation Vascular permeability ↑ Smooth muscle contraction in airways
Proteases	Digestion of basement membrane causes ↑vascular permeability and aids migration
Proteoglycans (e.g. heparan)	Anticoagulant activity
Synthesized de novo:	
Platelet activating factor (PAF)	Vasodilator
LTB_4, LTC_4, LTD_4	Neutrophil and eosinophil activators and chemoattractants Vascular permeability ↑ Bronchoconstrictors
Prostaglandins (mainly PGD_2)	Vascular permeability ↑ Bronchoconstrictors Vasodilators

Monocytes and macrophages

Cells of the monocyte/macrophage lineage (also called mononuclear phagocytes) are highly sophisticated phagocytes. *Monocytes* are the blood form of a cell that spends a few days in the circulation before entering into the tissues to differentiate into *macrophages*, and possibly some types of dendritic cells. This is especially likely to happen during episodes of inflammation.

Blood monocytes can be divided into two subsets: those expressing CD14 (a receptor for lipopolysaccharide, a bacterial cell wall component) and those expressing CD14 and CD16 (a receptor for IgG antibodies). In vitro, both subsets have the potential to differentiate into myeloid dendritic cells (after culture with IL-4 and granulocyte-monocyte colony simulating factor, GM-CSF) and into macrophages.

A key role of tissue macrophages is the maintenance of tissue homeostasis, through clearance of cellular debris, especially following infection or inflammation:

- *Alveolar macrophages* in the lung clear microorganisms and debris encountered in the air.
- *Gut macrophages* are specialized towards bactericidal activity.
- *Osteoclasts*, another form of tissue macrophage, are multinucleate giant cells that resorb bone.

In general, the mononuclear phagocytes are responsive to a range of proinflammatory stimuli, using a pattern recognition receptor (PRR) to recognize pattern associated molecular patterns (PAMP)s; once activated they engulf and kill microorganisms, especially bacteria and fungi; in doing so they release a range of proinflammatory cytokines and have the capacity to present fragments of the microorganisms to T lymphocytes (see below) in a process called antigen presentation.

Macrophages have proinflammatory and microbicidal capabilities similar to those of neutrophils. Under activation conditions, antigen presentation (see p. 62) is enhanced and a range of cytokines secreted, notably TNF-α, IL-1 and IFN-γ.

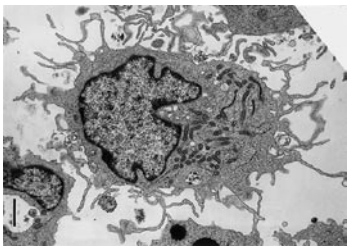


Fig. 3.4 **Electron micrograph of a mature dentritic cell.**

These are necessary for the removal of certain pathogens (e.g. mycobacteria) which live within mononuclear phagocytes. Thus while macrophages are capable of many of the bactericidal activities of neutrophils and have comparable phagocytic, chemotactic, opsonic and cytotoxic activities, they also play a role in the ingestion and killing of intracellular parasites, such as *Mycobacterium tuberculosis*. Tissue macrophages involved in such chronic inflammatory foci may undergo terminal differentiation into multinucleated giant cells, typically found at the site of the granulomata characteristic of tuberculosis and sarcoidosis (see p. 868).

Dendritic cells

The major function of dendritic cells (DCs) is that they alone are capable of activating naïve T lymphocytes to initiate adaptive immune responses. The definition of a dendritic cell is one that has:

- dendritic morphology (Fig. 3.4)
- machinery for sensing pathogens
- the ability to process and present antigens to CD4 and CD8 T lymphocytes, coupled with the ability to activate these T cells from a naïve state
- the ability to dictate the T cell's future function and differentiation.

This is a powerful cell type, that functions as a critical bridge between the innate and adaptive immune systems. In a sequence of events that spans 1–2 days, immature DCs are activated by PAMPs in the tissues and migrate to the local lymph node with the engulfed pathogen. During migration the DC matures, upregulating machinery required to activate T lymphocytes. Once in the lymph node, the DC interacts with naïve T lymphocytes (antigen presentation), resulting in two key outcomes. First is the activation of T cells with the ability to recognize peptide fragments (termed epitopes) of the pathogen; second is the polarization of the T cell towards a functional phenotype (see below) that is tailored to the particular pathogen.

The major types are the *myeloid DC (mDC)*, the *plasmacytoid DC (pDC)* and a variety of specialized DCs found in tissues that resemble mDCs (e.g. the Langerhans cell in the skin, see Fig. 23.1). DCs have several distinctive cell surface molecules, some of which have pathogen-sensing activity (e.g. DEC205 on mDCs) whilst others are inolved in interaction with T cells (Table 3.6). Immature mDCs and pDCs are present in the blood, but at very low levels (<0.5% of lymphocyte/monocyte cells).

Pathogen sensing is a key component of the function of immature DCs, as well as monocytes/macrophages, and is achieved through expression of a limited array of specialized PRR molecules capable of binding to structures common

Table 3.6 Molecules on dendritic cells

	Myeloid DC		Plasmacytoid DC	
	Immature	Mature	Immature	Mature
CD1c	+	++	–	–
CD123 (IL-3 receptor)	–	–	+	++
Molecules involved in co-stimulation of T cells (CD80, CD86)	+	+++	+	+++
HLA class I and II molecules for antigen presentation to T cells	+	+++	+	+++

Table 3.7 Toll-like receptors (TLR)

Pattern recognition receptor (PRR)	Pathogen associated molecular pattern (PAMP)	Pathogen	PRR expressed by
TLR2	Peptidoglycan	Gram positive bacteria	mDC
TLR3	Double stranded RNA	Viruses	mDC
TLR4	Lipopolysaccharide	Gram negative bacteria	mDC
TLR7	Single stranded RNA	Viruses	pDC
TLR9	Double stranded DNA	Viruses	pDC

mDC, myeloid dendritic cell; pDC, plasmacytoid dendritic cell

to pathogens, aided by long cell dendrites and pinocytosis (constant ingestion of soluble material).

PRRs include:

- *Mannan-binding lectin* which initiates complement activity inducing *opsonization* (p. 56).
- *Signalling receptors* that initiate nuclear factor kappa B (NF-κB) induction. For example, the PRR known as TLR4 (for toll-like receptor 4) binds lipopolysaccharide, a molecular pattern found in the cell walls of many Gram-negative bacteria (Table 3.7), whilst others bind double stranded and single stranded RNA from viruses. Innate immunity critically depends on toll-like receptor signalling. These receptors act through a critical adaptor molecule, myeloid differentiation factor 88 (MyD88) to regulate the activity of NF-κB pathways.
- *Endocyte pattern recognition receptors* which act by enhancing antigen presentation on macrophages, by recognizing microorganisms with mannose-rich carbohydrates on their surface or by binding to bacterial cell walls and scavenging bacteria from the circulation. All lead to *phagocytosis*.
- *TREM-1* (triggering receptor expressed on myeloid cells) is a cell surface receptor which, when bound to its ligand, triggers secretion of proinflammatory cytokines. It is upregulated by bacterial lipopolysaccharides but not in non-infective disorders.

Table 3.8 Myeloid dentritic cell (DC) maturation

Immature mDC	Mature mDC
Highly pinocytotic	Ceases pinocytosis
Low level expression of molecules required for T cell activation	Upregulates CD80, CD86 and HLA molecules
Low level expression of machinery required to process and present microbial antigens	Begins to process microbial antigens (break down into small peptides) in readiness to present them to T cells (using HLA molecules)
Generally localized and sedentary	Begins active migration to local lymph node
Minimal secretion of cytokines	Active secretion of cytokines in readiness to stimulate T cells; in particular IL-12

mDC, myeloid dendritic cell; pDC, plasmacytoid dendritic cell

The key principle at play here is that the immune system has devised a means of identifying most types of invading microorganisms by using a limited number of PRRs recognizing common molecular patterns, or PAMPs. This recognition event has been termed a 'danger signal': it alerts the immune system to the presence of a pathogen. Sensing danger is a key role of the DC and a key first step towards activation of the adaptive immune system.

DCs and T cell activation

The binding of a PAMP to a PRR activates the DC, resulting in one of the most amazing phenomena in cell biology as an immature DC changes its shape, gene and molecular profile and function within a matter of hours to take on a mature form, with altered functions (Table 3.8 and Fig. 3.5). Immature pDCs are small rounded cells that develop dendrites upon activation and secrete enormous quantities of IFN-α, a potent antiviral and proinflammatory cytokine.

The net result of these changes is that the mature DC is ready to activate naïve T cells in the lymph node. The mature DC provides three major signals to naïve T cells:

- Signal 1 = presentation of peptide fragments from the pathogen bound to surface HLA molecules.
- Signal 2 = co-stimulation through CD80 and CD86 interacting with CD28 on T cells.
- Signal 3 = secretion of cytokines, notably IL-12.

HLA MOLECULES AND ANTIGEN PRESENTATION

On the short arm of chromosome 6 is a collection of genes representing 0.1% of the whole genome and termed the major histocompatibility complex (MHC), which plays a critical role in immune function. In particular, the MHC contains a group of genes that code for proteins expressed on the surface of a variety of cell types that are involved in antigen recognition by T lymphocytes. In humans, these are known as the human leukocyte antigens, or HLA system. The T lymphocyte receptor for antigen recognizes its ligand pre-

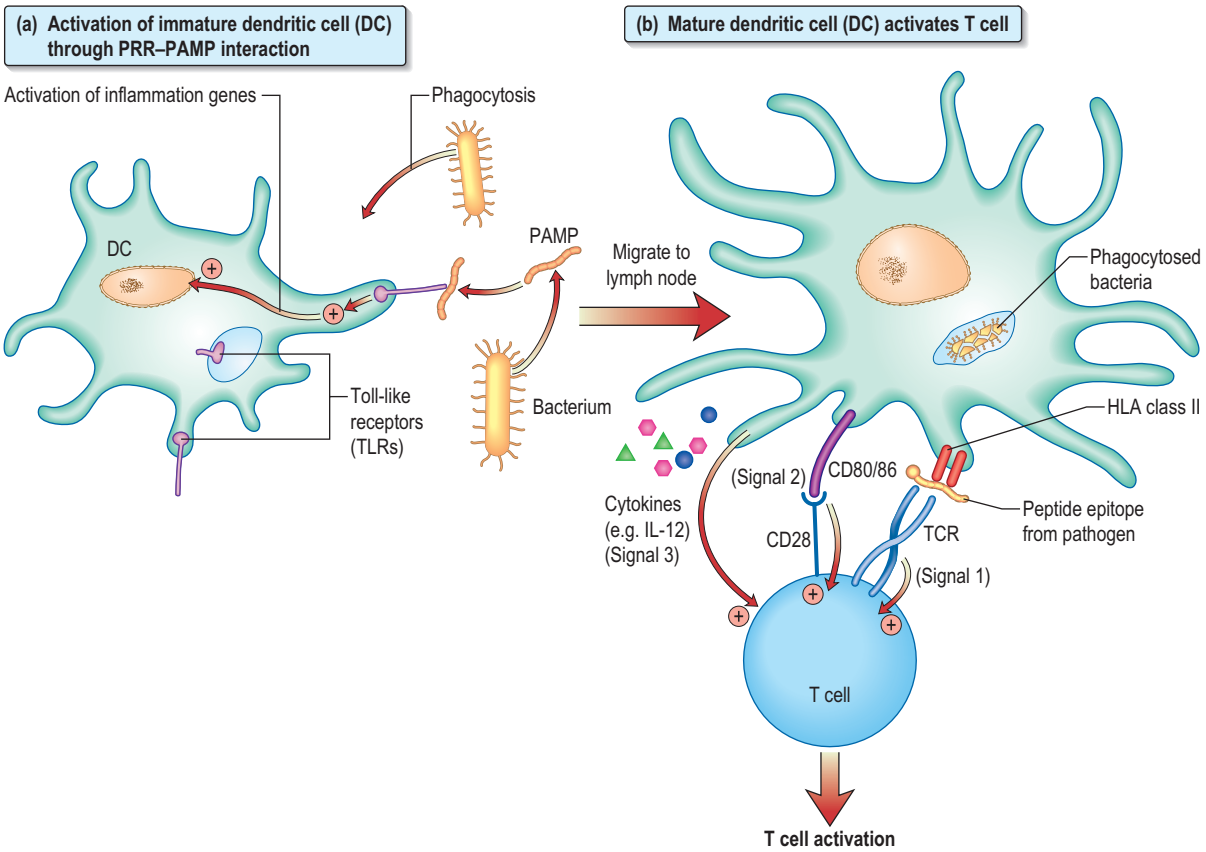


Fig. 3.5 **(a) Immature dendritic cells (DCs) in the tissues are activated by pathogens through PAMP–PRR interaction. (b) Multiple rapid changes in gene expression lead to migration to the lymph node as the DC takes on the mature phenotype.** During migration there is upregulation of the machinery required for activation of T cells (Table 3.8) shown here in response to signals 1–3.

Table 3.9 HLA associations with immune-mediated and infectious diseases

Disease process	Disease	HLA type
Autoimmunity	Type 1 diabetes	Class II: *DQA1*0301/DQB1*0302* (susceptibility) *DQA1*0501/DQB1*0201* (susceptibility) *DQA1*0102/DQB1*0602* (protection) Class I: *HLA-A*02; HLA-A*24; HLA-B*18; HLA-B*39*
	Multiple sclerosis	*DRB1*1501* (susceptibility)
	Rheumatoid arthritis	*DRB1*0404* (susceptibility)
	Autoimmune hepatitis	*DRB1*03* and *DRB1*04* (susceptibility)
	Goodpasture's syndrome (anti-glomerular basemement membrane disease)	*HLA-DRB1*1501*(susceptibility)
	Pemphigus vulgaris	*DRB1*0402; DQB1* 0503* (susceptibility)
Inflammatory	Coeliac disease	*DQA1*0501/DQB1*0201* (susceptibility)
	Ankylosing spondylitis	*HLA-B*27* (susceptibility)
	Psoriasis	*HLA-Cw*0602* (susceptibility)[1]
Infectious	Human immunodeficiency virus infection	*HLA-B*27; HLA-B*51; HLA-B*57* (associated with slow progression of disease) *HLA-B*35* (associated with rapid progression)

sented as a short antigenic peptide embedded within a physical groove at the extremity of the HLA molecule.

The HLA genes are particularly interesting for clinicians and biologists. First, differences in HLA molecules between individuals are responsible for tissue and organ graft rejection (hence the name histo- (*tissue*) compatibility). Second, the differences are many, as this is the most polymorphic region of the human genome. Third, possession of certain HLA genes is linked to susceptibility to particular diseases (Table 3.9).

The human major histocompatibility complex

The human MHC comprises three major classes (I, II and III) of genes involved in the immune response (Fig. 3.6 and see

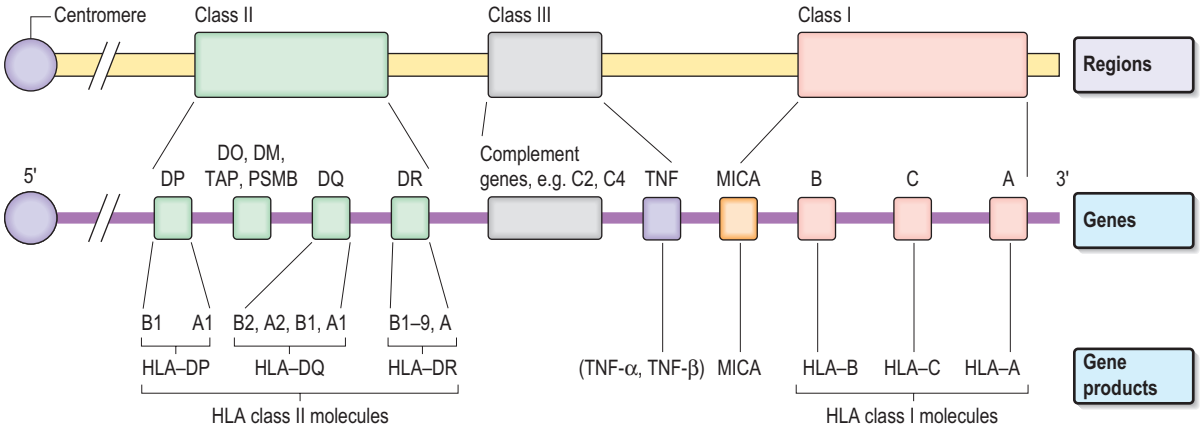


Fig. 3.6 **The HLA system in man.** On chromosome 6 are three major HLA regions (classes I–III) including genes that encode the HLA class I and II molecules, complement genes, the cytokine TNF and other genes involved in antigen presentation (HLA-DO, -DM; transporter associated with processing, TAP; proteasome subunit beta, PSMB; and the non-classical HLA class Ic molecule, MICA).

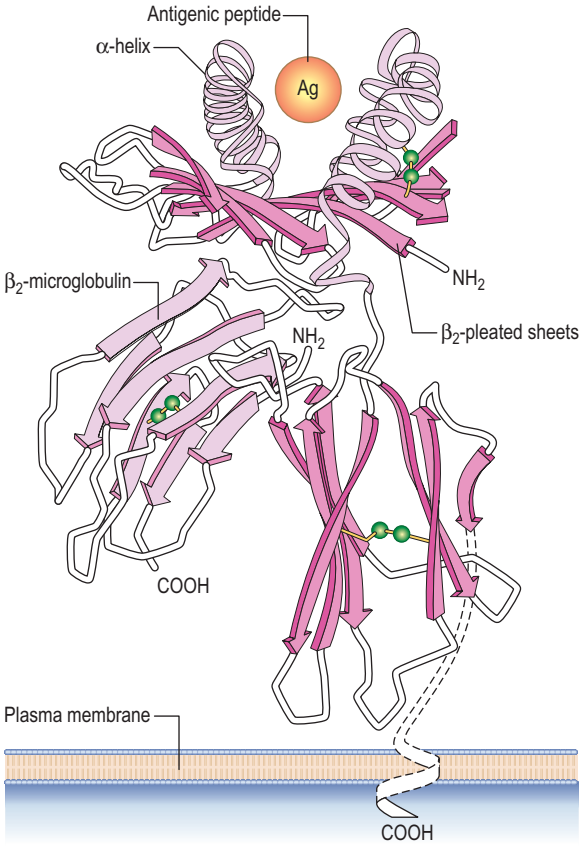


Fig. 3.7 **Structure of HLA class I molecule showing a short peptide antigen being presented in the peptide binding grove.**

http://www.sanger.ac.uk/HGP/Chr6/XMHC). In addition, several genes lie in the HLA region but do not fall into any of the three major categories.

- *Classical class I HLA genes* (also termed Ia) are designated *HLA-A, HLA-B* and *HLA-C*. Each encodes a class I α chain, which combines with a β chain to form the class I HLA molecule (the same β chain, called $\beta_2$microglobulin, is used for all α chains (Fig. 3.7)). The HLA class I molecule has the role of presenting short (8–10 amino acids) antigenic peptides to the T cell receptor on the subset of T lymphocytes that bear the co-receptor CD8. As an example of HLA polymorphism, there are nearly 200 allelelic forms at the *A* gene locus. Class I HLA molecules are expressed on all nucelated cells.
- *Non-classical HLA class I genes* are less polymorphic, have a more restricted expression on specialized cell types, and present a restricted type of peptide or none at all. These are the *HLA-E*, *F* and *G* (Ib genes) and MHC class I-related (MIC, or class Ic) genes, *A* and *B*. The products of these genes are predominantly found on epithelial cells, signal cellular stress and interact with lymphoid cells, especially natural killer cells (see p. 66).
- *The class II genes* have three major subregions, *DP, DQ* and *DR*. In these subregions are genes encoding *A* and *B* genes that combine to form dimeric αβ molecules that present short (12–15 amino acid) peptides to T lymphocytes that bear the CD4 co-receptor. Class II HLA genes (apart from *DRA*) are highly polymorphic. Other genes in this region encode proteins with key roles in antigen presentation (e.g. TAP, HLA-DM, HLA-DO, proteasome subunits; see below). Class II HLA genes are expressed on a restricted cohort of cells, that go by the general term of antigen presenting cells (APCs; DCs, monocyte/macrophages, B lymphocytes).
- *HLA class III genes* encode proteins that can regulate/modify immune responses, e.g. tumour necrosis factor (TNF), heat shock protein (HSPs) and complement protein (C2, C4).

HLA genotype is a risk factor for certain inflammatory diseases (Table 3.9) and also determines donor–recipient compatibility in transplantation and is denoted first by the letters that designate the locus (e.g. *HLA-A, DR, DQ*). For class I alleles, this is followed by an asterisk and then a two-four digit number defining the allelic variant at that locus, often called the HLA type (e.g. *HLA-A*02* is the 02 variant of the *HLA-A* gene). The class II nomenclature is the same, except that both *A* and *B* genes are named (HLA-DR molecules only require the name of the *B* gene, as the *A* gene is the same in all of us).

Some general principles apply to the HLA genes and their protein products.

First, the presence of multiple genes on each chromosome, and the fact that both maternal and paternal genes are co-dominantly expressed, allows considerable breadth in the number of HLA molecules that an individual expresses.

Ask the authors
www.studentconsult.com

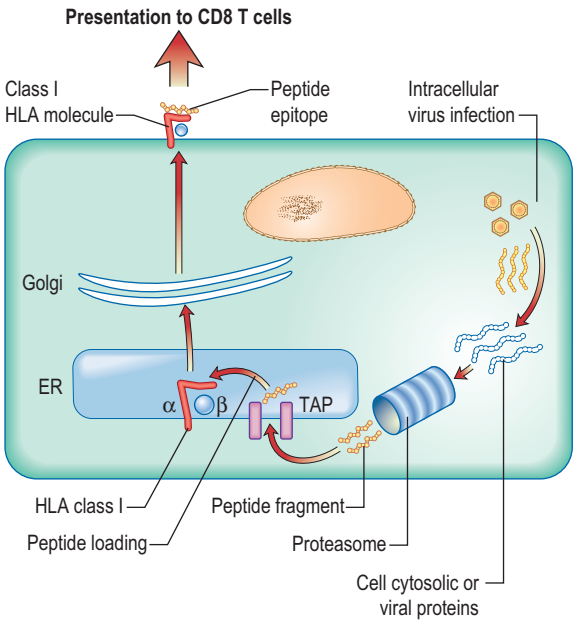


Fig. 3.8 **The endogenous route of antigen presentation to CD8 T cells.** Cytosolic proteins derived from self (resting cells) or viruses (virus-infected cells) are broken down by the proteasome (a sort of 'protein recycler') into fragments 2–25 amino acids long. Some of these are taken into the endoplasmic reticulum (ER) by a transporter (TAP), trimmed and loaded into empty HLA class I molecules. These are exported to the surface membrane for presentation to CD8 T cells.

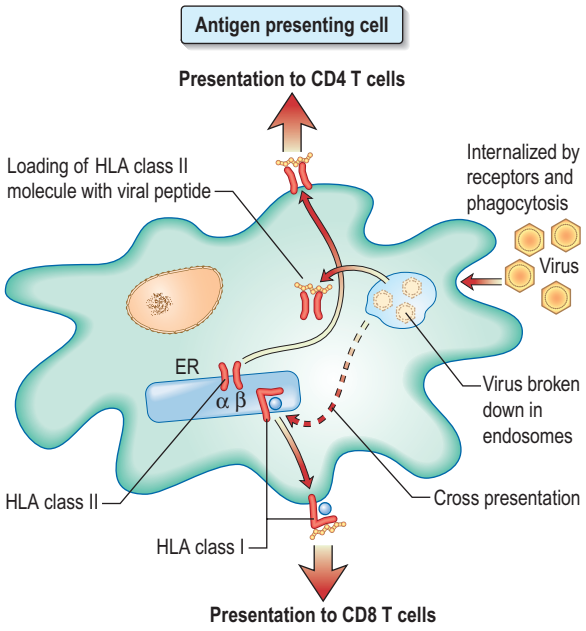


Fig. 3.9 **The exogenous route of antigen presentation and cross presentation.** External material (e.g. virus particles) is taken into an antigen presenting cell and broken down into specialized compartments by a combination of low pH and proteolysis. Peptides are then loaded into HLA class II molecules for presentation to CD4 T cells. Material may also be transferred across the cell so that it is loaded onto HLA class I molecules for presentation to CD8 T cells. This process of cross-presentation is restricted to specialized APCs such as DCs.

Second, the presence of polymorphisms at each locus provides great breadth in the number of HLA molecules expressed at a population level. The polymorphic forms of HLA molecules differ predominantly in the peptide-binding groove (Fig. 3.7).

Overall, then, each human can bind a range of peptide epitopes from pathogens to enhance individual protection, and the population an even greater range of protection, to ensure population survival.

Antigen presentation

HLA molecules bind short peptide fragments which are processed ('chopped up') from larger proteins (antigens) derived from pathogens. The peptide–HLA complex is presented on antigen presenting cells (APCs) for recognition by T cell receptors (TCRs) on T lymphocytes. There are three major routes to antigen processing and presentation:

- *The endogenous route* is a property of all nucleated cells and samples the internal mileu to generate peptide–HLA class I complexes for display on the cell surface. In a healthy cell, the peptides will be derived from self-proteins from the cytoplasm (Fig. 3.8), and will be ignored by the immune system. In a virus-infected cell, viral proteins are processed and presented. The resulting viral peptide–HLA class I complex is presented to CD8 T lymphocytes with cytotoxic (killer) function. In an immune response against a virus infection, CD8 T lymphocytes recognizing viral peptide–HLA complexes on the surface of an infected cell will kill, as a means to limit and eradicate infection.
- *The exogenous route* is a property of APCs and samples the external milieu (Fig. 3.9). Antigens are internalized, either in the process of phagocytosis of a pathogen, through pinocytosis, or through specialized surface receptors (e.g. for antigen/antibody/complement complexes). Within specialized intracellular compartments, the antigen is broken down by a combination of low pH and proteolytic enzymes for loading into HLA class II molecules. The pathogen peptide–HLA class II complex at the APC surface is able to interact with CD4 T lymphocytes. Presentation by DCs can initiate an adaptive immune response by activating a naïve, pathogen-specific CD4 T lymphocyte. Presentation by monocyte/macrophages and B lymphocytes can maintain and enhance this response by activating effector and memory pathogen-specific CD4 T lymphocytes.
- *Cross-presentation* refers to the ability of some APCs (mainly DCs) to internalize exogenous antigens and process them through the endogenous route (Fig. 3.8). This is an essential component in the activation of CD8 cytotoxic T lymphocyte responses against a virus.

CELLS AND MOLECULES OF THE ADAPTIVE IMMUNE SYSTEM

The information gained by DCs that interact with a pathogen is passed on, in the form of signals 1–3 (see Fig. 3.5b), to generate the adaptive arm of the immune response, a process through which specific recognition and removal of pathogens is organized by T lymphocytes. T lymphocytes may be involved in pathogen removal directly (e.g. by killing) or indirectly (e.g. by recruiting B lymphocytes to make specific antibody).

Antigen receptors on T and B lymphocytes

One of the key features of the adaptive immune system is specificity for antigen. For example, if you are immunized against measles virus, you do not have immunity to hepatitis B, and vice versa. Specificity is conferred by two types of receptor – the TCR on T lymphocytes and an equivalent on B lymphocytes, the BCR. BCRs are also termed surface *immunoglobulin* (sIg) and differ from TCRs in also being secreted in large quantities by end-stage B lymphocytes (plasma cells) as soluble immunoglobulins, commonly known as antibodies. To maintain protection against the multiplicity of pathogens in our environment, each of us generates great diversity amongst TCRs and antibodies. This is achieved through a mechanism shared by both TCR and antibody, in which distinct families of genes are encoded in the germline, each family (called constant, variable, diversity and joining) contributing a sequence to part of the receptor. Recombination between randomly selected members of each family ensures diversity in the end product. Recombination frequently involves base deletions and additions, adding to the diversity. In the case of antibodies, the result is a potential capacity of $>10^{14}$ different antibody molecules; for TCRs it may be as high as 10^{18}.

Immunoglobulins

These are produced from plasma cells which are end stage B cells. In structural terms, antibodies have four chains; two identical heavy and two identical light chains (Fig. 3.10). Each chain contains both highly variable *and* essentially constant regions. The variable parts of the heavy and light chains pair to form the potentially diverse part of the antibody molecule that binds antigen. The constant region of the heavy chain dictates the function of the antibody, and belongs to one of the classes M, G1-4, A1-2, D and E, giving rise to antibodies called IgM, IgG1-4, IgA1-2, IgD and IgE. The characteristics of these different isotypes are shown in Table 3.10.

Essential points about antibody production are:

- IgM is the first isotype to be made in a primary immune response and thus measurement of pathogen specific IgM is a useful diagnostic test for recent infection.
- IgG dominates in the second exposure to antigen.
- IgG and IgM are the most efficient complement activators when bound to antigen in an immune complex.
- IgG antibodies cross the placenta, and can carry both immunity and disease to the unborn fetus.

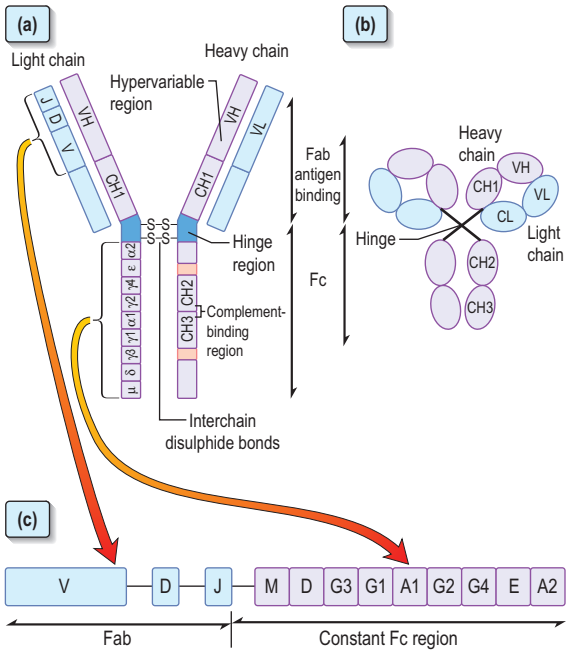


Fig. 3.10 **Immunoglobulin structure.**

(a) Basic subunit consisting of two heavy and two light chains.

(b) Schematic diagram of the same molecule. C and V, constant and variable domains; H and L, heavy and light chains; Fab, fragment antigen binding; Fc fragment crystalline.

(c) Genes on the Fab and Fc regions of an immunoglobulin. The chain is made up of a V (variable) gene which is translocated to the J (joining) chain. The VJ segment is then spliced to the C (constant) gene. Heavy chains have an additional D (diversity) segment which forms the VDJ segment that bears the antigen-binding site determinants.

Table 3.10 Characteristics of the immunoglobulins

	IgG Dominant class of antibody	IgM Produced first in immune response	IgA Found in mucous membrane secretions	IgE Responsible for symptoms of allergy: used in defence against nematode parasites	IgD Found almost solely on B lymphocyte surface membranes
Heavy chain	γ	μ	α	ε	δ
Mean adult serum levels (g/L)	IgG (total) = 8–16 G_1 = 6.5 G_2 = 2.5 G_3 = 0.7 G_4 = 0.3	0.5–2	IgA (total) = 1.4–4 A1 = 1.5 A2 = 0.2	17–450 ng/mL	0–0.4
Half-life (days)	21	10	6	2	3
Complement fixation					
Classical	++	+++	–	–	–
Alternative	–	–	+	–	–
Binding to mast cells	–	–	–	+	–
Crosses placenta	+	–	–	–	–

- IgA antibodies are present in secretions (tears, saliva, GI tract) to give protection to the mucosae.
- As it matures, and under the instruction of T lymphocytes, a B lymphocyte may change the class (class switching), but never the specificity, of the antibody it makes.
- As B lymphocytes mature and are stimulated to undergo further division, minor changes in antibody gene sequence can arise (somatic mutation) potentially allowing antibodies with higher affinity to arise and be selected for the effector response (affinity maturation).

Essential points about *antibody function* are:

- In host defence, antibodies target, neutralize and remove from the circulation and tissues infectious organisms and toxins, often through recruitment of innate host effector mechanisms such as complement, phagocytes and mast cells (by binding to specific surface receptors).
- In clinical medicine, specific antipathogen antibody levels are used in diagnosing/monitoring infectious disease, and may also be administered as serum pools to passively provide host protection.
- Antibodies can be raised in animals to generate monoclonal antibodies, which are commonly used in diagnostic immunology tests and increasingly used as therapeutics (e.g. to target cancer cells), often after 'humanization' (see below).

The genomic organization of TCR genes and principles of generation of receptor diversity are similar to those of immunoglobulin genes. The TCR exists as a heterodimer, with a similar overall structure as the antibody molecule. There are two TCR types:

- α and β chain (αβ TCR; expressed on all CD4 T lymphocytes and ~90% of CD8 T lymphocytes play a role in adaptive immune responses)
- γ and δ chains (γδ TCR fewer in number, mainly on intraepithelial lymphocytes; involved in epithelial defence).

The chains of each type of TCR are divided into variable and constant domains, each domain being encoded by separate gene pools. Like the B lymphocyte producing a single clone of immunoglobulin molecules, the T cell expresses only one form of TCR once the genes have been rearranged. Unlike antibodies, TCRs do not undergo somatic hypermutation and are not secreted.

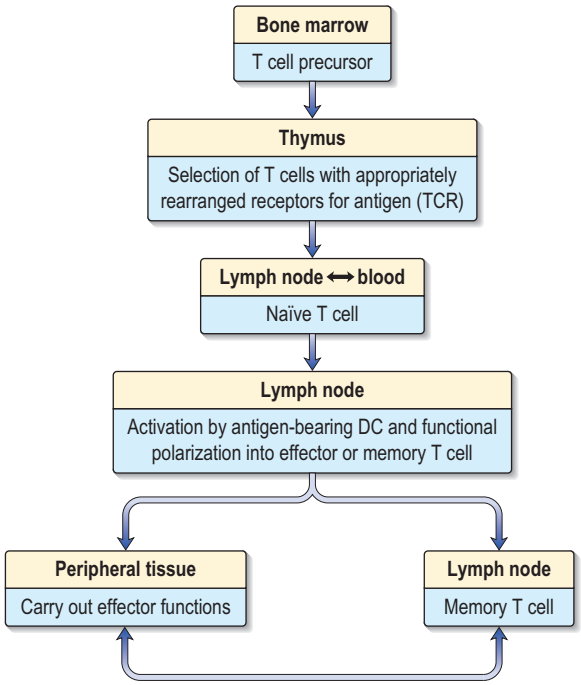


Fig. 3.11 **Development of T lymphocytes in the thymus and life cycle in the secondary lymphoid tissue.** DC, dendritic cell; TCR, T cell receptor.

Table 3.11 Identification of T cells and their major subsets by surface molecules

T cell population	Marker	Typical percentages in blood	Additional information
T lymphocytes	T cell receptor CD3	100% of T cells (70% of lymphocytes)	All T cells are thymus-derived
Helper T lymphocytes (Th)	CD4	66% of T cells	Interact with antigen presented by MHC class II molecules
Cytotoxic T lymphocytes (CTL)	CD8	33% of T cells	Interact with antigen presented by MHC class I molecules

T lymphocyte development and activation

T cells are generated from precursors in the bone marrow, which migrate to the thymus (Fig. 3.11). Only 1% of the cells that enter will leave the thymus as naïve T lymphocytes to populate the lymph nodes. This process (termed thymic selection) leads to a cohort of cells (Table 3.11) with:

- functionally rearranged genes allowing surface expression of a receptor for antigen (the T cell receptor, TCR) alongside the CD3 accessory molecule involved in transducing the antigen-specific signal
- selection of a co-receptor, either CD4 or CD8, to stabilize the interaction between TCR and peptide–HLA:
  - CD4 T cell responses require presentation of peptide antigens by *self* HLA class II molecules
  - CD8 T cell responses require presentation of peptide antigens by *self* HLA class I molecules
- a reduced or absent tendency of the selected TCR to recognize self antigens (and thus avoid autoimmunity).

Thus, during thymic education most TCRs are rejected for further use (negative selection), either because they are unable to bind self HLA molecules, or because they bind with too strong an affinity, which would run the risk of self-reactivity and autoimmune disease. The chosen TCRs (positive selection) have low/intermediate affinity for self HLA molecules. During post-thymic activation in the lymph node, the interaction with HLA will need to be bolstered by addi-

tional signals (co-stimulation) provided by DCs. This ensures that T lymphocytes are only activated when the checkpoint of DC maturation has been passed, which will only happen in the presence of pathogens.

Most naïve T lymphocytes are resident in the lymph nodes or spleen, whilst 2% are present in the blood, representing a recirculating pool. Naïve T lymphocytes are activated for the first time in the lymph node by antigens presented to their TCRs as short peptides bound to MHC molecules on the surface of DCs. Of all the cells equipped and capable of antigen presentation (monocyte/macrophages, DCs and B lymphocytes), only DCs are capable of activating naïve T cells. Provision of signals 1–3 (see p. 59) sets off an intracellular cascade of signalling molecule activation leading to induction of gene transcription. Key transcription factors are NF-AT (nuclear factor of activated T cells), NF-κB and AP-1 (activator protein-1), initiating transcription of, amongst others, the IL-2 gene which is critical for survival and expansion.

Nuclear factor kappa B (NF-κB) is a pivotal transcription factor in chronic inflammatory diseases and malignancy. It is a heterodimer of two proteins (p. 59) and is found in the cytoplasm bound to an inhibitor IκB which prevents it from entering the nucleus (see Fig. 2.4). It is released from IκB on stimulation of the cell and passes into the nucleus. Here it binds to specific sequences in the promoter regions of target genes. It is stimulated by, for example, cytokines, protein C activators and viruses and itself regulates various proteins (e.g. proinflammatory cytokines, chemokines, adhesion molecules, inflammatory enzymes and receptors).

The outcome is T cell activation, cell division and functional polarization, which is the acquired ability to promote a selected type of adaptive immune response. These processes take several days to achieve. The best described polarities of T cell responses (Table 3.12) are:

- $CD4^+$ proinflammatory T cells; T helper 1 (Th1), Th2 and Th17
- $CD8^+$ cytotoxic T lymphocytes (CTLs)
- $CD4^+$ regulatory responses (Treg).

The **Th17** subset of T cells is newly described and derives its name from the secretion of IL-17. The role of Th17 cells in the immune response is not yet entirely clear. They are capable of secreting numerous proinflammatory cytokines (e.g. IL-6, IL-8) and growth factors (GM-CSF). Studies in mice suggest that these cells are critically involved in the protection from pathogens and inflammation associated with the murine versions of diseases such as multiple sclerosis and rheumatoid arthritis; it is unknown whether this is also true in man but it does appear to have some vole in Crohn's disease.

Through cell division, a proportion of the T lymphocytes that are activated in response to a pathogen undertake these effector or regulatory functions, whilst a proportion is assigned to a memory pool. Once established, effector and memory T cells have lesser requirements for subsequent activation which can be mediated by mononuclear phagocytes and B lymphocytes.

CD4 T lymphocyte functions

As the pivotal cell in immune responses, the CD4 T cell influences most aspects of immunity, either through the release of cytokines, or via direct cell–cell interaction, a process often termed 'licensing'. Licensing critically involves the pairing of CD40 on the APC and CD40 ligand (CD154) on the CD4 T cell; defects in this process result in antibody deficiency (see below).

Major functions of CD4 T lymphocytes are:

- licensing of DCs during antigen presentation to activate CD8 T lymphocytes and generate cytotoxic cells
- licensing of B lymphocytes to initiate and mature antibody responses, leading to class switching, affinity maturation of antibodies and generation of plasma cells or memory cells
- secretion of cytokines responsible for growth and differentiation of a range of cell types, especially other T cells, macrophages and eosinophils
- regulation of immune reactions.

Regulatory T cells

The generation of B and T lymphocytes provides a potentially vast array of rearranged antigen receptors. Although there are selection processes to remove lymphocytes with 'dangerous' avidity for 'self' these are not foolproof and the potential for autoreactivity remains. The fact that there is no self-destruction in the vast majority of people implies a state of immunological self tolerance; the controlled inability to respond to self. Several mechanisms operate to maintain this state, including CD4 T lymphocytes that respond to antigenic stimulation by suppressing ongoing immune responses. These regulatory T cells (Tregs) have different origins, phenotypes and modes of action:

Table 3.12 Identification of CD4 T cell subsets by function

T cell type	Main cytokines causing polarization	Main cytokines produced	Functions	Major role in physiological immune response
T helper 1 (Th1) cells	IL-12	IFN-γ, IL-2, TNF-α	Proinflammatory	Organize killing of bacteria, fungi and viruses; activate macrophages to kill intracellular bacteria; instruct cytotoxic T lymphocyte responses
T helper 2 (Th2) cells	IL-4	IL-4, IL-5, IL-13	Proinflammatory	Organize killing of parasites by recruiting eosinophils; promote antibody responses, especially switching to IgE
T helper 17 (Th17) cells	IL-6, IL-23, TGF-β	IL-17	Proinflammatory	Not yet fully defined; capable of recruiting cells and damaging targets; may be more resistant to Treg than Th1/Th2 cells
Regulatory T cells (Treg)	IL-10, TGF-β	IL-10, TGF-β	Regulatory	Regulation of inflammation

- $CD4^+$ $CD25^{hi}$ Tregs; express high levels of CD25, the receptor for IL-2; regulate other T cells by cell–cell contact and also secrete the immune suppressive cytokines IL-10 and TGF-β; can be generated in the thymus or post-thymically in the periphery; have high expression of the transcription factor Foxp3.
- Tr1 Tregs; generated naturally or induced (e.g. by repeated antigen injection); regulate through production of IL-10.

Evidence that Tregs are clinically relevant is given by the example of immune deficiency states in which they are defective. For example, genetic defects in the Foxp3 gene give rise to **IPEX** (Immune dysregulation, Polyendocrinopathy, Enteropathy, X-linked syndrome). Patients with this rare syndrome have defective Tregs and develop a range of conditions soon after birth, including organ-specific autoimmune disease such as type 1 diabetes. Much research effort is directed at harnessing this natural regulatory potential, for example to control organ graft rejection and autoimmune disease.

CD8 T lymphocyte functions

Cytotoxic CD8 T lymphocytes (CTL) are involved in defence against viruses. CTLs kill virus-infected cells following recognition of viral peptide–HLA class I complexes. CTL must be activated first in the lymph node by a DC cross-presenting the *same* viral peptide and licensed by a CD4 T lymphocyte recognizing viral peptide–HLA class II complexes. The same defence mechanism may also apply in tumour surveillance. This is a checkpoint that ensures that CTL responses, which have great destructive power, are only activated against a target for which there is also a CD4 T cell response.

CTLs kill via three mechanisms:

- cytotoxic granule proteins (cytolysins such as perforin, granzyme B)
- toxic cytokines (e.g. IFN-γ, TNF-α)
- death-inducing surface molecules (e.g. Fas ligand binds Fas on target cells mediating apoptosis via caspase activation) (see p. 28).

Natural killer (NK) cells

These are bone marrow-derived, present in the blood and lymph nodes and represent 5–10% of lymphoid cells. The name is derived from two features. Unlike B and T lymphocytes, NK cells are able to mediate their effector function (i.e. *killing* of target cells) *spontaneously* in the absence of previous known sensitization to that target. Also, unlike B and T lymphocytes, NK cells achieve this with a very limited repertoire of germ line-encoded receptors that do not undergo somatic recombination. The lack of requirement for sensitization and the absence of gene rearrangement to derive receptors for target cells mean that NK cells are part of the innate immune system. For identification purposes, the main surface molecules associated with NK cells are CD16 (see below) and CD56 (note, NK cells are CD3- and TCR-negative).

The role of NK cells is to kill 'abnormal' host cells, typically cells that are virus-infected, or tumour cells. Killing is achieved in similar ways to CTLs. NK cells also secrete copious amounts of IFN-γ and TNF-α, through which they can mediate cytotoxic effects and activate other components of the innate and adaptive immune system. To become activated, NK cells integrate the signal from a potential target cell through a series of receptor-ligand pairings (Table 3.13).

Table 3.13 Examples of natural killler (NK) cell receptors

Receptors on NK cells	Ligand
Inhibitory:	
KIR2DL1	HLA-C molecules
KIR3DL1	HLA-B molecules
KIRDL2	HLA-A molecules
NKG2A	HLA-E molecules
Activating:	
CD16 (low affinity receptor for IgG)	IgG
NKG2D	MHC class I chain related gene A (MICA)
KIR2DS1	HLA-C

KIR, killer cell immunoglobulin-like receptors.

These pairings provide activating and inhibitory signals, and it is the overall balance of these that determines the outcome for the NK cell. The balance might be abnormal on a virus-infected or tumour cell, which might have altered expression of HLA molecules, for example, that mark them out for NK cytotoxicity. In addition, CD16 is the low affinity receptor for IgG (FcRγIIIA) and through this NK cells can kill IgG-coated target cells in a process termed antibody-dependent cellular cytotoxicity (ADCC).

As for T cells, NK cell activation requires initial contact, adhesion and polarization. It is not known which of the inhibitory or activating receptors (Table 3.13), if any, is the first to initiate target cell contact. Once contact is made, adhesion occurs and is predominantly mediated by LFA-1: ICAM-1 interactions (see below). These events polarize the NK cell on to its target and direct the release of granules and cytokines. One of the targets of NK cell activating receptors, MICA, is induced under conditions of cellular 'stress', such as might arise during infection or neoplastic transformation.

In practice, in vivo, how does the process of NK killing occur? Many viruses have developed immunoevasion strategies that avoid presentation of viral proteins to CTLs by interfering with the MHC class I presentation pathway. This renders infected cells susceptible to killing by NK cells, which detect the reduced levels of MHC class I molecules. Likewise, tumours that escape immune surveillance by CTLs through the outgrowth of daughter cells that have low MHC expression, then become NK cell targets. The relatively recent development of NK cell receptor diversity implies a strong evolutionary driving force, probably related to viral immunoevasion strategies. KIR–HLA interactions (KIR = killer cell immunoglobulin-like receptors) are thus vital to antiviral immunity, and this is confirmed by the numerous associations between KIR and HLA genotype combinations and susceptibility to chronic infection with viruses, notably the human immunodeficiency virus (see p. 184).

CELL MIGRATION

Immune cells are mobile. They can migrate into the lymph node to participate in an evolving immune response (e.g. a pathogen-loaded dendritic cell (DC) from the skin, or a recir-

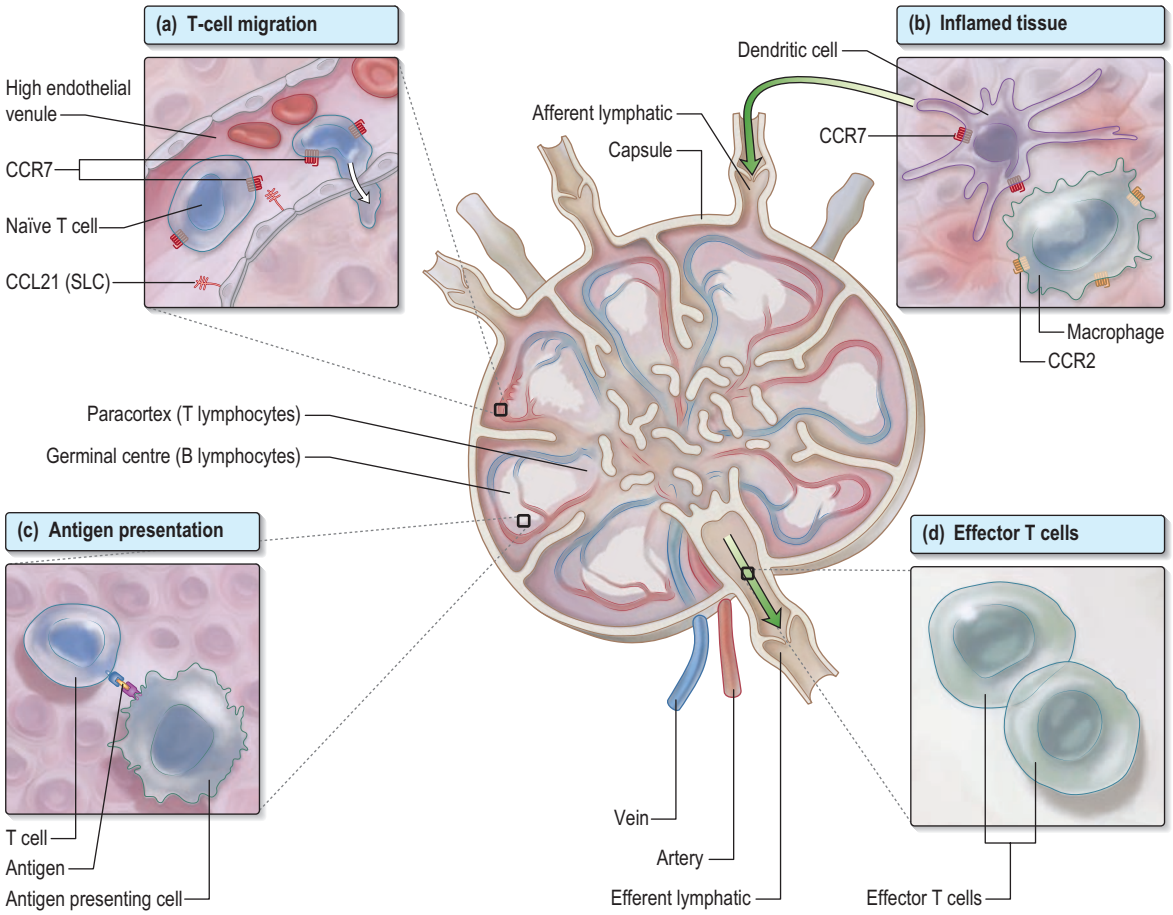


Fig. 3.12 **Movement of an immune cell through a lymph node.**
(a) T cells migrate though endothelial venules which express the chemokine CCL21 (secondary lymphoid tissue chemokine or SLC). **(b)** Dendritic cells (expressing CCR7), macrophages (expressing CCR2) enter lymph nodes via afferent lymphatics and localize in the paracortex. **(c)** Antigen presentation (see also Figs 3.8 and 3.9) results in activation of T cells. **(d)** Effector T cells leave lymph node via efferent lymphatic. Adapted from Charo IF, Ransohoft RM. *New England Journal of Medicine* 2006; **354**: 610–621, figure 1. Copyright © 2006 Massachusetts Medical Society. All rights reserved.

culating naïve T cell), or can migrate from the lymph node, via the blood, to the site of a specific infection in the tissues. Such migration takes place along blood and lymphatic vessels, and is a highly regulated process.

Taking the example of a DC migrating into the lymph node from the tissues via the lymphatics (Fig. 3.12), this is highly dependent upon expression of the chemokine receptor CCR7 by the migrating cell and of its ligand (secondary lymphoid tissue chemokine, SLC or CCL21) by the target tissue. Likewise, circulating naïve lymphocytes are $CCR7^{+}$ and migrate with ease into the lymph nodes from the blood or via tissue recirculation. In addition, naïve lymphocytes express L-selectin which binds a glycoprotein cell adhesion molecule (GlyCAM-1) found on the high endothelial venules of lymph nodes. This system can be upregulated in an inflamed lymph node, leading to an influx of naïve lymphocytes and the typical symptom of a swollen node.

Migration into inflamed tissue requires that:

(i) an affected organ or tissue signals that there is a focus of injury/infection and
(ii) that responding immune cells bind and adhere specifically to that tissue.

This process is highly organized and has a similar basis for all immune cells, involving three basic steps: rolling, adhesion and trans-migration. Each of these is dependent on specialized adhesion molecules, as shown in Figure 3.13.

- *Rolling* is mediated by selectins. P-selectin, E-selectin and L-selectin (for *P*latelet, *E*ndothelium and *L*eucocyte, respectively) are named after the predominant cell type that expresses them and bind to carbohydrate moieties such as sialyl-$Lewis^{X}$ (CD15s) on immune cells inducing them to slow and roll along the vessel lumen.
- *Adhesion* is mediated by integrins on immune cells binding to intercellular adhesion molecules (ICAMs) on the endothelium. The best known integrin is the heterodimer of CD11a/CD18 (called leucocyte function associated antigen-1; LFA-1) which binds ICAM-1.
- *Trans-migration* (diapedesis) is a complex process that includes structures known as platelet-endothelial cell adhesion molecule-1 (PECAM-1) and junctional adhesion molecules (JAMs).

The controlled regulation of these molecules (e.g. LFA-1 expression) is upregulated on T cells after activation in the lymph node. ICAM-1 expression on tissue endothelium is sensitive to numerous proinflammatory molecules and allows immune cells to be guided from the blood into the tissues. Once there, cells move along a gradient of increasing

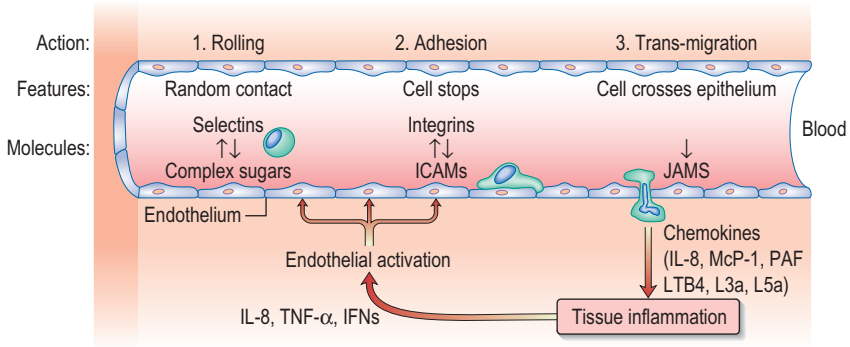


Fig. 3.13 **Neutrophil migration.** Neutrophils arrive at the site of the inflammation, attracted by various chemoattractants. They roll along the blood vessel wall through interaction between L-selectin, on their surface, binding to a carbohydrate structure (e.g. sialyl-Lewis^x) on the endothelium. E-selectin on the endothelium mediates a similar effect.

Firm adhesion is then mediated via interaction between integrins on the neutrophil and ICAM-I/VCAM-1 on the endothelium.

Following diapedesis, mediators such as interleukin-8, macrophage-chemotactive factor and TNF-α attract and activate the neutrophil to the infected tissue where it phagocytoses and destroys the C3b-coated bacteria.

LFA, leucocyte function antigen; PAF, platelet activating factor; LTB_4, leukotriene B4; ICAM, intercellular adhesion molecule; VCAM, vascular cell adhesion molecule.

concentration of mediators such as chemokines in the process of chemotaxis.

THE IMMUNE SYSTEM IN CONCERT

Acute inflammation: events and symptoms

This is the early and rapid host response to tissue injury. Taking a bacterial infection as the classic example, local expansion of pathogen numbers leads to direct activation of complement in the tissues, with ensuing degranulation of mast cells. Inflammatory mediators (from mast cells and complement) change the blood flow and attract and activate granulocytes. Concomitantly there are the local symptoms of heat, pain, swelling and redness, and perhaps more systemic symptoms such as fever due to the effect of circulating cytokines (IL-1, IL-6, TNF-α) on the hypothalamus. Systemically active mediators (especially IL-6) also initiate the production of C reactive protein (CRP) in the liver. Bacterial lysis follows through the actions of complement and neutrophils, leading to formation of fluid in the tissue space containing dead and dying bacteria and host granulocytes ('pus'). The inflammation may become organized and walled off through local fibrin deposition to protect the host. Antigens from the pathogen travel via the lymphatics (which may become visible as red tracks in the superficial tissues – lymphangitis) in soluble form or are carried by dendritic cells to establish an adaptive immune response which, at the first host–pathogen encounter, takes approximately 7–14 days. This will lead to activation of pathogen-specific T cells, B cells and production of pathogen-specific antibody, initially of the IgM class and of low-moderate affinity and subsequently of the IgG class (or IgA if the infection is mucosal) and high affinity. Resolution of the infection is aided by the scavenging activity of tissue macrophages.

Chronic inflammation: events and symptoms

Inflammation arising in response to immunological insults that cannot be resolved in days/weeks gives rise to chronic inflammation. Examples include infectious agents (chronic viral infections such as hepatitis B and C or intracellular bacteria such as mycobacteria) and environmental toxins such as asbestos and silicon.

Chronic inflammation is also a hallmark of some forms of allergic disease, autoimmune disease and organ graft rejection. The common feature of these pathological processes is that the inciting stimulus is not easy to remove. For example, some viruses and mycobacteria remain hidden intracellularly; antigens that incite allergy (allergens) may be constantly present in the host's environment; self or donated organs are a resident source of antigens. In many ways, the pathology that results is thus inadvertent; the immune system is caught between the repercussions of not dealing with the infection/insult, and the tissue damage that is caused by chronic activation of lymphoid and mononuclear cells.

Chronic inflammation may lead to permanent organ damage or impaired vascular function and can be fatal. If the inciting stimulus is removed, inflammation resolves. However, inflammation returns rapidly (24–48 hours) on re-exposure. This rapid recall response is the basis for patch testing to identify the cause of contact dermatitis, another form of chronic inflammation, and also for the Mantoux test of tuberculosis immunity. Together, these forms of immunopathology used to go under the umbrella term 'Type IV hypersensitivity'.

The main immunological event is the presence of a pro-inflammatory focus comprising T and B lymphocytes and APCs, especially macrophages. If antigen persists, inflammation becomes chronic and the macrophages in the lesion fuse to form giant cells and epithelioid cells. Both Th1 and Th2 reactivity is recognized, but specific syndromes may be polarized towards one or the other (e.g. chronic mycobacterial or viral infection engenders Th1 responses, chronic allergic inflammation Th2).

When the inflammation is sufficiently chronic it may take on the appearance of organized lymphoid tissue resembling a lymph node germinal centre (e.g. in the joints in rheumatoid arthritis). There is massive cytokine production by T cells and APCs, which contributes to local tissue damage. Granulomata, which 'wall off' the inciting stimulus, may also arise and result in fibrosis and calcification. Symptoms typically relate to the site of the inflammation and the type of pathology, but there may also be systemic effects such as fever and weight loss.

Table 3.14 Examining the immune system in the clinical immunology laboratory

	Measurement	Interpretation
Proteins	C-reactive protein	Raised levels indicate infection or inflammation
	Immunoglobulins	Low levels indicate antibody deficiency, usually a result of underlying disease or primary immunodeficiency. High levels, e.g. ↑ IgM seen in acute viral infection (hepatitis A)
	Complement C3 and C4	Low levels indicate consumption of complement in immune complex disease
	IgE	Raised levels in allergy; allergen-specific IgE useful to pinpoint the inciting stimulus (e.g. pollen, grass)
Cells	Neutrophils	High levels in bacterial infection; low levels in secondary immune deficiency
	Eosinophils	High levels in allergic or parasitic disease
	CD4 T cells	Low levels in HIV infection
Function	Neutrophil respiratory burst	Absent in the immune deficiency chronic granulomatous disease
	T cell proliferation	Abnormally low in primary T cell immune deficiency disease
Autoantibodies (see also Table 10.2 and Box 10.11)	Rheumatoid factor, anti-cyclic citrullinated peptide (CCP) antibodies	Rheumatoid arthritis
	Double-stranded DNA autoantibodies	Systemic lupus erythematosus
	Acetylcholine receptor antibodies	Myasthenia gravis
	Anti-neutrophil cytoplasmic antibodies (ANCA)	Vasculitis
	Mitochondrial	Primary biliary cirrhosis

LABORATORY INVESTIGATIONS OF THE IMMUNE SYSTEM

In the clinical immunology laboratory, proteins and cells can be measured to ascertain the status of the immune system. The results may indicate an undiagnosed inflammatory or infectious disease (e.g. through high C-reactive protein level); a state of immune deficiency (e.g. low concentration of IgG); or a state of immune pathology (e.g. the presence of autoantibodies or allergen-specific IgE).

Examples of the commoner tests and their interpretation are shown in Table 3.14.

CLINICAL IMMUNODEFICIENCY

Secondary (acquired) versus primary immunodeficiency (Table 3.15)

Most forms of immunodeficiency are secondary – to infection (mainly HIV) or therapy (e.g. corticosteroids, anti-TNF-α monoclonal antibody therapy, cytotoxic anticancer drugs, bone marrow ablation pre-transplant). *Acquired neutropenias* are common (e.g. due to myelosuppression by disease or drugs or the increased rate of destruction in hypersplenism or autoimmune neutropenia) and carry a high risk of infection once the neutrophil count falls below 0.5 $\times 10^9$/L.

Acquired reductions in levels of immunoglobulins (hypogammaglobulinaemia) are seen in patients with myeloma and chronic lymphocytic leukaemia or lymphoma. Splenectomy causes impairment of defence against capsulated bacteria, especially pneumococcus, and such patients should receive pneumococcal, meningococcal and Hib vaccinations as a matter of course (see p. 423).

Primary immunodeficiency is rare and arises at birth as the congenital effect of a developmental defect or as a result of genetic abnormalities. Gene defects may not become manifest until later in infancy or childhood, and some forms of immunodeficiency typically present in adolescence or adulthood.

Table 3.15 Classification of immunodeficiencies and the main diseases in each category

Immune component	Examples of diseases
T lymphocyte deficiency	DiGeorge syndrome Acquired immunodeficiency syndrome/HIV infection T cell activation defects (e.g. CD3γ chain mutation) X-linked hyper-IgM syndrome (XHIM; CD40L deficiency)
B lymphocyte deficiency	X-linked agammaglobulinaemia (XLA) Common variable immunodeficiency (CVID) Selective IgA deficiency (IgAD)
Combined T and B cell defects	Severe combined immunodeficiency (SCID) (e.g. due to defects in common γ chain receptor for IL-2, -4, -7, -9, -15)
T cell–APC interactions	IFN-γ receptor deficiency IL-12 and IL-12 receptor deficiency
Neutrophil defects	Chronic granulomatous disease (CGD) Leucocyte adhesion deficiency (LAD)
Deficiency of complement components	Classical pathway Alternative pathway Common pathway Regulatory proteins Mannan binding lectin

Clinical features of immunodeficiency

The infections associated with immunodeficiency have several typical features:

- often chronic, severe or recurrent
- resolve only partially with antibiotic therapy or return soon after cessation of therapy
- the organisms involved are often unusual ('opportunistic' or 'atypical').

The pattern of infection, in terms of the type of organism involved, is indicative (Table 3.16):

- *'Opportunistic' organisms* are of low virulence but become invasive in immunodeficient states, e.g. atypical mycobacteria, *Pneumocystis jiroveci*, *Staphylococcus epidermis*.
- *Phagocyte defects* cause, for example, deep skin infections, abscesses, osteomyelitis.
- *Defective antibody* producers experience infections with pyogenic ('pus-forming') bacteria.
- *T cell deficiency* causes infection with fungi, protozoa and intracellular microorganisms.
- *Congenital deficiencies* of antibody production are not revealed for several months after birth, due to the 28-day half-life of maternal IgG.

The *family history* may reveal unexplained sibling death; only females of the family affected; or consanguinity, each of which make a primary genetic syndrome more likely. Graft-versus-host disease (GVHD) may arise as a complication of primary or secondary T cell immunodeficiency. For GVHD to arise, there must be impaired T cell function in the recipient and the transfer of immunocompetent T lymphocytes from an HLA non-identical donor (see below). GVHD usually arises from therapeutic interventions such as transfusions or transplantation.

Primary immunodeficiency

T lymphocyte deficiency

DiGeorge syndrome is due to T cell deficiency. The third and fourth pharyngeal arches, which normally give rise to the parathyroid glands, aortic arch and the thymus, fail to develop. DiGeorge arises in 1–5 per 100 000 of the population and presents at birth with dysmorphic facies, hypoparathyroidism and cardiac defects, followed by infections (fungi, protozoa) in later months. The lack of a location for T cell development, i.e. the thymus, means that affected children have reduced/absent T lymphocyte number and proliferation responses. Apart from calcium supplementation, correction of cardiac abnormalities and prophylactic antibiotics, cure has been reported with thymic transplantation using fetal tissue or stem cell transplant (SCT) from HLA identical siblings.

Other T cell deficiencies caused by single gene defects have been characterized and typically present from the age of 3 months with candidal infections of the mouth and skin, protracted diarrhoea, fever and failure to thrive. Examples include deficiency of CD3 itself; defects in signal transduction pathways; and deficiency of IL-2, the T cell growth factor. These disorders are similar in presentation and management to the combined (T and B lymphocyte) immune deficiencies.

Table 3.16 Immune defects and associated infections

Neutropenia and defective neutrophil function
Staphylococcus aureus
Staphylococcus epidermidis
Escherichia coli
Klebsiella pneumoniae
Proteus mirabilis
Pseudomonas aeruginosa
Serratia marcescens
Bacteroides spp.
Aspergillus fumigatus
Candida spp. (systemic)

Opsonin defects (antibody/complement deficiency)
Pneumococcus
Haemophilus influenzae
Meningococcus
Streptococcus spp. (capsulated)

Antibody deficiency only
Campylobacter spp.
Mycoplasma spp.
Ureaplasma spp.
Echovirus

Complement lytic pathway defects (C5–9)
Meningococcus
Gonococcus (disseminated)

Defect in T cell or T cell/APC responses
Listeria monocytogenes
Legionella pneumophila
Salmonella spp. (non-typhi)
Nocardia asteroides
Mycobacterium tuberculosis
Atypical mycobacteria, especially *M. avium-intracellulare*
Candida spp. (mucocutaneous)
Cryptococcus neoformans
Histoplasma capsulatum
Pneumocystis jiroveci
Toxoplasma gondii
Herpes simplex
Herpes zoster
Cytomegalovirus
Epstein–Barr virus

T and B lymphocyte deficiency

Severe combined immunodeficiencies (SCID) are a heterogeneous group of rare (1–2 live births per 100 000), genetically determined disorders resulting from impaired T, NK and B cell immunity. The most common form of SCID is an X-linked defect in the IL-2 receptor γ chain, interfering with the function of not just IL-2 but also IL-4, IL-7, IL-9 and IL-15 which share this component in their receptors. Clinical features are similar to pure T cell deficiency and in the blood T and NK cells are lacking whilst B cell numbers may be normal. The most successful treatment for all forms of SCID is SCT, preferably from HLA-identical donors such as siblings.

Hyper IgM syndrome (HIGM) is a milder mixed deficiency, which may not be diagnosed until later in life. It results from an X-linked defect in the CD40 ligand (CD40L (ligand); CD154) gene. Signalling between CD40 on B cells and CD40L on T cells is necessary for the generation of class switched B cells bearing high affinity immunoglobulin, as well as for the maturation of the T cell response. In CD40L deficiency, B cells are trapped at the immature level producing only IgM. Levels of other types of immunoglobulin IgG and IgA are absent/low and there are mild defects in T cell function.

Opportunistic infections occur, for example pneumocystis, cryptosporidium, herpes viral infections, candida and cryptococcus.

Wiskott–Aldrich syndrome. This is an X-linked defect (at Xp11–23) in a gene involved in signal transduction and cytoskeletal function with associated eczema and thrombocytopenia; a mainly cell-mediated defect with falling immunoglobulins is seen, and autoimmune manifestations and lymphoreticular malignancy may develop.

Ataxia telangiectasia. These patients have defective DNA repair mechanisms and have cell-mediated defects with low IgA and IgG_2; lymphoid malignancy is again common.

EBV associated immunodeficiency (Duncan's syndrome). Apparently normal, but genetically predisposed (usually X-linked) individuals develop overwhelming EBV infection, polyclonal EBV-driven lymphoproliferation, combined immunodeficiency, aplastic anaemia and lymphoid malignancy. EBV appears to act as a trigger for the expression of a hitherto silent immunodeficiency.

B lymphocyte deficiency

X-linked agammaglobulinaemia (XLA) presents soon after maternal IgG protection falls, with recurrent infections of the upper and lower respiratory tract with pyogenic organisms, enteritis and malabsorption. Morbidity and mortality are high, mainly due to chronic lung disease and CNS infections with enteroviruses. IgG levels are usually much less than 2.0 g/L (a 1–2-year-old child will usually have an IgG level of 3.5–14 g/L) and all five classes of immunoglobulins are affected, with total levels of less than 2.5 g/L (typically >4 g/L at age 1–2 years). B cell development is arrested at the pre-B stage and B cells are absent in the blood. The cause is a loss of function mutation on chromosome Xq22 affecting the gene for a tyrosine kinase involved in cell activation and maturation.

Treatment is by i.v. immunoglobulins (IVIg) to a level that controls infections. Trough IgG concentrations should remain within normal limits (i.e. >5–6 g/L). Prognosis has improved in recent years as more patients survive into adulthood, but chronic lung disease and lymphomas are life-threatening complications.

Selective IgA deficiency (IgAD; serum IgA <0.05 g/L), with normal levels of IgG and IgM, is found in approximately 1/600 northern Europeans, making it the most common primary immunodeficiency, although the underlying defect is unknown. Patients typically present at any age with recurrent infections caused by pyogenic organisms and affecting mucosal sites. ***Treatment*** mainly comprises antibiotics. Some IgAD patients produce anti-IgA antibodies of the IgG and IgE class. Infusion of exogenous IgA (e.g. during a blood transfusion) could therefore result in anaphylaxis. IgAD patients should therefore be screened for anti-IgA antibodies, and transfused if necessary with washed red cells, blood from an IgAD donor or with stored aliquots of their own blood.

Common variable immunodeficiency (CVID) is a heterogeneous disease affecting 1/50 000 and arising during late childhood and early adulthood, with IgG levels <0.5 g/L. Presenting with recurrent upper and lower respiratory tract infections is typical and may progress to chronic bronchiectatic lung disease, malabsorption and diarrhoea. There may be additional features of T cell deficiency. Genetic defects underlying the condition have revealed a mutation in the gene encoding ICOS-L (ligand for inducible co-stimulator on activated T cells in some cases). IVIg and antibiotics are the treatments of choice.

Defects in antigen presenting cell function

A series of rare genetic defects have been uncovered in which APCs demonstrate an inability to mount protective responses to intracellular bacteria, particularly low virulence mycobacteria and salmonella. The axis affected is the interplay between CD4 T cells and APCs which drives Th1 responses, and therefore in turn acivates mononuclear cells such as macrophages to kill and eradicate intracellular pathogens. Defective genes so far identified include those encoding a component chain of IL-12; a component chain of the IL-12 receptor; and IFN-γ receptor chains 1 or 2.

Neutrophil defects

Chronic granulomatous disease (CGD) is a rare (1/250 000) immunodeficiency due to a defect in neutrophil killing and characterized by deep-seated infections. The functional defect is an inability to generate antibacterial metabolites through the respiratory burst (see Table 3.4). Typical onset is at toddler age. Neutrophil numbers are normal or increased. A simple respiratory burst test of neutrophil function is diagnostic. ***Treatment*** of infections and prophylactic antibiotic and antifungal therapy are required. Immunotherapy with interferons may have a role in patients with intractable infections, and in some cases SCT is required.

Leucocyte adhesion deficiency (LAD) results from defects in integrins (see p. 26). Numerous underlying defects in the gene encoding one of the component chains, CD18, have been described in LAD-I. LAD-I has an autosomal recessive inheritance presenting almost immediately after birth, with delayed umbilical cord separation. Recurrent infections similar to those in chronic granulomatous disease appear during the first decade of life. Blood neutrophil levels are high but cells are absent from the sites of infection, which require aggressive antimicrobial and antifungal treatment, and SCT for cure.

Hyper IgE syndrome (or **Job's syndrome**) is an autosomal dominant (occasionally sporadic) immune disorder with high serum IgE levels, dermatitis, boils, pneumonias with cyst formation and bone and dental abnormalities. Mutations in STAT3 have been found.

Shwachman's syndrome can resemble cystic fibrosis clinically with exocrine pancreatic insufficiency and pyogenic infections. A mild neutropenia is associated with a defect of neutrophil migration.

Chédiak–Higashi syndrome is a rare, recessive disorder due to a mutation in the lysosomal trafficking regulator gene (LYST) on chromosome 1q42–45. There are defects in neutrophil function with defective phagolysosome fusion and large lysosome vesicles are seen in phagocytes. Patients have recurrent infections, neutropenia, anaemia and hepatomegaly. Similar abnormalities in melanocytes cause partial oculocutaneous albinism.

Complement deficiency

The consequences of deficiency of complement proteins can be predicted from their functions (see Fig. 3.2).

- Failures in innate response components (e.g. the alternative pathway) lead to impaired non-specific immunity with an increase in bacterial infections.
- Genetically determined low levels of mannose binding lectin (MBL) are associated with a number of inflammatory and infectious diseases.
- Failure of the classical pathway results in a tendency towards infection but also towards diseases in which immune complex deposition causes inflammation, such as systemic lupus erythematosus (e.g. C1 and C4 deficiency), vasculitis and glomerulonephritis.

An unexpected finding is that neisserial infections (e.g. meningitis due to *N meningitidis*) are often encountered in patients with complement defects of the membrane attack complex.

Complement regulatory proteins. Deficiency of C1 inhibitor *(C1 esterase) deficiency* (see also p. 1244) is relatively rare. Since this enzyme is involved in regulation of several plasma enzyme systems (e.g. the kinin system) and is continuously consumed, a single parental chromosome defect resulting in 50% of normal production barely copes with the demand and fails under stress (hence an autosomal dominant effect). As a result, uncontrolled activation of com-

plement and the kinins may occur, leading to oedema of the deep tissues affecting the face, trunk, viscera and the airway, hence the alternative name of hereditary angioedema (HAE).

TYPE I HYPERSENSITIVITY AND ALLERGIC DISEASE

Normally host defence can cope with potentially harmful cells and molecules. Under some circumstances, a harmless molecule can initiate an immune response that can lead to tissue damage and death. Such exaggerated, inappropriate responses are termed hypersensitivity reactions or allergic disease.

In Type I (immediate) hypersensitivity the binding of an antigen to specific IgE bound to its high-affinity receptor on a mast cell surface results in massive and rapid cell degranulation and the inflammatory response outlined on page 57. The antigens involved are typically inert molecules present in the environment (termed allergens; see p. 831 and http://www.allergen.org/allergen.aspx). The immediate effects of allergen exposure are often very florid (early phase response), allergic disorders also have a second phase, occurring a few hours after exposure and lasting up to several days. These 'late phase responses' (LPR) are mediated by Th2 cells recognizing peptide epitopes of the allergen. Recruitment of eosinophils is often a prominent feature. From a pathological and therapeutic viewpoint, the LPR gives rise to chronic inflammation which is difficult to control. *In asthma* the LPR gives rise to the prolonged wheezing that can be fatal. Immediate hypersensitivity is usually responsive to antihistamines but the LPR is not, requiring powerful immune modulators such as corticosteroids.

In immunopathological terms neutrophils and eosinophils are prominent in the first 6–18 hours of the LPR and may persist for 2–3 days. Th2 cells accumulate around small blood vessels and persist for 1–2 days. Mediators responsible for the cellular infiltrate include platelet activating factor and leukotrienes (Table 3.17), whilst the Th2 cytokines IL-4 and IL-5 and chemokines such as eotaxin act as growth and activation stimuli for eosinophils, which are capable of extensive tissue damage. Th2 cytokines are also responsible for the class switch of Ig production towards IgE, maintaining the cycle of immediate and late responses.

What makes an allergen so powerful? Several allergens are proteolytic enzymes allowing them to cross skin and mucosal barriers. They are often contained within small, aerodynamic particles (e.g. pollen grains) that gain access to nasal and bronchial mucosa.

Why do some people react and others not? The tendency to develop allergic responses (known as atopy) shows strong heritability. Between 20% and 30% of the UK population is atopic and two, one or no atopic parents pass on the atopic trait to their children with a risk of 75%, 50% and 15%, respectively. Amongst the predisposing genes are those encoding the β chain of the high affinity receptor for IgE and IL-4, both strongly associated with Th2 pathways. The presence of Th2 cells recognizing allergens is the pathological hallmark of allergy.

What environmental factors are involved? Early exposure to allergens (even in utero) may be a factor in developing atopy. Over-zealous attention to cleanliness (the hygiene hypothesis) in developed societies (use of antibiotics, reduced exposure to pathogens which might favour a Th1-like environment) may favour a Th2 response. This environmental factor is shown by the rapid increase of allergy in the Eastern part of Germany following reunification.

In clinical terms, approximately two-thirds of atopic individuals (who can be identified as those with circulating allergen-specific IgE) have clinical allergic disease (equating to 15–20% of the UK population). Allergy accounts for up to one-third of school absences because of chronic illness. Other allergic disorders include allergic rhinitis (hay fever), allergic eczema, bee and wasp venom allergy and some forms of food allergy, urticaria and angioedema.

Diagnosis of allergic disease is usually made on the history and backed up by skin prick testing (insertion of a tiny quantity of allergen under the skin and measurement of the size of the weal) and/or measurement of allergen-specific IgE. Mast cell tryptase serum levels peak 1–2 hours after an event remaining high for 24 hours.

Table 3.17 Mediators involved in the allergic response

Preformed mediators
Histamine and serotonin:
- Bronchoconstriction
- Increased vascular permeability

Neutrophil and eosinophil chemotactic factors (NCF and ECF):
- Induce inflammatory cell infiltration

Newly formed mediators (membrane-derived)
Leukotriene (LT) B_4:
- Chemoattractant

LTC_4, $-D_4$, $-E_4$ (slow reacting substance of anaphylaxis, SRS-A):
- Sustained bronchoconstriction and oedema

Prostaglandins and thromboxanes:
- Platelet-activating factor (PAF)
- Prolonged airway hyperactivity

Treatment

Avoidance is the first line of therapy.

- *Antihistamines* are effective for many immediate hypersensitivity reactions (but have no role in the treatment of asthma). *A monoclonal antibody* that mops up serum IgE (omalizumab) is also used (see p. 856).
- *Corticosteroids* have several well-identified modifying actions in the allergic process: production of prostaglandin and leukotriene mediators is suppressed, inflammatory cell recruitment and migration is inhibited and vasoconstriction leads to reduced cell and fluid leakage from the vasculature.
- *Cysteinyl leukotriene receptor antagonists (LTRAs)* inhibit leukotrienes (LTs) by blocking the type I receptor (e.g. montelukast – used in asthma, particularly aspirin induced).
- *Desensitization (allergen immunotherapy)* can be used. The principle is that allergy can be prevented by inoculation by giving the allergen in a controlled way. It is indicated for disorders in which the hypersensitivity is IgE mediated, e.g. life-threatening allergy to insect stings, drug allergy and allergic rhinitis. An induction

Table 3.18 Mechanism of action of desensitization for allergy

Probable mechanism	Effect
IgG blocking antibodies	During repeated exposure to desensitizing allergen, IgG class antibodies develop (especially IgG_4 subclass); these compete with the pathogenic IgE for allergen binding, and/or prevent IgE–allergen complexes binding to mast cell high affinity IgE receptors
Regulation	Exposure to repeated desensitizing allergen induces T_{reg} cells which recognize allergen but invoke regulatory immune responses, dampening down migration, infiltration and inflammation
Immune deviation	A shift away from Th2- to Th1-producing CD4 cells results in the generation of cytokines (e.g. IFN-γ) which are inhibitory to IgE production

course of subcutaneous injections of increasing doses of the allergen extract, given once every 1–2 weeks, is followed by maintenance injections monthly for 2–3 years. A recent systematic review of 51 published randomized placebo-controlled clinical trials, enrolling a total of nearly 3000 participants, showed a low risk of adverse events with consistent clinical benefit. From an immunological viewpoint, desensitization seems capable of modifying the allergic response at several levels (Table 3.18). A new development has been sublingual allergen immunotherapy and early results (using grass pollen extract tablets of Phl p 5 from Timothy grass are promising) (see p. 833).

Anaphylaxis

The term anaphylaxis describes 'a serious allergic reaction that is rapid in onset and may cause death'. It arises as an acute, generalized IgE-mediated immune reaction involving specific antigen, mast cells and basophils. The reaction requires priming by the allergen, followed by re-exposure. To provoke anaphylaxis, the allergen must be systemically absorbed, either after ingestion or parenteral injection and a range of allergens have been identified:

- *foods*: nuts (peanuts (protein-arachis hypogaea Ara h 1–3), Brazil, cashew), shellfish (shrimp (allergen Met e 1), lobster), dairy products, egg (and more rarely citrus fruits, mango, strawberry, tomato)
- *venoms*: wasps, bees, yellow-jackets, hornets
- *medications*: antisera (tetanus, diphtheria), dextran, latex, some antibiotics.

Anaphylaxis is rare, and the symptom/sign constellation ranges from widespread urticaria to cardiovascular collapse, laryngeal oedema, airway obstruction and respiratory arrest leading to death. Fatal reactions to penicillin occur once every 7.5 million injections; between 1 in 250 and 1 in 125 individuals have severe reactions to bee and wasp stings, and a death takes place every 6.5 million stings; such stings cause between 60 and 80 deaths per year in North America, and 5–10 in the UK. Central to the pathogenesis of anaphylaxis is the activation of mast cells and basophils, with systemic release of some mediators and generation of others. The initial symptoms may appear innocuous: tingling, warmth and itchiness. The ensuing effects on the vasculature give vasodilatation and oedema. The consequence of these may be no more than a generalized flush, with urticaria and angio-oedema. More serious sequelae are hypotension, bronchospasm, laryngeal oedema and cardiac arrhythmia or infarction. Death may occur within minutes. Serum platelet-activating factor (PAF) levels correlate directly with the severity of anaphylaxis whereas PAF acetylhydrolase (the enzyme that inactivates PAF) correlated inversely and was significantly lower in peanut sensitive patients with fatal anaphylactic reactions.

Early recognition and treatment are essential (Emergency Box 3.1).

The best treatment is prevention. Avoidance of triggering foods, particularly nuts and shellfish, may require almost obsessive self-discipline. Patient education is necessary and many are instructed in the self-administration of adrenaline (epinephrine) and carry pre-loaded syringes. Desensitization has a well-established place in the management of this disorder, particularly if exposure is unavoidable or unpredictable, as in insect stings.

Emergency Box 3.1

Treatment of acute anaphylaxis

Clinical features:
- Bronchospasm
- Facial and laryngeal oedema
- Hypotension
- Nausea, vomiting and diarrhoea.

Management:
- Position the patient lying flat with feet raised
- Ensure the airway is free
- Give oxygen
- Monitor BP
- Establish venous access
- Administer 0.5 mg intramuscular epinephrine (adrenaline) and repeat every 5 min if shock persists.
- Administer intravenous antihistamine (e.g. 10–20 mg chlorphenamine) slowly
- Administer 100 mg intravenous hydrocortisone.

If hypotension persists, give 1–2 L of intravenous fluid. If hypoxia is severe, assisted ventilation may be required.

AUTOIMMUNE DISEASE

Autoimmunity is when the immune response turns against self, i.e. recognizes 'self' antigens. The vast array of possible TCRs and antibodies that can be generated by the host make it highly probable that at least a small proportion can recognize self (i.e. are autoreactive). Moreover, a degree of autoreactivity is physiological – the TCR is designed to interact both with the peptide epitope in the HLA molecule binding groove, *and* with the HLA molecule itself.

The critical event in the development of autoimmune disease is when T and B lymphocytes bearing these recep-

tors become activated. The following are the major checkpoints that the immune system has in place to prevent this:

1. removal of TCRs with very strong affinity for 'self' in the thymus
2. the presence of naturally arising regulatory T cells (Tregs)
3. the requirement for a danger signal to licence dendritic cells to activate CD4 T cells.

Autoimmune diseases affect 5% of the population at some stage of their life.

Failure of Checkpoint 1, thymic education

During thymic education, TCRs with dangerously high affinity for self are deleted. It has become apparent that such a system relies upon the thymic expression of self antigens. Situations which compromise the expression of a self protein would be expected to favour the development of autoimmunity. In a rare group of patients who develop multiple autoimmune disorders affecting the adrenal and parathyroid glands (Autoimmune Polyglandular Syndrome Type 1; APS Type 1: see p. 1026), there is a defect in the *AIRE* gene (for autoimmune regulator) which controls thymic expression of a host of self genes. When the gene malfunctions, there is reduced expression of self proteins in the thymus and autoimmune disease is a consequence.

Failure of Checkpoint 2, regulatory T cells

An example of Treg failure is the defect in the gene encoding Foxp3, a critical transcription factor in Tregs (see p. 65) that leads to IPEX (see p. 66). IPEX is very rare, but it serves to indicate how Treg defects can lead to autoimmune disease. Laboratory studies in this area are revealing subtle Treg defects in several autoimmune diseases (e.g. type 1 diabetes, multiple sclerosis, rheumatoid arthritis).

Failure of Checkpoint 3, CD4 T cell activation against an autoantigen (or its mimic)

For an autoimmune disease to develop, there must be presentation of autoantigens to a naïve, potentially autoreactive CD4 T cell by activated DCs. This could happen in one of two ways:

- Tissue damage due to infection leads to both the release of hidden self antigens and the provision of sufficient danger signals to activate DCs, which in turn activate autoreactive CD4 T cells as well as the pathogen-specific ones. This is often termed 'bystander activation'.
- A pathogen mimics a self antigen. In the process of making an entirely appropriate immune response against the pathogen, T or B cells are generated that also have the capacity to recognize self. This is termed molecular mimicry.

It is unlikely that for the common autoimmune diseases (Table 3.19) there is a 'single checkpoint' explanation. Rather, it is likely that multiple subtle defects, at various checkpoints, are at play.

Table 3.19 Some autoimmune diseases and their autoantigens

Disease	Antigens
Addison's disease	21α-hydroxylase
Goodpasture's syndrome	Type IV collagen (located in GBM)
Graves' thyroiditis	Thyroid-stimulating hormone receptor
Hashimoto's thyroiditis	Thyroid peroxidase, thyroglobulin
Multiple sclerosis	Myelin basic protein Myelin oligodendrocyte glycoprotein
Myasthenia gravis	Acetylcholine receptor
Pemphigus vulgaris	Desmoglein-3
Pernicious anaemia	H^+/K^+-ATPase, intrinsic factor
Polymyositis/ dermatomyositis	tRNA synthases
Primary biliary cirrhosis	Pyruvate dehydrogenase complex
Rheumatoid arthritis	Citrullinated cyclic peptide, IgM
Scleroderma	Topoisomerase
Sjögren's syndrome	Ro/La ribonuclear proteins
Systemic lupus erythematosus	Sm/RNP, Ro/La (SS-A/SS-B), histone and native DNA
Type 1 diabetes	Proinsulin, glutamic acid decarboxylase, IA-2, ZNT8
Vitiligo	Pigment cell antigens
Wegener's granulomatosis	Neutrophil proteinase 3

GBM, glomerular basement membrane.

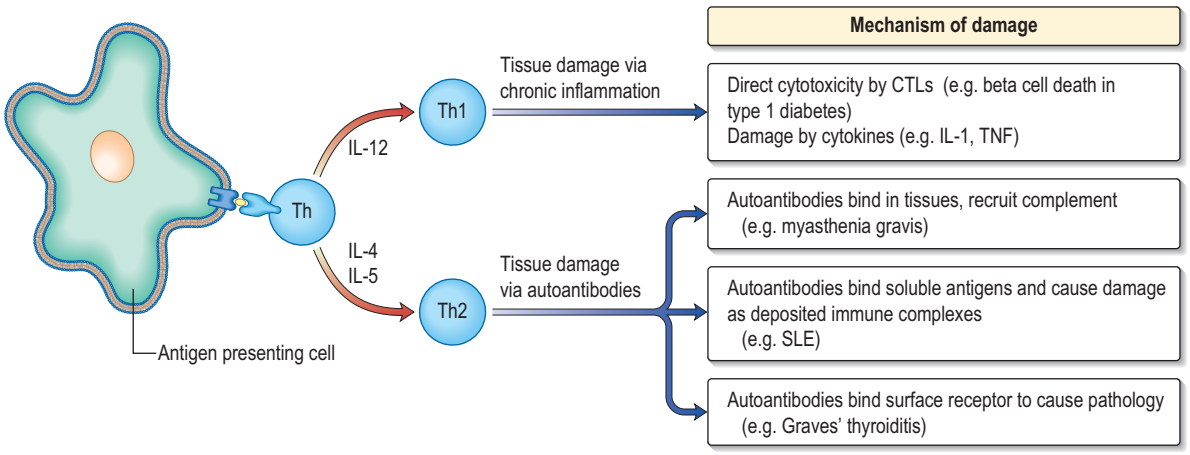


Fig. 3.14 **Potential mechanisms of immune damage in autoimmune disease.**

Mechanisms of tissue damage in autoimmune disease

The scheme shown in Figure 3.14 illustrates potential mechanisms of immune damage in autoimmune disease.

Autoantibodies can cause damage directly, through direct binding to a target cell or structure, with recruitment of complement and other destructive processes. Disease examples of this process include:

- *autoimmune haemolytic anaemia* (autoantibody targets red blood cell autoantigens leading to lysis)
- *myasthenia gravis* (autoantibodies against the acetylcholine receptor cause damage to the neuromuscular junction)
- *Goodpasture's syndrome* (antiglomerular basement membrane autoantibodies damage glomerular integrity).

Autoantibody binding may not only cause damage – in Graves' thyroiditis, for example, the antithyroid stimulating hormone receptor antibody stimulates follicular cells to produce thyroid hormones, leading to hyperthyroidism.

Autoantibodies can also bind their antigen in the circulation to form immune complexes. When these become deposited in the tissues, complement and cell-mediated immune reactions are initiated. Immune complexes preferentially deposit in sites such as the kidney glomerulus, leading to chronic kidney disease. This is a feature of SLE, in which the autoantigen within the immune complexes is DNA.

Common autoimmune diseases

There are now over 80 diseases classified as autoimmune. Some of the diseases are very clear cut in their autoimmune origin, but others are not and both major and minor factors must be defined.

Major criteria:

- Evidence of autoreactivity (e.g. activated or memory autoreactive T cells or autoantibodies)
- clinical response to immune suppression
- passive transfer of the putative immune effector (e.g. autoreactive T cell or autoantibody) causes the disease (hardest criterion to satisfy in man but most stringent).

Minor criteria:

- An animal model exists that resembles the human condition, and in which there is a similar loss of immunological tolerance to self
- evidence that, in the animal model, passive transfer of the putative immune effectors reproduces the disease in a naïve animal
- HLA association (frequent indicator that a disease is autoimmune).

ORGAN REJECTION IN CLINICAL TRANSPLANTATION

The outcome of an allograft (i.e. a graft between genetically non-identical members of the same species) in the absence of adequate immunosuppressive therapy is *immunological rejection*. Histological analysis of rejected organs shows a range of immunological processes in action (Table 3.20). With modern tissue-matching approaches, hyperacute rejection is rare and acute rejection can usually be prevented or treated with immunosuppression. The process of chronic rejection typically takes place over several years and is the main reason for organ graft failure.

The antigens recognized in acute and chronic graft responses are donor HLA molecules (alloantigens). It is thought that in acute rejection, the predominant response is against intact HLA molecules (called direct allorecognition) and is mediated by $CD4^+$ and $CD8^+$ T lymphocytes. As the rejection process becomes more chronic, peptides from donor HLA molecules are processed and presented to T cells by host HLA molecules (indirect allorecognition).

As a consequence of the vigourous immunosuppression required to avoid allograft rejection, secondary immunodeficiency often arises in graft recipients, manifesting as infections and malignancies. Cytomegalovirus is particularly common, along with herpes viruses, fungal organisms (*Aspergillus* and *Nocardia* spp., cryptococcal infection), mycobacteria and parasites (*Pneumocystis, Toxoplasma* spp.). The most common cancers seen after organ transplantation are lymphoma, skin cancer and Kaposi's sarcoma.

IMMUNE-BASED THERAPIES

Manipulating the immune response in a therapeutic setting has seen many successes, as evidenced by the control of organ rejection in clinical transplantation through targeted immunosuppression (Table 3.20). Monoclonal antibodies offer the opportunity to neutralize the unwanted effects of

Table 3.20 Classification of rejection

Description	Timing of response	Immunological mechanism	Management
Hyperacute	Minutes to hours	Preformed antibodies (e.g. anti-blood group, anti-HLA)	Prevention
Accelerated acute	1–5 days	T lymphocytes	Combinations of immune suppression used prophylactically and as required: e.g. corticosteroids, ciclosporin, tacrolimus, sirolimus, polyclonal antibodies (e.g. against all T lymphocytes) and monoclonal antibodies (e.g. against activated T cells expressing CD25)
Acute	7–14 days onwards	T lymphocytes	
Chronic	Months to years	Antibodies, complement, endothelial cell changes	No specific therapy

Ask the authors
www.studentconsult.com

cytokines, or to direct immune responses, drugs, toxins or irradiation against a specific target, whether it be a tumour cell or an immune cell involved in a damaging autoimmune response. Natural antiviral mediators, such as the interferons, are already in the clinic as therapies for, amongst other things, chronic viral infection.

Monoclonal antibody therapy

The combined power of monoclonal antibody (MAb) and recombinant DNA technology has led to a series of 'designer' drugs which have been engineered so that they are:

(a) exquisitely targeted
(b) have optimal effector function, and
(c) do not carry antigenic segments that may incite a neutralizing response in the host.

In general this has meant a process of 'humanizing' antibodies of mouse origin and selecting an appropriate effector function. For example, a MAb designed to remove a subpopulation of lymphocytes from the patient should have good complement-fixing ability or bind well to receptors on phagocytes, whereas a MAb designed to 'modulate' a cell without depletion should have these functions removed. An example of the latter is an anti-CD3 MAb that has modified Fc regions and is designed to functionally downregulate, but not deplete, T cells. Examples of MAbs in current use are shown in Table 3.21.

Others include efalizumab (inhibits T cell activation) in psoriasis; omalizumab (binds to IgE) in asthma; basiliximab and daclizumab (anti-CD25 preventing T cell proliferation) in acute transplant rejection (p. 640).

Immunosuppressive drugs

Glucocorticosteroids (cortisone, hydrocortisone, prednisone and prednisolone) are the most commonly used steroids and have a variety of effects on immune function. These include potent effects on monocyte production of the proinflammatory cytokines IL-1 and TNF-α; blockade of T cell production of IL-2 and IFN-γ; and reduced activation and migration of a range of innate and adaptive immune cells.

Ciclosporin, *tacrolimus* and *rapamycin* (sirolimus) have similar effects on T cell function. Ciclosporin and tacrolimus are calcineurin inhibitors and inhibit Ca^{2+}-dependent second messenger signals in T cells following activation via TCR. By contrast, sirolimus achieves a similar effect but acts at the level of post-activation events in the nucleus.

Purine analogues such as azathioprine are also frequently used as anti-inflammatory drugs in conjunction with steroids and act by inhibiting DNA synthesis in dividing adaptive immune cells. Similar in mode of action, but more powerful, is mycophenolate mofetil (MMF).

Alkylating agents that interfere with DNA synthesis, such as cyclophosphamide, are also used for immunosuppression.

Table 3.21 Examples of monoclonal antibodies in therapeutic use*

Name	Target	Clinical application
OKT3	CD3	Graft rejection
Rituximab	CD20	Non-Hodgkin's lymphoma; also successful in autoimmune disease setting (e.g. rheumatoid arthritis) and multiple sclerosis
Basiliximab, Daclizumab	CD25 (receptor for IL-2)	Prophylaxis against acute transplant rejection
Infliximab	TNF-α	Rheumatoid arthritis, Crohn's disease
Omalizumab	IgE	Asthma
Efalizumab	CD11a (block binding of LFA-1 to ICAM-1)	Psoriasis

For further monoclonal therapy for malignancy, see p. 464.
*For use in oncology, see Table 9.11.

Cytokines and anticytokines

Cytokines are pleiotropic agents with powerful proinflammatory and immunosuppressive effects and are attractive targets for therapies that inhibit or enhance their function. In addition to TNF-α targeting, agents that block IL-1 have also been tested in rheumatoid arthritis, although with much less success. Nonetheless, the good safety profile of the drug (recombinant IL-1 receptor antagonist; IL-1ra, Anakinra) prompted its use in other diseases, and a recent study showed a beneficial effect in type 2 diabetes, which may have an innate inflammatory component.

IL-2 and its receptor (CD25) are also obvious candidates for immune therapies, and in addition to blocking function (e.g. through anti-CD25), some IL-2-based strategies are aimed to boost immunity, for example against tumours. Interferons have not only been used to boost proinflammatory immunity, but also as immune modulators (e.g. IFN-β in the autoimmune disease multiple sclerosis), although their efficacy in this respect remains to be fully established.

Restoring tolerance in autoimmune diseases and allergy

One of the goals for immune-based therapies for autoimmune disease is not to simply achieve immunosuppression, but also to restore immunological tolerance against the relevant autoantigens. In animal models an effective means of achieving this is to administer the autoantigen itself, or key peptide epitopes from it. This is known as antigen-specific or peptide immunotherapy and is under trial in several autoimmune diseases. Giving antigen without danger signals (i.e. no co-stimulation) could provide an excessive amount of TCR-dependent signal 1 in the absence of signal 2 (p. 59) – this is known to lead to non-responsive T cells. There is also evidence that this safe antigen administration biases T cells towards a regulatory phenotype. The diseases that have the most advanced clinical data are in the field of allergy. For example, administration of cocktails of peptides of the cat allergen Fel d 1 has led to a reduction in detectable skin-prick responses and improved clinical scores.

Intravenous immunoglobulin

Intravenous immunoglobulin (IVIg) is a preparation of polyspecific IgG chemically purified from the plasma of large numbers (>20 000) of healthy donors. IVIg is used as a

replacement therapy in patients with primary and secondary antibody deficiencies. However, when used for inflammatory conditions there is also a therapeutic benefit, although randomized placebo-controlled studies are few. In the US, IVIg is only recommended for a small number of diseases in addition to antibody deficiency. These include immune-mediated thrombocytopenia; Kawasaki syndrome, chronic inflammatory demyelinating polyneuropathy and post-transfusion purpura. The mechanism of action is not known, but may include blockade of Fc receptors to prevent pathogenic antibodies binding to phagocytes; inhibition of autoantibody synthesis by B cells; inhibition of complement activation; induction of T cell regulation.

CHAPTER BIBLIOGRAPHY

Breart B, Bousso P. Cellular orchestration of T cell priming in lymph nodes. *Current Opinion in Immunology* 2006; **18**: 483–490.

Corthay A. A three-cell model for activation of native T helper cells. *Scandinavian Journal of Immunology* 2006; **64**: 93–96.

Goodnow C. Multistep pathogenesis of autoimmune disease. *Cell* 2007; **130**: 25–35.

O'Connor GM, Hart OM, Gardiner CM. Putting the natural killer cell in its place. *Immunology* 2006; **117**: 1–10.

Ogawa Y, Calhoun WJ. The role of leukotrienes in airway inflammation. *Journal of Allergy and Clinical Immunology* 2006; **118**: 789–798; quiz 799–800.

Peters-Golden M, Henderson WR. Leukotrienes. *New England Journal of Medicine* 2007; **357**: 1841–1854.

Reis e Sousa C. Dendritic cells in a mature age. *Nature Reviews. Immunology* 2006; **6**: 476–483.

Sakaguchl S, Wing K, Miyara M. Regulatory T cells – a brief history and perspective. *European Journal of Immunology* 2007; **37** Suppl 1: S116–123.

Wood P, Stanworth S, Burton J et al. Recognition, clinical diagnosis and management of patients with primary antibody deficiencies: a systematic review. *Clinical and Experimental Immunology* 2007; **149**: 410–423.

SIGNIFICANT WEBSITES

http://www.antibodyresource.com/
Source of antibody resources for educators and researchers

http://www.microbiologybytes.com/iandi/16.html
MCQs and immunology glossary (University of Leicester)

http://www.copewithcytokines.org
Information resource for cytokines and chemokines

http://www-immuno.path.cam.ac.uk/~immuno/part1.html
Immunology resource from Department of Pathology, Immunology Division, University of Cambridge

This page intentionally left blank

4 Infectious diseases, tropical medicine and sexually transmitted infection

Infection and infectious disease 79

Principles and basic mechanisms 84

Approach to the patient with a suspected infection 87

Antimicrobial chemotherapy 91
Antibacterial drugs 94
Antituberculosis drugs 100
Antifungal drugs 100
Antiviral drugs 100

Immunization against infectious diseases 102

Viral infections 104

DNA viruses 104
Adenoviruses 104
Herpesviruses 104
Papovaviruses 108
Erythroviruses 109
Poxviruses 109
Hepadna viruses 109

RNA viruses 109
Picornaviruses 109
Reoviruses 112
Caliciviruses 112
Togaviruses 113
Flaviviruses 114
Bunyaviruses 116
Orthomyxoviruses 116
Paramyxoviruses 117
Coronaviruses 119
Rhabdoviruses 119
Retroviruses 120
Arenaviruses 120

Filoviruses 121

Postviral/chronic fatigue syndrome 121

Transmissible spongiform encephalopathies (TSE or prion disease) 121

Bacterial infections 122
Skin and soft tissue infection 123
Respiratory tract infections 125
Gastrointestinal infections 128
Infections of the cardiovascular system 133
Infections of the nervous system 133
Bone and joint infections 135
Urinary tract infections 135
Systemic/multisystem infections 135

Bacterial infections seen in developing and tropical countries 140
Skin, soft tissue and eye disease 140
Gastrointestinal infections 143
Systemic infections 145

Fungal infections 149
Systemic fungal infections 149
Subcutaneous infections 152
Superficial infections 153

Protozoal infections 153
Blood and tissue protozoa 153
Gastrointestinal protozoa 160

Helminthic infections 163
Nematodes 163
Trematodes 168
Cestodes 170

Arthropod ectoparasites 171

SEXUALLY TRANSMITTED INFECTIONS

Clinical syndromes 175

Infestations 183

Human immune deficiency virus (HIV) and AIDS 184
Clinical features of HIV infection 187
Effects of HIV infection 188
Conditions associated with HIV immunodeficiency 189
Investigations and monitoring of HIV-infected patients 190
Management of the HIV-infected patient 191
Specific conditions associated with HIV infection 200
Prevention and control 205

INFECTION AND INFECTIOUS DISEASE

Infection remains the main cause of morbidity and mortality in man, particularly in developing areas where it is associated with poverty and overcrowding.

In the **developed world** increasing prosperity, universal immunization and antibiotics have reduced the prevalence of infectious disease. However, antibiotic-resistant strains of microorganisms and diseases such as human immunodeficiency virus (HIV) infection, variant Creutzfeldt–Jakob disease (vCJD), avian influenza and severe acute respiratory syndrome (SARS) have emerged. There is increased global mobility, both enforced (as a result of war, civil unrest and natural disaster), and voluntary (for tourism and economic benefit). This has aided the spread of infectious disease and allowed previously localized pathogens such as SARS and West Nile virus to establish themselves world-wide. An increase in the movement of livestock and animals has enabled the spread of zoonotic diseases like monkeypox, while changes in farming and food-processing methods have

contributed to an increase in the incidence of food- and water-borne diseases. Deteriorating social conditions in the inner city areas of our major conurbations have facilitated the resurgence of tuberculosis and other infections. Prisons and refugee camps, where large numbers of people are forced to live in close proximity, often under poor conditions, are providing a breeding ground for devastating epidemics of infectious disease. There are new concerns about the deliberate release of infectious agents such as smallpox or anthrax by terrorist groups or national governments.

In the **developing world** successes such as the eradication of smallpox have been balanced or outweighed by the new plagues. Infectious diseases cause nearly 25% of all human deaths (Table 4.1). Two billion people – one-third of the world's population – are infected with tuberculosis (TB), 500 million people catch malaria every year, and 200 million are infected with schistosomiasis. 39.5 million people are living with HIV/AIDS, with 4.3 million new HIV infections in 2006 (65% in sub-Saharan Africa). Infections are often multiple, and there is synergy both between different infections, and between infection and other factors such as malnutrition. Many of the infectious diseases affecting developing countries are preventable or treatable, but continue to thrive owing to lack of money and political will.

The eight Millennium Development Goals (MDG) hope to be achieved by 2015. These are: eradicating extreme poverty and hunger, universal primary education, gender equality and empowerment of women, reduction in child mortality, improvement in maternal health, combating HIV/AIDS, malaria and other diseases, ensuring environment sustainability and developing global partnership for development. The Global Fund was established to combat AIDS, tuberculosis and malaria and has achieved much by providing the means for treatment for TB, insecticide-treated bed nets for malaria and antivirals for HIV. Several other funding streams (governmental, non-governmental and charitable) have also contributed to the fight against infection.

The climate also has consequences on diseases. Both natural climatic events and the gradual global change in weather conditions are affecting the spread and transmission of infections such as malaria and cholera.

Infectious agents

The causative agents of infectious diseases can be divided into four groups.

Table 4.1 World-wide mortality from infectious diseases

Disease	Estimated deaths (annual)
Acute lower respiratory infection	3.8 million
HIV/AIDS	2.9 million
Tuberculosis	2.5 million
Diarrhoeal disease	1.8 million
Malaria	1 million
Measles	350 000
Whooping cough	301 000
Tetanus	292 000
Meningitis	175 000
Leishmaniasis	51 000
Trypanosomiasis	10 000

Prions are the most recently recognized and the simplest infectious agents, consisting of a single protein molecule. They contain no nucleic acid and therefore no genetic information: their ability to propagate within a host relies on inducing the conversion of endogenous prion protein PrP^{c} into a protease-resistant isoform PrP^{res}.

Viruses contain both protein and nucleic acid, and so carry the genetic information for their own reproduction. However, they lack the apparatus to replicate autonomously, relying instead on 'hijacking' the cellular machinery of the host. They are small (usually less than 200 nanometres in diameter) and each virus possesses only one species of nucleic acid (either RNA or DNA).

Bacteria are usually, though not always, larger than viruses. Unlike the latter they have both DNA and RNA, with the genome encoded by DNA. They are enclosed by a cell membrane, and even bacteria which have adopted an intracellular existence remain enclosed within their own cell wall. Bacteria are capable of fully autonomous reproduction, and the majority are not dependent on host cells.

Eukaryotes are the most sophisticated infectious organisms, displaying subcellular compartmentalization. Different cellular functions are restricted to specific organelles, e.g. photosynthesis takes place in the chloroplasts, DNA transcription in the nucleus and respiration in the mitochondria. Eukaryotic pathogens include unicellular protozoa, fungi (which can be unicellular or filamentous) and multicellular parasitic worms.

Other higher classes, notably the insects and the arachnids, also contain species which can parasitize man and cause disease: these are discussed in more detail on page 171.

Host–organism interactions

Each of us is colonized with huge numbers of microorganisms (10^{14} bacteria, plus viruses, fungi, protozoa and worms) with which we coexist. The relationship with some of these organisms is symbiotic, in which both partners benefit, while others are commensals, living on the host without causing harm. Infection and illness may be due to these normally harmless commensals and symbiotes evading the body's defences and penetrating into abnormal sites. Alternatively, disease may be caused by exposure to exogenous pathogenic organisms which are not part of our normal flora.

The symptoms and signs of infection are a result of the interaction between host and pathogen. In some cases, such as the early stages of influenza, symptoms are almost entirely due to killing of host cells by the invading organism. Usually, however, the harmful effects of infection are due to a combination of direct microbial pathogenicity, and the body's response to infection. In meningococcal septicaemia, for example, much of the tissue damage is caused by cytokines released in an attempt to fight the bacteria. In a few instances, such as chronic South American trypanosomiasis (Chagas' disease), morbidity is almost entirely immunological, with the parasite itself having little effect once the inflammatory process has been triggered. The molecular mechanisms underlying host–pathogen interactions are discussed in more detail on page 85.

Sources of infection

The endogenous skin and bowel commensals can cause disease in the host, either because they have been transferred to an inappropriate site (e.g. bowel coliforms causing urinary tract infection), or because host immunity has been attenuated (e.g. candidiasis in an immunocompromised host). Many infections are acquired from other people, who may be symptomatic themselves or be asymptomatic carriers. Some bacteria, like the meningococcus, are common transient commensals, but cause invasive disease in a small minority of those colonized. Infection with other organisms, such as the hepatitis B virus, can be followed in some cases by an asymptomatic but potentially infectious carrier state.

Zoonoses are infections that can be transmitted from wild or domestic animals to man. Infection can be acquired in a number of ways: direct contact with the animal, ingestion of meat or animal products, contact with animal urine or faeces, aerosol inhalation, via an arthropod vector, or by inoculation of saliva in a bite wound. Many zoonoses can also be transmitted from person to person. Some zoonoses are listed in Table 4.2.

Most microorganisms do not have a vertebrate or arthropod host but are free-living in the environment. The vast majority of these environmental organisms are non-pathogenic, but a few can cause human disease (Table 4.3). Person-to-person transmission of these infections is rare. Some parasites may have a stage of their life cycle which is environmental (e.g. the free-living larval stage of *Strongyloides stercoralis* and the hookworms) even though the adult worm requires a vertebrate host. Other pathogens can survive for periods in water or soil and may be transmitted from host to host via this route (see below): these should not be confused with true environmental organisms.

Routes of transmission

Endogenous infection

The body's own endogenous flora can cause infection if the organism gains access to an inappropriate area of the body. This can happen by simple mechanical transfer, for example colonic bacteria entering the female urinary tract. The non-specific host defences may be breached, for example by cutting or scratching the skin and allowing surface commensals to gain access to deeper tissues; this is frequently the aetiology of cellulitis. There may be more serious defects in host immunity owing to disease or chemotherapy, allowing normally harmless skin and bowel flora to produce invasive disease.

Table 4.2 Zoonotic infections

Disease	Pathogen	Animal reservoir	Mode of transmission
Prions			
vCJD	Prion protein	Cattle	Ingestion (CNS tissue)
Viral			
Lassa fever	Arenavirus	Multimammate rat	Direct contact
Avian influenza	Influenza H5N1	Birds	Inhalation
Japanese encephalitis	Flavivirus	Pigs	Mosquito bite
Rabies	Rhabdovirus	Dog and other mammals	Saliva, faeces (bats)
Yellow fever	Flavivirus	Primates	Mosquito bite
Monkey pox	Orthopox virus	Rodents, small mammals	Uncertain
SARS	Coronavirus	Unsure, most likely bats	Droplet
Bacterial			
Gastroenteritis	*Escherichia coli* 0157	Cattle, chickens	Ingestion (contaminated food)
	Salmonella enteritidis and others	Chickens, cattle	Ingestion (meat, eggs)
	Campylobacter jejuni	Various, e.g. chicken	Ingestion (meat, milk, water)
Leptospirosis	*Leptospira interrogans*	Rodents	Ingestion (urine)
Brucellosis	*Brucella abortus*	Cattle	Contact; ingestion of milk/cheese
	Brucella melitensis	Sheep, goats	
Anthrax	*Bacillus anthracis*	Cattle, sheep	Contact; ingestion
Lyme disease	*Borrelia burgdorferi*	Deer	Tick bite
Cat scratch disease	*Bartonella henselae*	Cats	Flea bite
Plague	*Yersinia pestis*	Rodents	Flea bite
Typhus	Various *Rickettsia* spp.	Various	Arthropod bite
Psittacosis (ornithosis)	*Chlamydia psittaci*	Psittacine and other birds	Aerosol
Others			
Toxoplasmosis	*Toxoplasma gondii*	Cats and other mammals	Ingestion (meat, faeces)
Hydatid disease	*Echinococcus granulosus*	Dogs	Ingestion (faeces)
Trichinosis	*Trichinella spiralis*	Pigs, bears	Ingestion (meat)
Toxocariasis	*Toxocara canis*	Dogs	Ingestion
Cutaneous larva migrans	*Ancylostoma caninum*	Dogs	Penetration of skin by larvae
Leishmaniasis	*Leishmania* spp.	Dogs	Ingestion

vCJD, variant Creutzfeldt–Jakob disease; SARS, severe acute respiratory syndrome.

Table 4.3 Environmental organisms which can cause human infection

Organism	Disease (most common presentations)
Bacteria	
Burkholderia pseudomallei	Melioidosis
Burkholderia cepacia	Lung infection in cystic fibrosis
Pseudomonas spp.	Various
Legionella pneumophila	Legionnaires' disease (pneumonia)
Bacillus cereus	Gastroenteritis
Listeria monocytogenes	Various
Clostridium tetani	Tetanus
Clostridium perfringens	Gangrene, septicaemia
Mycobacteria other than tuberculosis (MOTT)	Pulmonary infections
Fungi	
Candida spp.	Local and disseminated infection
Cryptococcus neoformans	Meningitis, pulmonary infection
Histoplasma capsulatum	Pulmonary infection
Coccidioides immitis	Pulmonary infection
Mucor spp.	Mucormycosis (rhinocerebral, cutaneous)
Sporothrix schenckii	Lymphocutaneous sporotrichosis
Blastomyces dermatitidis	Pulmonary infection
Aspergillus fumigatus	Pulmonary infections

Table 4.4 Infections transmitted by arthropod vectors

Vector	Disease	Microorganism
Mosquito	Malaria	*Plasmodium* spp.
	Lymphatic filariasis	*Wuchereria bancrofti*, *Brugia malayi*
	Yellow fever	Flavivirus
	West Nile fever	Flavivirus
	Dengue	Flavivirus
Sandfly	Leishmaniasis	*Leishmania* spp.
Blackfly	Onchocerciasis	*Onchocerca volvulus*
Tsetse fly	Sleeping sickness	*Trypanosoma brucei*
Flea	Plague	*Yersinia pestis*
	Endemic typhus	*Rickettsia typhi*
	Carrion's disease	*Bartonella bacilliformis*
Reduviid bug	Chagas' disease	*Trypanosoma cruzi*
Louse	Epidemic typhus	*Rickettsia prowazekii*
	Louse-borne relapsing fever	*Borrelia recurrentis*
Hard tick	Lyme disease	*Borrelia burgdorferi*
	Typhus (spotted fever group)	*Rickettsia* spp.
	Babesiosis	*Babesia* spp.
	Tick-borne relapsing fever	*Borrelia duttonii*
	Tick-borne encephalitis	Flavivirus
	Congo-Crimean haemorrhagic fever	Nairovirus (Bunyavirus)

Airborne spread

Many respiratory tract pathogens are spread from person to person by aerosol or droplet transmission. Secretions containing the infectious agent are coughed, sneezed or breathed out, and are then inhaled by a new victim. Some enteric viral infections may also be spread by aerosols of faeces or vomit. Environmental pathogens such as *Legionella pneumophila*, and zoonoses such as psittacosis, are also acquired by aerosol inhalation, while rabies virus may be inhaled in the dust from bat droppings.

Faeco-oral spread

Transmission of organisms by the faeco-oral route can occur by direct transfer (usually in small children), by contamination of clothing or household items (usually in institutions or conditions of poor hygiene), or most commonly via contaminated food or water. Human and animal faecal pathogens can get into the food supply at any stage. Raw sewage is used as fertilizer in many parts of the world, contaminating growing vegetables and fruit. Poor personal hygiene can result in contamination during production, packaging, preparation or serving of foodstuffs. In the western world, the centralization of food supply and increased processing of food has allowed the potential for relatively minor episodes of contamination to cause widely disseminated outbreaks of food-borne infection.

Water-borne faeco-oral spread is usually the result of inadequate access to clean water and safe sewage disposal, and is common throughout the developing world. Worldwide 1.1 billion people have no access to clean water, and 2.6 billion do not have basic sanitation.

Vector-borne disease

Many tropical infections, including malaria, are spread from person to person or from animal to person by an arthropod vector. Vector-borne diseases are also found in temperate climates, but are relatively uncommon. In most cases part of the parasite life cycle takes place within the body of the arthropod, and each parasite species requires a specific vector. Simple mechanical transfer of infective organisms from one host to another can occur, but is rare. Some vector-borne diseases are shown in Table 4.4.

Direct person-to-person spread

Organisms can be passed on directly in a number of ways. Sexually transmitted infections are dealt with on page 175. Skin infections such as ringworm, and ectoparasites such as scabies and head lice, can be spread by simple skin-to-skin contact. Other organisms are passed on by blood- (or occasionally other body fluid) to-blood transmission. Blood-to-blood transmission can occur during sexual contact, from mother to infant peripartum, between intravenous drug users sharing any part of their injecting equipment, when infected medical or other (e.g. tattoo needles) equipment is reused, if contaminated blood or blood products are transfused, or in any sporting or accidental contact when blood is spilled.

Direct inoculation

Infection can occur when pathogenic organisms breach the normal mechanical defences by direct inoculation. Some of the circumstances in which this can occur are covered under endogenous infection and blood-to-blood transmission

above. Some environmental organisms may be inoculated by accident: this is a common mode of transmission of tetanus and certain fungal infections. Rabies virus may be inoculated by the bite of an infected animal.

Consumption of infected material

Although many food-related zoonotic infections are due to contamination of food with animal faeces (and are thus, strictly speaking, faeco-oral), several diseases are transmitted directly in animal products. These include some strains of salmonella (eggs, chicken meat), brucellosis (unpasteurized milk), and the prion diseases kuru and vCJD (neural tissue).

Prevention and control

Methods of preventing infection depend upon the source and route of transmission, as described above.

- *Infection control measures*. Poor infection control practice in hospitals and other healthcare environments can cause the transfer of infection from person to person. This may be airborne, via fomites, or a direct contact route. It is essential that all healthcare workers wash or clean their hands before and after patient contact, and whenever necessary they should wear gloves, aprons and other protective equipment. This is particularly necessary when performing invasive procedures, or manipulating indwelling devices such as cannulae.
- *Eradication of reservoir*. In a few diseases, for which man is the only natural reservoir of infection, it may be possible to eliminate disease by an intensive programme of case finding, **treatment and immunization**. This has been achieved in the case of smallpox. If there is an animal or environmental reservoir complete eradication is unlikely, but local control methods may decrease the risk of human infection (e.g. killing of rodents to control plague, leptospirosis and other diseases).
- For *arthropod-vector-borne infections*:
  - destroying vector species (which may be practical in certain circumstances)
  - taking measures to avoid being bitten (e.g. insect repellent sprays, impregnated bed nets).
- For *food-borne infections*. Improvements in food handling and preparation result in less contamination during processing, transport or preparation. Organisms intrinsically present in food can be killed by appropriate preparation and cooking. Improved surveillance and regulation of the food industry, as well as better health education for the public is necessary.
- For *faeco-oral infections*. Improvements in water supply and sanitation (recognized in the Millenium Development Goals) could dramatically decrease the prevalence of faeco-oral infections.
- For *blood-borne infections*. Prevention of blood transfer, e.g. in blood transfusions and contaminated medical equipment. Donated blood is routinely tested for infection in most developed countries.
- For *infections spread by airborne and direct contact*. Some airborne-transmitted respiratory infections, and some infections spread by direct contact, can be controlled by isolating patients. This is often difficult, but isolation is useful in patients with severe immunodeficiency to protect them from infection.
- Immunization (see p. 102).

Classification of outbreaks

The type of outbreak has a bearing on public health measures that need to be instituted for its control.

- *Person to person* where infection is passed from one infected individual to another and outbreaks of infection are separated by the incubation period.
- '*Point source*' is where there is a single source of infection, e.g. food eaten at a social function. All those infected will develop symptoms at the same time, around the expected incubation period.
- *Common source* where there is a single source of infection but over a period of time, e.g. a symptomatic carrier of infection working with food preparation. Many people will be exposed over a long period of time.
- *Epidemic*. An increased unusual widespread infection in the community, causing waves of infection. These spread through communities and affect all people who have no active immunity to that infection.

Cases of some infectious diseases should be notified to the public health authorities so that they are aware of cases and outbreaks. Diseases that are notifiable in England and Wales are listed in Table 4.5.

Table 4.5 Notifiable diseases in England & Wales under the Public Health (Control of Diseases) Act 1984 and the Public Health (Infectious Diseases) Regulations 1988

Acute encephalitis
Acute poliomyelitis
Anthrax
Cholera
Diphtheria
Dysentery (amoebic or bacillary)
Food poisoning
Leprosy
Leptospirosis
Malaria
Measles
Meningitis
Meningococcal septicaemia (without meningitis)
Mumps
Ophthalmia neonatorum
Paratyphoid fever
Plague
Rabies
Relapsing fever
Rubella
Scarlet fever
Smallpox
Tetanus
Tuberculosis
Typhoid fever
Typhus
Viral haemorrhagic fever, e.g. Yellow fever
Viral hepatitis
Whooping cough

Recent infections endangering public health should also be notified (e.g. SARS, Avian flu).

FURTHER READING

Chomel B, Belotto A, Meslin FX. Wildlife, exotic pets, and emerging zoonoses. *Emerging Infectious Diseases* 2007; **13**: 6–11.

Epstein PR. Climate change and human health. *New England Journal of Medicine* 2005; **353**: 1433–1436.

Health Protection Agency. http://www.hpa.org.uk.

Jarvis WR. The United States approach to strategies in the battle against healthcare associated infections. *Journal of Hospital Infection* 2007; **65**(S2): 3–9.

WHO Infectious Diseases. http://www.who.int/topics/infectious_diseases/en/.

PRINCIPLES AND BASIC MECHANISMS

Man constantly interacts with the world of microorganisms from birth to death. The majority cause no harm and some play a role in the normal functioning of the mouth, vagina and intestinal tract. However, many microorganisms have the potential to produce disease. This may result from inoculation of damaged tissues, tissue invasion, a variety of virulence factors or toxin production.

Specificity

Microorganisms are often highly specific with respect to the organ or tissue they infect (Fig. 4.1). For example, a number of viruses are hepatotropic, such as those responsible for hepatitis A, B, C and E and yellow fever. This predilection for specific sites in the body relates partly to the presence of appropriate receptors on different cell types, and partly to the immediate environment in which the organism finds itself; for example, anaerobic organisms colonize the anaerobic colon, whereas aerobic organisms are generally found in the mouth, pharynx and proximal intestinal tract. Other organisms that show selectivity include:

- *Streptococcus pneumoniae* (respiratory tract)
- *Escherichia coli* (urinary and alimentary tract).

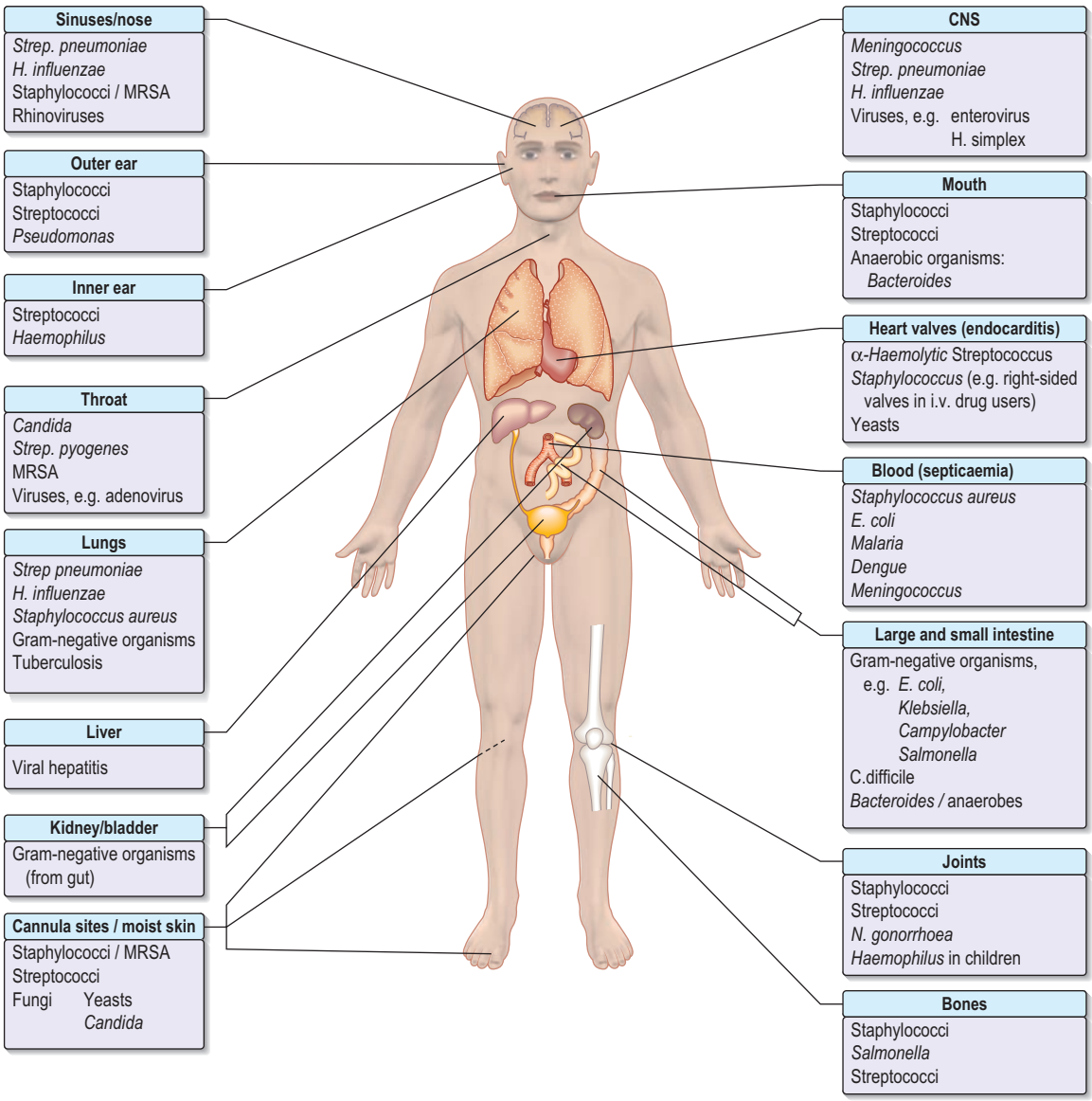


Fig. 4.1 **Sites of infection with common organisms.**

Even within a species of bacterium such as *E. coli*, there are clear differences between strains with regard to their ability to cause gastrointestinal disease (see p. 130), which in turn differ from uropathogenic *E. coli* responsible for urinary tract infection.

Within an organ a pathogen may show selectivity for a particular cell type. In the intestine, for example, rotavirus predominantly invades and destroys intestinal epithelial cells on the upper portion of the villus, whereas reovirus selectively enters the body through the specialized epithelial cells, known as M cells, that cover the Peyer's patches (see p. 274).

Pathogenesis

Figure 4.2 summarizes some of the steps that occur during the pathogenesis of infection. In addition, pathogens have developed a variety of mechanisms to evade host defences. For example, some pathogens produce toxins directed at phagocytes – *Staphylococcus aureus* (α-toxin), *Streptococcus pyogenes* (streptolysin) and *Clostridium perfringens* (α-toxin), while others such as *Salmonella* spp. and *Listeria monocytogenes* can survive within macrophages. Several pathogens possess a capsule that protects against complement activation (*Strep. pneumoniae*). Antigenic variation is an additional mechanism for evading host defences that is recognized in viruses (antigenic shift and drift in influenza), bacteria (flagella of salmonella and gonococcal pili) and protozoa (surface glycoprotein changes in *Trypanosoma*).

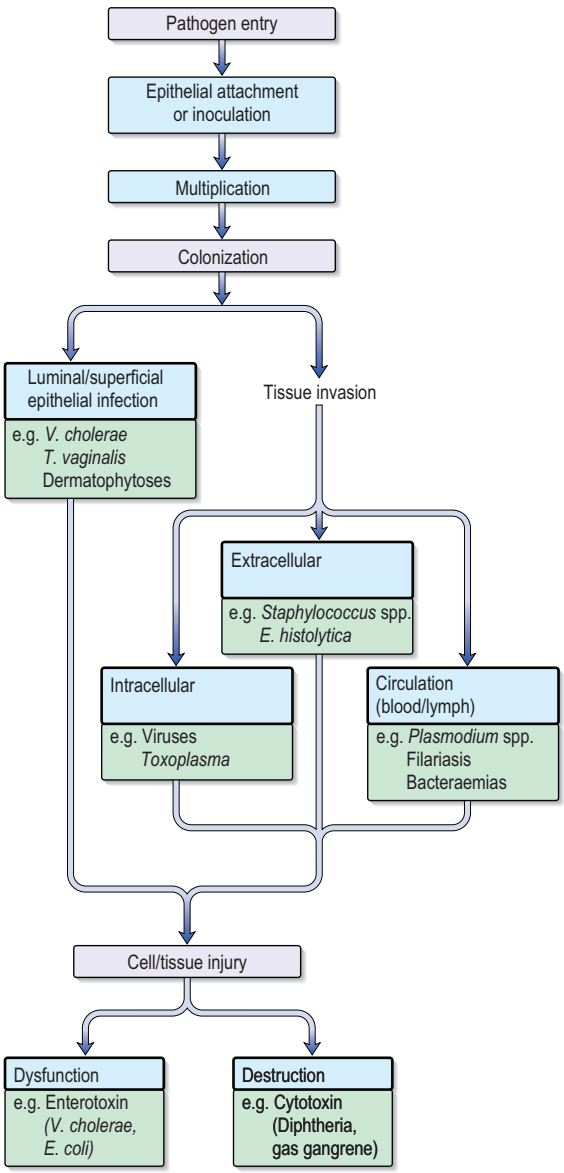


Fig. 4.2 **The pathogenesis of infection.**

Epithelial attachment

Many bacteria attach to the epithelial substratum by specific organelles called pili (or fimbriae) that contain a surface lectin(s) – a protein or glycoprotein that recognizes specific sugar residues on the host cell. This family of molecules is known as adhesins (see p. 26). Following attachment, some bacteria, such as coagulase-negative staphylococci (*Staph. epidermidis*), produce an extracellular slime layer and recruit additional bacteria, which cluster together to form a biofilm. These biofilms can be difficult to eradicate, and are a frequent cause of medical device-associated infections which affect prosthetic joints and heart valves as well as indwelling catheters. Many viruses and protozoa (e.g. *Plasmodium* spp., *Entamoeba histolytica*) attach to specific epithelial target-cell receptors. Other parasites such as hookworm have specific attachment organs (buccal plates) that firmly grip the intestinal epithelium.

Colonization

Following epithelial attachment, pathogens may either remain on the surface epithelium or within the lumen of the organ they have colonized. Tissue invasion may follow.

Invasion may result in:

- an intracellular location for the pathogen (e.g. viruses, *Mycobacterium* spp., *Toxoplasma gondii*, *Plasmodium* spp.)
- an extracellular location for the pathogen (e.g. pneumococci, *E. coli* and *Entamoeba histolytica*)
- invasion directly into the blood or lymph circulation (e.g. schistosome schistosomula and trypanosomes).

Once the pathogen is firmly established in its target tissue, a series of events follows that usually culminates in damage to the host.

Tissue dysfunction or damage

Microorganisms produce disease by a number of well-defined mechanisms:

Cell lysis

The presence of replicating viruses within a cell may interfere with host cell metabolism such that the cell dies – so-called cytolytic or cytocidal infection

Exotoxins and endotoxins

- *Exotoxins* have many diverse activities including inhibition of protein synthesis (diphtheria toxin), neurotoxicity (*Clostridium tetani* and *C. botulinum*) and enterotoxicity, which results in intestinal secretion of water and electrolytes (*E. coli*, *Vibrio cholerae*). Colonization and secretion in many classical diarrhoeal diseases is the result of virulence-associated genes which encode protein secretion systems (Fig. 4.3).
- *Endotoxin* is a lipopolysaccharide (LPS) in the cell wall of Gram-negative bacteria. It is responsible for many of

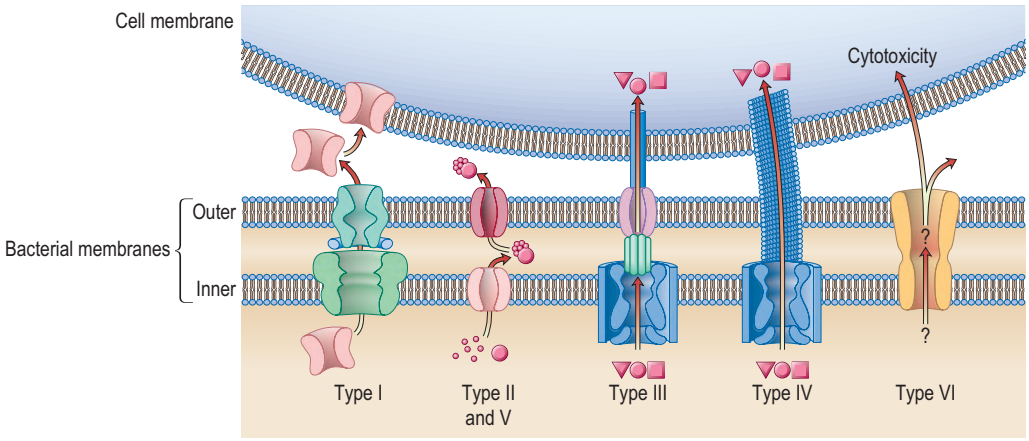


Fig. 4.3 **Protein secretion pathways for Gram-negative bacteria.** These specialized pathways are necessary for virulence. Types I and II facilitate protein (toxin) secretion into the extracellular space (**Type I**). **Type II** does this in two stages. **Types III and IV** pathways secrete toxins as well as inject toxins directly into the host cells via multiprotein complexes across the bacterial cell envelope and host cell membrane. **Type V** is a minor variation of Type II. **Type VI** pathway is unclear but is involved in cytotoxicity. Reproduced with permission from Yarh TL. A critical new pathway for toxin secretion? *New England Journal of Medicine* 2006; **355**: 1171–1172. Copyright © 2006 Massachusetts Medical Society. All rights reserved.

Table 4.6 Clinical conditions produced by *Staphylococcus aureus*

Due to invasion

Skin
Furuncles
Cellulitis
Impetigo
Carbuncles

Lungs
Pneumonia
Lung abscesses

Heart
Endocarditis
Pericarditis

Central nervous system
Meningitis
Brain abscesses

Bones and joints
Osteomyelitis, arthritis

Miscellaneous
Parotitis
Pyomyositis
Septicaemia
Enterocolitis

Due to toxin
Staphylococcal food poisoning
Scalded skin syndrome
Bullous impetigo
Staphylococcal scarlet fever
Toxic shock syndrome

Table 4.7 Examples of host factors that increase susceptibility to staphylococcal infections (predominantly *Staphylococcus aureus*)

Injury to skin or mucous membranes
Abrasions
Trauma (accidental or surgical)
Burns
Insect bites

Metabolic abnormalities
Diabetes mellitus
Uraemia

Foreign bodies*
Intravenous and other indwelling catheters
Cardiac and orthopaedic prosthesis
Tracheostomies

Abnormal leucocyte function
Job's syndrome
Chediak–Higashi syndrome
Steroid therapy
Drug-induced leucopenia

Postviral infections
Influenza

Miscellaneous conditions
Excess alcohol consumption
Malnutrition
Malignancies
Old age
Recent antibiotic therapy

*Often *Staphylococcus epidermidis.*

the features of septic shock (see p. 899), namely hypotension, fever, intravascular coagulation and, at high-doses, death. The effects of endotoxin are mediated predominantly by release of tumour necrosis factor.

Staphylococcus aureus presents an excellent example of the repertoire of microbial virulence. The clinical expression of disease varies according to site, invasion and toxin production and is summarized in Table 4.6. Furthermore, host susceptibility to infection may be linked to genetic or acquired defects in host immunity that may complicate intercurrent infection, injury, ageing and metabolic disturbances (Table 4.7).

Host response to infection

Natural defences

The natural host defences to infection are those of an intact surface epithelium with local production of secretions, antimicrobial enzymes (e.g. lysozyme in the eye) and in the stomach, gastric acidity. The mucociliary escalator of the large airways is unique to the lung.

Immunological defences

Antibody and cell-mediated immune mechanisms play a vital role in combating infection. All organisms can initiate secondary immunological mechanisms, such as complement activation, immune complex formation and antibody-mediated cytolysis of cells. The immunological response to infection is described in Chapter 3.

Metabolic and immunological consequences of infection

Fever

Body temperature is controlled by the thermoregulatory centre in the anterior hypothalamus in the floor of the third ventricle. Body temperature is maintained at 36.8°C in health, with a diurnal variation of ±0.5°C. Gram-negative bacteria contain lipopolysaccharide (LPS) and peptidoglycan, which is also a component of Gram-positive bacterial cell walls. Toll-like receptors (TLR, see p. 59) on monocytes and den-

dritic cells recognize these lipopolysaccharides and generate signals leading to formation of inflammatory cytokines, e.g. IL-1, -6, -12, TNF-α and many others. These cytokines act on the thermoregulatory centre by increasing prostaglandin (PGE_2) synthesis. The antipyretic effect of salicylates is brought about, at least in part, through its inhibitory effects on prostaglandin synthase.

Fever production has a positive effect on the course of infection. However, for every 1°C rise in temperature, there is a 13% increase in resting metabolic rate and oxygen consumption. Fever therefore leads to increased energy requirements at a time when anorexia leads to decreased food intake. The normal compensatory mechanisms in starvation (e.g. mobilization of fat stores) are inhibited in acute infections. This leads to an increase in skeletal muscle breakdown, releasing amino acids, which, via gluconeogenesis, are used to provide energy.

The inflammatory response

The inflammatory response is a fundamental biological response to a variety of stimuli including micro-organisms or their products, such as endotoxin which acts on monocytes and macrophages. Non-phagocytic cells (lymphocytes, natural killer cells) are also involved. The release of cytokines, notably TNF-α, IL-1, IL-6 and interferon-γ lead to the release of a cascade of other mediators involved in inflammation and tissue remodelling, such as interleukins prostaglandins, leukotrienes and corticotropin. TNF is therefore responsible for many of the effects of an infection.

The biological behaviour of the pathogen and the consequent host response are responsible for the clinical expression of disease that often allows clinical recognition. The incubation period following exposure can be helpful (e.g. chickenpox 14–21 days). The site and distribution of a rash may be diagnostic (e.g. shingles), while symptoms of cough, sputum and pleuritic pain point to lobar pneumonia. Fever and meningismus characterize classical meningitis. Infection may remain localized or become disseminated and give rise to the sepsis syndrome and disturbances of protein metabolism and acid–base balance (see Ch. 15). Many infections are self-limiting, and immune and non-immune host defence mechanisms will eventually clear the pathogens. This is generally followed by tissue repair, which may result in complete resolution or leave residual damage.

FURTHER READING

Bassler BL. Small talk: cell-to-cell communication in bacteria. *Cell* 2002; **109**: 421–424.

Ghannoum M, O'Toole GA. *Microbial Biofilm*. Washington DC: ASM Press, 2004.

Opal SM, Huber CE. The toll-like receptors and their role in septic shock. *Critical Care* 2002; **6**: 125–136.

Yahr TL. A critical new pathway for toxin secretion? *New England Journal of Medicine* 2006; **355**: 1171–1172.

APPROACH TO THE PATIENT WITH A SUSPECTED INFECTION

Infectious diseases can affect any organ or system, and can cause a wide variety of symptoms and signs. Fever is often regarded as the cardinal feature of infection, but not all febrile illnesses are infections, and not all infectious diseases present with a fever. History-taking and examination should aim to identify the site(s) of infection, and also the likely causative organism(s).

History

A detailed history is taken with specific questions about epidemiological risk factors for infection. These are based on the sources of infection and routes of transmission discussed above.

- Travel history: some diseases are more prevalent in certain geographical locations, and many infections common in the tropics are seen rarely if at all in the UK.
- Food and water history: systemic as well as gastroenteric infections can be caught via this route.
- Occupational history.
- Animal contact: domestic, farm and wild animals can all be responsible for zoonotic infection.
- Sexual activity: as well as the traditional sexually transmitted diseases, HIV, hepatitis B and rarely hepatitis C can all be transmitted sexually. Some enteric infections are more common among male homosexuals.
- Intravenous drug use: as well as blood-borne viruses, drug injectors are susceptible to a variety of bacterial and fungal infections due to inoculation. Tattooing, body piercing and receipt of blood products (especially outside the UK) are also risk factors for blood-borne viruses.
- Leisure activities: certain pastimes may predispose to water-borne infections or zoonoses.

Clinical examination

A thorough examination covering all systems is required. Skin rashes and lymphadenopathy are common features of infectious diseases, and the ears, eyes, mouth and throat should also be inspected. Infections commonly associated with a rash are listed in Box 4.1. Rectal, vaginal and penile examination is required in sexually transmitted infections.

The fever pattern may occasionally be helpful, e.g. the tertian fever of falciparum malaria, but too much weight should not be placed on the pattern or degree.

Investigations

In some infections such as chickenpox the clinical presentation is so distinctive that no investigations are normally necessary to confirm the diagnosis. Other cases require investigation.

General investigations (to assess health and identify organ(s) involved)

These will vary depending on circumstances:

- **Blood tests**. Routine blood count, ESR and C-reactive protein, biochemical profile, urea and electrolytes are performed in the majority of cases (Box 4.2).
- **Imaging**. X-ray, ultrasound, echocardiography, CT and MR scanning are used to identify and localize infections. Positron emission tomography (PET) and single photon emission tomography (SPECT) have proved useful in localizing infection, especially when combined with CT scanning. However the sensitivity and specificity of these tests in diagnosing infection has yet to be determined, and their use remains limited.

Box 4.1 Infections commonly associated with a rash

Macular/maculopapular

Measles
Rubella
Enteroviruses
Human herpesvirus 6
Epstein–Barr virus
Cytomegalovirus
Human erythrovirus B19
Human immunodeficiency virus (HIV)
Dengue
Typhoid
Secondary syphilis
Rickettsiae spotted fevers

Vesicular

Chickenpox (herpes zoster virus)
Shingles (herpes zoster virus)
Herpes simplex virus
Hand, foot and mouth disease (Coxsackie virus)
Herpangina (Coxsackie virus)

Petechial/haemorrhagic

Meningococcal septicaemia
Any septicaemia with disseminated intravascular coagulation (DIC)
Rickettsiae
Viruses (Table 4.22)

Erythematous

Scarlet fever
Lyme disease (erythema chronicum migrans)
Toxic shock syndrome
Erythrovirus B19

Urticarial

Toxocaria
Strongyloides
Schistosoma
Cutaneous larva migrans

Others

Tick typhus (eschar)
Primary syphilis (chancre)
Anthrax (ulcerating papule)

Box 4.2 General investigations for a patient with suspected infection

Investigation	Possible cause
Full blood count	
Neutrophilia	Bacterial infection
Neutropenia	Viral infection
	Brucellosis
	Typhoid
	Typhus
	Overwhelming sepsis
Lymphocytosis	Viral infection
Lymphopenia	HIV infection (not specific)
Atypical lymphocytes	Infectious mononucleosis
Eosinophilia	Invasive parasitic infection
Thrombocytopenia	Overwhelming sepsis Malaria
ESR or C-reactive protein	Elevated in all
Urea and electrolytes	Potentially deranged in severe illness from any cause
Liver enzymes	
Minor elevation of transferases	Non-specific feature of many infections
	Mild viral hepatitis
High transferases, elevated bilirubin	Viral hepatitis (usually A, B or E)
Coagulation	May be deranged in hepatitis, and in overwhelming infection of any type

Biopsy or aspiration of tissue for microbiological examination may also be facilitated by ultrasound or CT guidance.

- **Radionuclide scanning** after injection of indium- or technetium-labelled white cells (previously harvested from the patient) may occasionally help to localize infection. It is most effective when the peripheral white cell count is raised, and is of particular value in localizing occult abscesses.

Microbiological investigations (to identify causative organism)

Diagnostic services range from simple microscopy to molecular probes. It is often helpful to discuss the clinical problem with a microbiologist to ensure that appropriate tests are performed, and that specimens are collected and transported correctly.

Microscopy and culture

Specimens to be sent for microscopy and culture (Box 4.3).

- *Blood and urine* should routinely be sent for bacterial culture regardless of whether fever is present at the time.
- *Cerebrospinal fluid, sputum and biopsy* specimens are sent if clinically indicated.
- *Special culture techniques* are required for fungi, mycobacteria, and some other bacteria such as *Brucella* spp., and the laboratory must be informed if these are suspected.
- *Faecal culture* for viruses is not helpful in the investigation of gastroenteritis – the viruses responsible for this do not grow in routine tissue culture. Antigen or nucleic acid detection techniques (see below) are more appropriate, especially in the investigation of an outbreak of diarrhoea and vomiting. Protozoa should be considered as a cause of diarrhoea in returning travellers, immunocompromised patients, toddlers, men having sex with men, farm workers, and in any cases of prolonged unexplained diarrhoea. Detection of a specific clostridial toxin is a more reliable test for diarrhoea caused by *Clostridium difficile* than culture of the organism itself. Stool culture is a costly routine test and is often requested unnecessarily.

Immunodiagnostic tests

These can be divided into two types:

- tests that detect viral or bacterial antigen, using a polyvalent antiserum or a monoclonal antibody
- tests that detect an antibody response to infection (serological tests).

These investigations are valuable in the identification of organisms that are difficult to culture, especially viruses and

Box 4.3 Specimens and indications for microscopy and culture

Specimen	Investigation	Indication
Blood	Giemsa stain for malaria	Any symptomatic traveller returning from a malarious area
	Stains for other parasites	Specific tropical infections
	Culture	All suspected bacterial infections
Urine	Microscopy and culture	All suspected bacterial infections
	Tuberculosis (TB) culture	Suspected TB
		Unexplained leucocytes in urine
Faeces	Microscopy ± iodine stain	Suspected protozoal diarrhoea
	Culture	All unexplained diarrhoea
	PCR/antigen detection (not usually necessary to do both)	Suspected viral diarrhoea in children
	Clostridium difficile toxin	Diarrhoea following hospital stay or antibiotic treatment
Throat swabs	Culture	Suspected bacterial tonsillitis and pharyngitis
	PCR	Viral meningitis
	PCR/antigen detection	Viral respiratory infections where urgent diagnosis is considered necessary
Sputum	Microscopy and culture	Unusual chest infections; pneumonia
	Auramine stain/TB culture (liquid culture – see p. 866)	Suspected TB
	Other special stains/cultures	Immunocompromised patients
		Suspected fungal infections
Cerebrospinal fluid	Microscopy and culture	Suspected meningitis
	Auramine stain/TB culture	Suspected TB, meningitis
	Other special stains/cultures	Immunocompromised patients
		Suspected fungal infections
	PCR	Suspected encephalitis or viral or bacterial meningitis
Rash aspirate:		
Petechial	Microscopy and culture	Meningococcal disease
Vesicular	PCR/antigen detection/viral culture	Herpes simplex/zoster

PCR, polymerase chain reaction.

fungi, and can also be helpful when antibiotics have been administered before samples were obtained. However, care is needed in the interpretation of serological tests. Elevated antibody titres on a single occasion (especially of IgG) are rarely diagnostic, and in some infections it may be difficult to distinguish between past and acute infection. Paired serological tests a few weeks apart, or specific assays for IgM antibodies (indicating an acute infection), are more helpful. Numerous serological tests are available: they should only be used in the light of the clinical picture, and not as a general 'trawl' for a diagnosis.

Nucleic acid detection

Specific genes from many pathogenic microorganisms have been cloned and sequenced. Nucleic acid probes can be designed to detect these sequences, identifying pathogen-specific nucleic acid in body fluids or tissue. The utility of this approach has been greatly enhanced by the development of amplification techniques such as the polymerase chain reaction (PCR), which increases the amount of target DNA/RNA in the sample to be tested.

Treatment

Many infections, particularly those caused by viruses, are self-limiting and require no treatment. The mainstay of treatment for most infectious diseases is antimicrobial chemotherapy. The choice of antibiotic should be governed by:

- the clinical state of the patient
- the likely cause of the infection.

Serious infections may require supportive therapy in addition to antibiotics. It is always preferable to have a definite microbial diagnosis before starting treatment, so that an antibiotic with the most appropriate spectrum of activity and site of action can be used. However, some patients are too unwell to wait for results (which in the case of culture may take days). In diseases such as meningitis or septicaemia delay in treatment may be fatal and therapy must be started on an empirical basis. Appropriate samples for culture should be taken before the first dose of antibiotic, and an antibiotic regimen chosen on the basis of the most likely causative organisms. Usually patients are less unwell, and specific therapy can be deferred pending results. Antibiotic therapy is discussed in more detail on page 91.

Special circumstances

Returning travellers. A detailed travel itinerary, including any flight stopovers, should be taken from anyone who is unwell after arriving from another country. Previous travel should also be covered as some infections may be chronic or recurrent. It is necessary to find out not just which countries were visited but also the type of environment: a stay in a remote jungle village carries different health risks from a holiday in an air-conditioned coastal holiday resort. Food and water consumption, bathing and swimming habits, animal and insect contact, and contact with human illness all need to be established. Enquiry should be made about sexual contacts, drug use and medical treatment (especially parenteral) while

abroad. In some parts of the world over 90% of professional sex workers are HIV-positive, and hepatitis B and C are very common in parts of Africa and Asia. In addition to the investigations described in the previous section, special tests may be needed depending on the epidemiological risks and clinical signs, and malaria films are mandatory in anyone who is unwell after being in a malarious area. Some of the more common causes of a febrile illness in returning travellers are listed in Table 4.8.

Table 4.8 Causes of febrile illness in travellers returning from the Tropics and world-wide

WHO advises that fever occurring in a traveller 1 week or more after entering a malaria risk area and up to 3 months after departure is a medical emergency

Developing countries	**Specific geographical areas (see text)**
Malaria	Histoplasmosis
Schistosomiasis	Brucellosis
Dengue	
Tick typhus	**World-wide**
Typhoid	Influenza
Tuberculosis	Pneumonia
Dysentery	URTI
Hepatitis A	UTI
Amoebiasis	Traveller's diarrhoea
	Viral infection

URTI, upper respiratory tract infection; UTI, urinary tract infection.

Immunocompromised patients. Advances in medical treatment over the past three decades have led to a huge increase in the number of patients living with immunodeficiency states. Cancer chemotherapy, the use of immunosuppressive drugs and the world-wide AIDS epidemic have all contributed to this. The presentation may be very atypical in the immunocompromised patient with few, if any, localizing signs or symptoms. Infection can be due to organisms which are not usually pathogenic, including environmental bacteria and fungi. The normal physiological responses to infection (e.g. fever, neutrophilia) may be diminished or absent. The onset of symptoms may be sudden, and the course of the illness fulminant. A high index of suspicion for infections in people who are known to be immunosuppressed is required. These patients should usually be given early and aggressive antibiotic therapy without waiting for the results of investigations. Samples for culture should be sent before starting treatment, but therapy should not be delayed if this proves difficult. The choice of antibiotics should be guided by the likely causative organisms: these are shown in Box 4.4.

Injecting drug users. Parenteral drug use is associated with a variety of local and systemic infections. HIV, HBV and HCV can all be transmitted by sharing injecting equipment. Abscesses and soft tissue infections at the site of injection are common, especially in the groin, and may involve adjacent vascular and bony structures. Systemic infections are also common, most frequently caused by staphylococci and Group A streptococci, but a wide variety of other bacterial and fungal pathogens may be implicated.

Box 4.4 Common causes of infection in immunocompromised patients

Deficiency	**Causes**	**Organisms**
Neutropenia	Chemotherapy	*Escherichia coli*
	Myeloablative therapy	*Klebsiella pneumoniae*
	Immunosuppressant drugs	*Staph. aureus*
		Staph. epidermidis
		Aspergillus spp.
		Candida spp.
Cellular immune defects	HIV infection	Respiratory syncytial virus
	Lymphoma	Cytomegalovirus
	Myeloablative therapy	Epstein–Barr virus
	Congenital syndromes	Herpes simplex and zoster
		Salmonella spp.
		Mycobacterium spp. (esp. *M. avium-intracellulare*)
		Cryptococcus neoformans
		Candida spp.
		Cryptosporidium parvum
		Pneumocystis jiroveci
		Toxoplasma gondii
Humoral immune deficiencies	Congenital syndromes	*Haemophilus influenzae*
	Chronic lymphocytic leukaemia	*Streptococcus pneumoniae*
	Corticosteroids	Enteroviruses
Terminal complement deficiencies (C5–C9)	Congenital syndromes	*Neisseria meningitidis*
		N. gonorrhoeae
Splenectomy	Surgery	*Strep. pneumoniae*
	Trauma	*N. meningitidis*
		H. influenzae
		Malaria

Highly transmissible infections. Relatively few patients with infectious disease present a serious risk to healthcare workers (HCW) and other contacts. However, the appearance of diseases like severe acute respiratory syndrome (SARS), avian influenza, the occasional importation of zoonoses like Lassa fever, and concerns about the bioterrorist use of agents such as smallpox mean that there is still the potential for unexpected outbreaks of life-threatening disease. During the world-wide SARS outbreak in 2003, scrupulous infection control procedures reduced spread of infection. However, in the 'inter-epidemic' period it is difficult to maintain the same level of 'alert'. HCWs should remain vigilant because the early symptoms of many of these diseases are non-specific.

Pyrexia of unknown origin

History, clinical examination and simple investigation will reveal the cause of a fever in most patients. In a small number, however, no diagnosis is apparent despite continuing symptoms. The term pyrexia (or fever) of unknown origin (PUO) is sometimes used to describe this problem. Various definitions have been suggested for PUO: a useful one is 'a fever persisting for >2 weeks, with no clear diagnosis despite intelligent and intensive investigation'. Patients who are known to have HIV or other immunosuppressing conditions are normally excluded from the definition of PUO, as the investigation and management of these patients is different.

Successful diagnosis of the cause of PUO depends on a knowledge of the likely and possible aetiologies. These have been documented in a number of studies, and are summarized in Box 4.5.

A detailed history and examination is essential, taking into account the possible causes, and the examination should be repeated on a regular basis in case new signs appear. Investigation findings to date should be reviewed, obvious omissions amended and abnormalities followed up. Confirm that the patient does have objective evidence of a raised temperature: this may require admission to hospital if the patient is not already under observation. Some people have an exaggerated circadian temperature variation (usually peaking in the evening), which is not pathological.

The range of tests available is discussed above. Obviously investigation is guided by particular abnormalities on examination or initial test results, but in some cases 'blind' investigation is necessary. Some investigations, especially cultures, should be repeated regularly, and serial monitoring of inflammatory markers such as C-reactive protein allows assessment of progress.

Improvements in imaging techniques have diminished the need for invasive investigations in PUO, and scanning has superseded the blind diagnostic laparotomy. Ultrasound, echocardiography, CT, MRI, PET and labelled white cell scanning can all help in establishing a diagnosis if used appropriately: the temptation to scan all patients with PUO from head to toe as a first measure should be avoided. Biopsy of bone marrow (and less frequently liver) may be useful even in the absence of obvious abnormalities, and temporal artery biopsy should be considered in the elderly (see p. 550). Bronchoscopy can be used to obtain samples for microbiological and histological examination if sputum specimens are not adequate. Serological tests have greatly improved the diagnosis of infectious causes of PUO, but these tests should only be ordered and interpreted in the context of the clinical findings and epidemiology.

Box 4.5 Causes of pyrexia of unknown origin

Infection (20–40%)
Pyogenic abscess
Tuberculosis
Infective endocarditis
Toxoplasmosis
Epstein–Barr virus (EBV) infection
Cytomegalovirus (CMV) infection
Primary HIV infection
Brucellosis
Lyme disease

Malignant disease (10–30%)
Lymphoma
Leukaemia
Renal cell carcinoma
Hepatocellular carcinoma

Vasculitides (15–20%)
Adult Still's disease
Rheumatoid arthritis
Systemic lupus erythematosus
Wegener's granulomatosis
Giant cell arteritis
Polymyalgia rheumatica

Miscellaneous (10–25%)
Drug fevers
Thyrotoxicosis
Inflammatory bowel disease
Sarcoidosis
Granulomatous hepatitis
Factitious fever
Familial Mediterranean fever

Undiagnosed (5–25%)

Treatment of a patient with a persistent fever is aimed at the underlying cause, and if possible only symptomatic treatment should be used until a diagnosis is made. Blind antibiotic therapy may make diagnosis of an occult infection more difficult, and empirical steroid therapy may mask an inflammatory response without treating the underlying cause. In a few patients no cause for the fever is found despite many months of investigation and follow-up. In most of these the symptoms do eventually settle spontaneously, and if no definite cause has been identified after 2 years the long-term prognosis is good.

Ask the authors www.studentconsult.com

FURTHER READING

Freedman DO, Weld LH, Kozarsky PE et al. Spectrum of disease and relation to place of exposure among ill returned travellers. *New England Journal of Medicine* 2006; **354**: 119–130.

Mourad O, Palda A, Detsky A. A comprehensive evidence-based approach to fever of unknown origin. *Archives of Internal Medicine* 2003; **163**: 545–551.

ANTIMICROBIAL CHEMOTHERAPY

Principles of use

Antibiotics are among the safest of drugs, especially those used to treat community infections. They have had a major impact on the life-threatening infections and reduce the morbidity associated with surgery and many common infectious diseases. This in turn is, in part, responsible for the overprescribing of these agents which has led to concerns with regard to the increasing incidence of antibiotic resistance.

Empirical 'blind' therapy

Most antibiotic prescribing, especially in the community, is empirical. Even in hospital practice, microbiological documentation of the nature of an infection and the susceptibility of the pathogen is generally not available for a day or two. Initial choice of therapy relies on a clinical diagnosis and, in turn, a presumptive microbiological diagnosis. Such 'blind therapy' is directed at the most likely pathogen(s) responsible for a particular syndrome such as meningitis, urinary tract infection or pneumonia. Examples of 'blind therapy' for these three conditions are ceftriaxone, trimethoprim and amoxicillin + erythromycin, respectively. Initial therapy in the severely ill patient is often broad spectrum in order to cover the range of possible pathogens but should be targeted once microbiological information becomes available.

Combination therapy

Combinations of drugs are occasionally required for reasons other than providing broad-spectrum cover. Tuberculosis is initially treated with three or four agents to avoid resistance emerging. Synergistic inhibition is achieved by using penicillin and gentamicin in enterococcal endocarditis or gentamicin and ceftazidime in life-threatening pseudomonas infection.

In the majority of infections there is no firm evidence that bactericidal drugs (penicillins, cephalosporins, aminoglycosides) are more effective than bacteriostatic drugs (tetracyclines), but it is generally considered necessary to use the former in the treatment of bacterial endocarditis and in patients in whom host defence mechanisms are compromised, particularly in those with neutropenia.

Pharmacokinetic factors

To be successful, sufficient antibiotic must penetrate to the site of the infection. Knowledge of the standard pharmacokinetic considerations of absorption, distribution, metabolism and excretion for the various drugs is required. Difficult sites include the brain, eye and prostate, while loculated abscesses are inaccessible to most agents.

Many mild-to-moderate infections can be treated effectively with oral antibiotics provided that the patient is compliant. Parenteral administration is indicated in the severely ill patient to ensure rapid high blood and tissue concentrations of drug. Some antibiotics can only be administered parenterally, such as the aminoglycosides and extended-spectrum cephalosporins. Parenteral therapy is also required in those unable to swallow or where gastrointestinal absorption is unreliable.

Dose and duration of therapy

This will vary according to the nature, severity and response to therapy.

Prolonged treatment (up to 6 weeks) is necessary for some varieties of infective endocarditis, while pulmonary tuberculosis is treated for at least 6 months. In treating many common infections, improvement occurs within 2–3 days; once the patient is afebrile or the leucocytosis has settled, oral administration should be used for those commenced on parenteral therapy. Treatment for 5–7 days is adequate for most infections. Shorter-course therapy (3 days or less) is appropriate for those with symptomatic uncomplicated bacteriuria (cystitis). Minimizing the duration of therapy lowers the risks of adverse reactions and superinfection with *Candida* spp. or *Clostridium difficile*, as well as the cost of therapy.

Drugs which are concentrated intracellularly, such as erythromycin, quinolones and tetracyclines, are used in treating mycoplasma, brucella and legionella infections.

Renal and hepatic insufficiency

Many drugs require dose reduction in renal failure to avoid toxic accumulation. This applies to the β-lactams and especially the aminoglycosides and vancomycin. Tetracyclines, other than doxycycline and minocycline, should be avoided. In those with hepatic insufficiency, dose reduction is often required for agents such as isoniazid, ketoconazole, clindamycin, interferon and rifampicin.

Therapeutic drug monitoring

To ensure therapeutic yet non-toxic drug concentrations, serum concentrations of drugs such as the aminoglycosides and vancomycin are monitored, especially in those with impaired or changing renal function. Peak (1 hour post-dose) and trough (pre-dose) serum samples are assayed. However, with the increasing use of once-daily aminoglycoside dosage regimens, random but timed serum assays are being adopted.

Antibiotic chemoprophylaxis

The value of antibiotic chemoprophylaxis has been questioned as there are few controlled trials to prove efficacy (see p. 773). The evidence for chemoprophylaxis against infective endocarditis (IE) is an example. The new guidelines recognize that procedures can cause bacteraemia but not infective endocarditis. Even patients at 'high risk', e.g. previous IE prosthetic heart valves and surgical shunts, do not require prophylaxis (Table 4.9). However, there are a number of indications for which the prophylactic use of antibiotics is still advised. These include surgical procedures where the risk of infection is high (colon surgery) or the consequences of infection are serious (post-splenectomy sepsis). The choice of agent(s) is determined by the likely infectious risk and the established efficacy and safety of the regimen.

Mechanisms of action and resistance to antimicrobial agents

Antibiotics act at different sites of the bacterium.

Resistance to an antibiotic can be the result of:

- impaired or altered permeability of the bacterial cell envelope, e.g. penicillins in Gram-negative bacteria
- active expulsion from the cell by membrane efflux systems
- alteration of the target site (e.g. single point mutations in *E. coli* or a penicillin-binding protein in *Strep. pneumoniae* leading to acquired resistance – see below)
- specific enzymes which inactivate the drug before or after cell entry (e.g. β-lactamases)
- development of a novel metabolic bypass pathway.

The development or acquisition of resistance to an antibiotic by bacteria invariably involves either a mutation at a single point in a gene or transfer of genetic material from another organism (Fig. 4.4).

Larger fragments of DNA may be introduced into a bacterium either by transfer of 'naked' DNA or via a bacteriophage (a virus) DNA vector. Both the former (transformation)

Table 4.9 Antibiotic chemoprophylaxis

(a) General

Clinical problem	Aim	Drug regimen*
Splenectomy/spleen malfunction	To prevent serious pneumococcal sepsis	Phenoxymethylpenicillin 500 mg 12-hourly
Rheumatic fever	To prevent recurrence and further cardiac damage	Phenoxymethylpenicillin 250 mg × 2 daily or sulfadiazine 1 g if allergic to penicillin
Meningitis:		
Due to meningococci	To prevent infection in close contacts	Adults: rifampicin 600 mg twice-daily for 2 days (Children < 1 year: 5 mg/kg; >1 year: 10 mg/kg) Alternatives (single dose) ciprofloxacin 500 mg (p.o.) or ceftriaxone 250 mg (i.m.)
Due to *H. influenzae* type b	To reduce nasopharyngeal carriage and prevent infection in close contacts	Adults: rifampicin 600 mg daily for 4 days (Children: < 3 months 10 mg/kg; > 3 months 20 mg/kg)
Tuberculosis	To prevent infection in exposed (close contacts) tuberculin-negative individuals, infants of infected mothers and immunosuppressed patients	Oral isoniazid 300 mg daily for 6 months (Children: 5–10 mg/kg daily)

(b) Endocarditis (NICE guidelines for adults and children undergoing interventional procedures March 2008)

- Antibacterial prophylaxis and chlorhexidine mouthwash are not recommended for the prevention of endocarditis in patients undergoing dental procedures.
- Antibacterial prophylaxis is not recommended for the prevention of endocarditis in patients undergoing procedures of the:
  - Upper and lower respiratory tract (including ear, nose and throat procedures and bronchoscopy)
  - Genitourinary tract (including urological, gynaecological and obstetric procedures)
  - Upper and lower gastrointestinal tract.

Any infection in patients at risk of endocarditis should be investigated promptly and treated appropriately to reduce the risk of endocarditis.

- If patients at risk of endocarditis are undergoing a gastrointestinal or genitourinary tract procedure at a site where infection is suspected, they should receive appropriate antibacterial therapy that includes cover against organisms that cause endocarditis.
- Patients at risk of endocarditis should be:
  - Advised to maintain good oral hygiene
  - Told how to recognize signs of infective endocarditis and advised when to seek expert advice.

*Unless stated, doses are those recommended in adults.
For surgical procedure – see individual procedures in text.
Adapted from Joint Formulary Committee *British National Formulary*, 55th edn. London: BMJ Group and RPS Publishing, 2008.

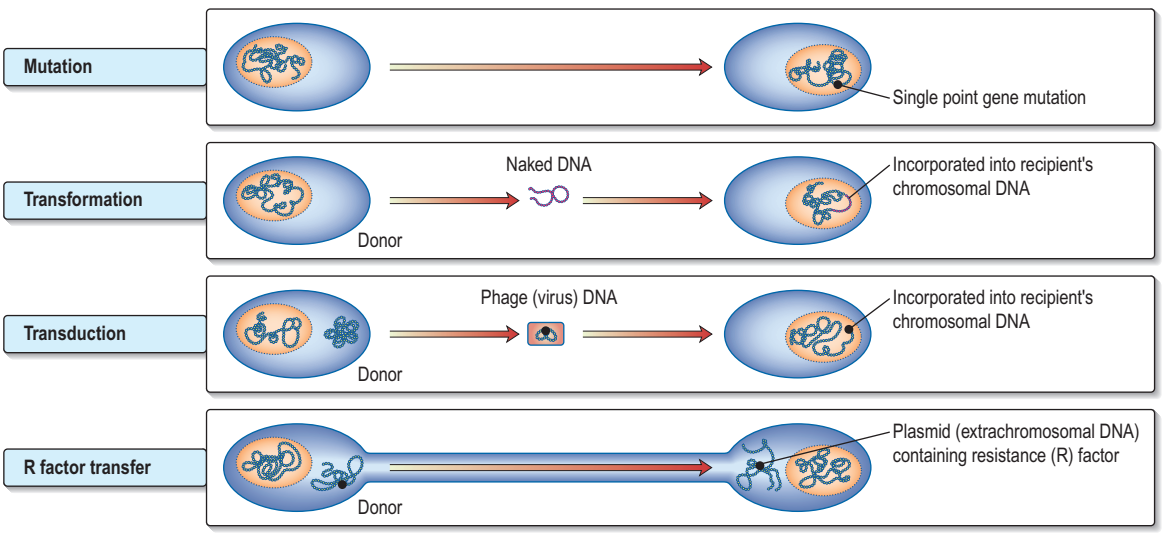


Fig. 4.4 **Some mechanisms for the development of resistance to antimicrobial drugs.** These involve either a single point mutation or transfer of genetic material from another organism (transformation, transduction or R factor transfer).

Table 4.10 Some bacteria that have developed resistance to common antibiotics

Pathogen	Previously fully sensitive to
Enterobacteria	Amoxicillin, trimethoprim, ciprofloxacin, gentamicin, glycopeptide (GRE), vancomycin (VRE)
Haemophilus influenzae	Amoxicillin, chloramphenicol
Neisseria gonorrhoeae	Penicillin, ciprofloxacin
Pseudomonas aeruginosa	Gentamicin
Salmonella spp.	Amoxicillin, sulphonamides, ciprofloxacin
Shigella spp.	Amoxicillin, trimethoprim, tetracycline
Staphylococcus aureus	Penicillin, meticillin (MRSA), vancomycin (VRSA), ciprofloxacin
Streptococcus pneumoniae	Penicillin, erythromycin, cefotaxime
Streptococcus pyogenes	Erythromycin, tetracycline
Vibrio cholerae	Quinolones, azithromycin

Table 4.11 Classification of penicillins

Benzylpenicillin and its long-acting parenteral relatives
Benzylpenicillin
Benethamine penicillin*
Benzathine penicillin*
Clemizole penicillin*
Procaine benzylpenicillin (procaine penicillin)

Oral alternatives to benzylpenicillin
Azidocillin*
Phenoxymethylpenicillin (penicillin V)

β-Lactamase-stable penicillins
Flucloxacillin
Temocillin
Oxacillin*
Meticillin*
Nafcillin*

Extended-spectrum penicillins
Ampicillin, pivampicillin,* talampicillin*
Amoxicillin
Co-amoxiclav, pivmecillinam
Co-fluampicil

Penicillins active against *Pseudomonas*
Azlocillin,* mezlocillin,* piperacillin, ticarcillin

*Not available in the UK.

and the latter (transduction) are dependent on integration of this new DNA into the recipient chromosomal DNA. This requires a high degree of homology between the donor and recipient chromosomal DNA.

Finally, antibiotic resistance can be transferred from one bacterium to another by conjugation, when extrachromosomal DNA (a plasmid) containing the resistance factor (R factor) is passed from one cell into another during direct contact. Transfer of such R factor plasmids can occur between unrelated bacterial strains and involve large amounts of DNA and often codes for multiple antibiotic resistance, e.g. as for the quinolones.

Transformation is probably the least clinically relevant mechanism, whereas transduction and R factor transfer are usually responsible for the sudden emergence of multiple antibiotic resistance in a single bacterium. Increasing resistance to many antibiotics has developed (Table 4.10).

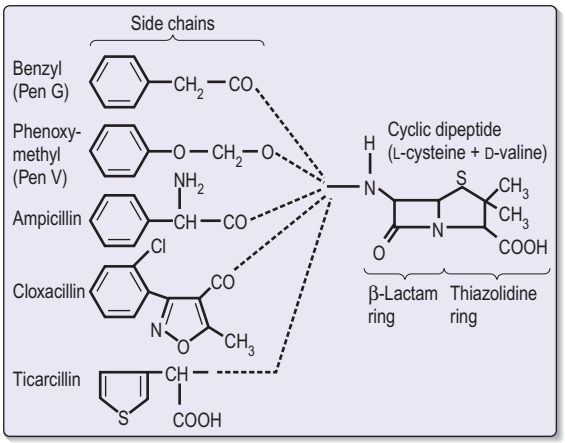


Fig. 4.5 **The structure of penicillins.**

FURTHER READING

Diekema DJ, Boots Miller BJ et al. Antimicrobial resistance trends and outbreak frequency in United States hospitals. *Clinical Infectious Diseases* 2004; **38**: 78–85.

McDonald LC. Trends in antimicrobial resistance in healthcare associated pathogens and effect on treatment. *Clinical Infectious Diseases* 2006; **42**: S65–S71.

National Institute for Health and Clinical Excellence (NICE) www.nice.org.uk.

Robicsek A, Jacoby GA, Hooper DC. The world-wide emergence of plasmid-mediated quinolone resistance. *Lancet Infectious Diseases* 2006; **6**: 629–640.

ANTIBACTERIAL DRUGS

β-Lactams (penicillins, cephalosporins and monobactams)

Penicillins (Table 4.11)

Structure. The β-lactams share a common ring structure (Fig. 4.5). Changes to the side-chain of benzylpenicillin (penicillin G) render the phenoxymethyl derivative (penicillin V) acid resistant and allow it to be orally absorbed. The presence of an amino group in the phenyl radical of benzylpenicillin increases its antimicrobial spectrum to include many Gram-negative and Gram-positive organisms. More extensive modification of the side-chain (e.g. as in flucloxacillin) renders the drug insensitive to bacterial penicillinase. This is useful in treating infections caused by penicillinase (β-lactamase)-producing staphylococci.

Mechanisms of action. β-lactams block bacterial cell wall mucopeptide formation by binding to and inactivating specific penicillin-binding proteins (PBPs), which are peptidases involved in the final stages of cell wall assembly and division. Meticillin-resistant *Staph. aureus* (MRSA) (see p. 124) produce a low-affinity PBP which retains its peptidase activity even in the presence of high concentrations of meticillin.

Indications for use. Benzylpenicillin can only be given parenterally and is often the drug of choice for serious infections,

notably infective endocarditis, meningococcal, streptococcal, gas gangrene, actinomycosis, anthrax and spirochaetal infections (syphilis, yaws).

Phenoxymethylpenicillin (penicillin V) is an oral preparation that is used chiefly to treat streptococcal pharyngitis and as prophylaxis against rheumatic fever.

Flucloxacillin is used in infections caused by β-lactamase (penicillinase)-producing staphylococci.

Ampicillin is susceptible to β-lactamase, but its antimicrobial activity includes streptococci, pneumococci and enterococci as well as Gram-negative organisms such as *Salmonella* spp., *Shigella* spp., *E. coli*, *H. influenzae* and *Proteus* spp. Drug resistance has, however, eroded its efficacy against these Gram-negatives. It is widely used in the treatment of respiratory tract infections. Amoxicillin has similar activity to ampicillin, but is better absorbed when given by mouth.

The extended-spectrum penicillin, ticarcillin, is active against pseudomonas infections, as is the acylureidopenicillin piperacillin in combination with sulbactam.

Clavulanic acid is a powerful inhibitor of many bacterial β-lactamases and when given in combination with an otherwise effective agent such as amoxicillin (co-amoxiclav) or ticarcillin can broaden the spectrum of activity of the drug. Sulbactam acts similarly and is available combined with ampicillin, while tazobactam in combination with piperacillin is effective in appendicitis, peritonitis, pelvic inflammatory disease and complicated skin infections. The penicillin β-lactamase combinations are also active against β-lactamase-producing staphylococci.

Pivmecillinam has significant activity against Gram-negative bacteria including *E. coli*, *Klebsiella*, *Enterobacter* and *Salmonellae* but no activity against pseudomonas.

Temocillin is active against Gram-negative bacteria, including β-lactamase producers. It is not active against pseudomonas or *Acinetobacter* spp.

Interactions. Penicillins inactivate aminoglycosides when mixed in the same solution.

Toxicity. Generally, the penicillins are very safe. Hypersensitivity (skin rash (common), urticaria, anaphylaxis), encephalopathy and tubulointerstitial nephritis can occur. Ampicillin also produces a hypersensitivity rash in approximately 90% of patients with infectious mononucleosis who receive this drug. Co-amoxiclav causes a cholestatic jaundice six times more frequently than amoxicillin, as does flucloxacillin.

Cephalosporins and cephamycins (Fig. 4.6)

The cephalosporins have an advantage over the penicillins in that they are resistant to staphylococcal penicillinases (but are still inactive against meticillin-resistant staphylococci) and have a broader range of activity that includes both Gram-negative and Gram-positive organisms, but excludes enterococci and anaerobic bacteria. Ceftazidime and cefpirome are active against *Pseudomonas aeruginosa*. Cefoxitin is a cephamycin with activity against anaerobic bacteria but is no longer available in the UK.

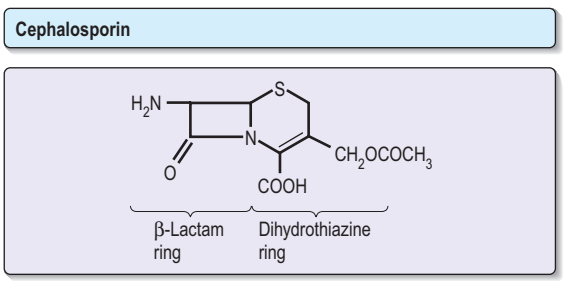


Fig. 4.6 **The structure of a cephalosporin.**

Indications for use (Table 4.12). These potent broad-spectrum antibiotics are useful for the treatment of serious systemic infections, particularly when the precise nature of the infection is unknown. They are commonly used for serious sepsis in postoperative and immunocompromised patients, particularly during cytotoxic chemotherapy of leukaemia and other malignancies. They are used in pneumonia, meningitis, peritonitis and urinary tract infections.

Interactions. There are relatively few interactions.

Toxicity. The toxicity is similar to that of the penicillins but is less common. Some 10% of patients are allergic to both groups of drugs. The early cephalosporins caused proximal tubule damage, although the newer derivatives have fewer nephrotoxic effects.

Monobactams

Aztreonam is currently the only member of this class available. It is a synthetic β-lactam and, unlike the penicillins and cephalosporins, has no ring other than the β-lactam, hence its description as a monobactam.

Its mechanism of action is by inhibition of bacterial cell wall synthesis. It is resistant to most β-lactamases and does not induce β-lactamase production.

Indications for use. Aztreonam's spectrum of activity is limited to aerobic Gram-negative bacilli. With the exception of urinary tract infections, aztreonam should be used in combination with metronidazole (for anaerobes) and an agent active against Gram-positive cocci (a penicillin or erythromycin). It is a useful alternative to aminoglycosides in combination therapy, largely for the treatment of intra-abdominal sepsis.

Toxicity. As for the β-lactam antibiotics.

Carbapenems

The carbapenems are semisynthetic β-lactams and include imipenem, meropenem and ertapenem. They are currently the most broad spectrum of antibiotics, being active against the majority of Gram-positive and Gram-negative as well as anaerobic bacterial pathogens. Ertapenem, unlike the others, is not active against *Pseudomonas* or *Acinetobacter* spp. They differ in their dosage and frequency of administration. Imipenem is partially inactivated in the kidney by enzymatic inactivation and is therefore administered in combination with cilastatin.

Indications for use. They are used for serious nosocomial infections when multiple-resistant Gram-negative bacilli or mixed aerobe and anaerobe infections are suspected.

Toxicity. This is similar to that of β-lactam antibiotics. Nausea, vomiting and diarrhoea occur in less than 5% of

Table 4.12 Some examples of cephalosporins

	Activity	Use
First generation	Gram-positive cocci and Gram-negative organisms	Urinary tract infections
Cefalexin (oral)		Penicillin allergy
Cefradine (oral)		
Cefadroxil (oral)		
Second generation	Extended spectrum	Prophylaxis and treatment of Gram-negative infections and mixed aerobic–anaerobic infections
Cefuroxime Cefamandole*	More effective than first generation against *E. coli*, *Klebsiella* spp. and *Proteus mirabilis*, but less effective against Gram-positive organisms	
Cefoxitin*	Includes *Bacteroides fragilis*	Peritonitis
Cefaclor (oral)		
Cefuroxime (oral)		
Cefprozil (oral)*		
Third generation Cefotaxime Ceftazidime	Broad-spectrum More potent against aerobic Gram-negative bacteria than first or second generation	Especially severe infection with Enterobacteriaceae, *Pseudomonas aeruginosa* (ceftazidime, cefpirome) and *Neisseria gonorrhoeae*, *N. meningitidis*, Lyme disease (ceftriaxone)
Cefpirome*		Urinary tract infections and infections with neutropenia
Ceftriaxone		Once daily. Septicaemia, pneumonia, meningitis
Cefpodoxime (oral)		
Cefixime (oral)		
Fourth generation	Aerobic Gram-negative bacteria including *P. aeruginosa*	Febrile, neutropenic patients
Cefepime*		

*Unavailable in the UK.

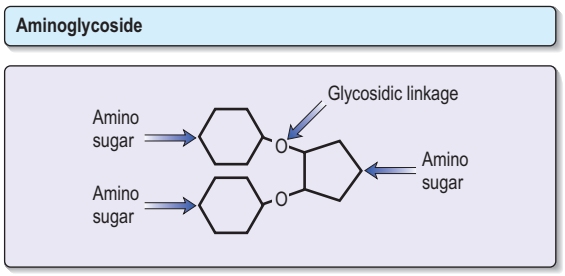


Fig. 4.7 **The structure of an aminoglycoside.**

cases. Imipenem may cause seizures and should not be used to treat meningitis; meropenem is safe for this indication.

Aminoglycosides

Structure. These antibiotics are polycationic compounds of amino sugars (Fig. 4.7).

Mechanism of action. Aminoglycosides interrupt bacterial protein synthesis by inhibiting ribosomal function (messenger and transfer RNA).

Indications for use. Streptomycin is bactericidal and is rarely used except for the treatment of tuberculosis or the plague. Occasional indications include endocarditis (with penicillin/ampicillin). Neomycin is used only for the topical treatment of eye and skin infections. Even though it is poorly absorbed, prolonged oral administration can produce ototoxicity.

Gentamicin and tobramycin are given parenterally. They are highly effective against many Gram-negative organisms including *Pseudomonas* spp. They are synergistic with a penicillin against *Enterococcus* spp. Netilmicin and amikacin have a similar spectrum but are more resistant to the aminoglycoside-inactivating enzymes (phosphorylating, adenylating or acetylating) produced by some bacteria. Their use should be restricted to gentamicin-resistant organisms.

Interactions. Enhanced nephrotoxicity occurs with other nephrotoxic drugs, ototoxicity with some diuretics, and neuromuscular blockade with curariform drugs.

Toxicity. This is dose-related. Aminoglycosides are nephrotoxic and ototoxic (vestibular and auditory), particularly in the elderly. Therapeutic drug monitoring is necessary in ensuring therapeutic and non-toxic drug concentrations.

Tetracyclines

Structure. These are bacteriostatic drugs possessing a four-ring hydronaphthacene nucleus (Fig. 4.8). Included among the tetracyclines are tetracycline, oxytetracycline, demeclocycline, lymecycline, doxycycline and minocycline. Tigecycline is a new injectable glycylcycline which is structurally related to the tetracyclines.

Mechanism of action. Tetracyclines inhibit bacterial protein synthesis by interrupting ribosomal function (transfer RNA).

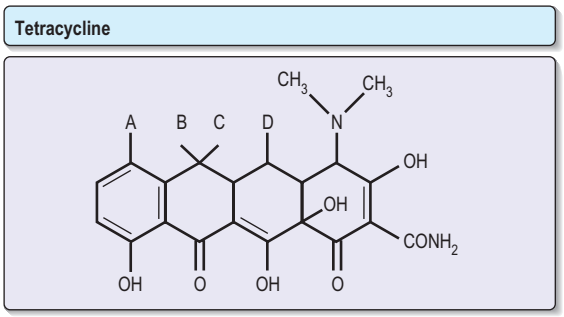


Fig. 4.8 **The structure of a tetracycline.** Substitution of CH_3, OH or H at positions A to D produces variants of tetracycline.

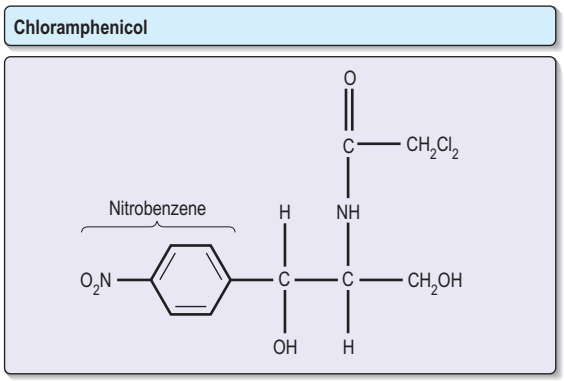


Fig. 4.9 **The structure of chloramphenicol.**

Indications for use. Tetracyclines are active against Gram-positive and Gram-negative bacteria but their use is now limited, partly owing to increasing bacterial resistance. Resistance is now common with *Strep. pneumoniae*. A tetracycline is used for the treatment of acne and rosacea. Tetracyclines are also active against *V. cholerae*, *Rickettsia* spp., *Mycoplasma* spp., *Coxiella burnetii*, *Chlamydia* spp. and *Brucella* spp. Tigecycline is active against many organisms resistant to tetracycline and other antibiotics. Most notably, these include vancomycin-resistant enterococci, MRSA and multidrug-resistant Gram-negative bacilli, including *Acinetobacter* spp. Its indications include complicated skin and soft tissue infections and intra-abdominal sepsis.

Interactions. The efficacy of tetracyclines is reduced by antacids and oral iron-replacement therapy.

Toxicity. Tetracyclines are generally safe drugs, but they may enhance established or incipient renal failure, although doxycycline is safer than others in this group. They cause brown discoloration of growing teeth, and thus these drugs are not given to children or pregnant women. Photosensitivity can occur. Nausea and vomiting are the most frequent adverse effects of tigecycline.

Macrolides

Erythromycin

Structure. Erythromycin consists of a lactone ring with two sugar side-chains, one of which is an aminosugar.

Mechanism of action. Erythromycin inhibits protein synthesis by interrupting ribosomal function.

Indications for use. Erythromycin has a similar (but not identical) antibacterial spectrum to penicillin and is useful in individuals with penicillin allergy especially in the management of bacterial respiratory and skin and soft tissue infections. It can be given orally or parenterally. It is the preferred agent in the treatment of pneumonias caused by *Legionella* spp. and *Mycoplasma* spp. It is also effective in the treatment of infections due to *Bordetella pertussis* (whooping cough), *Campylobacter* spp. and *Chlamydia* spp.

Other macrolides

These include clarithromycin, azithromycin and telithromycin. They have a broad spectrum of activity that includes selective Gram-negative organisms, mycobacteria and *Toxoplasma gondii*. Compared with erythromycin, they have superior pharmacokinetic properties with enhanced tissue and intracellular penetration and longer half-life that allows once or twice daily dosage. Clarithromycin is used in the treatment of bacterial respiratory infections and also as a component of triple therapy regimens (usually with a proton pump inhibitor and amoxicillin) for the eradication of *Helicobacter pylori*. Azithromycin has similar indications and is also used for trachoma and cholera (see p. 143).

Interactions. Erythromycin and other macrolides interact with theophyllines, carbamazepine, digoxin and ciclosporin, occasionally necessitating dose adjustment of these agents.

Toxicity. Diarrhoea, vomiting and abdominal pain are the main side-effects of erythromycin (less with clarithromycin and azithromycin) and are, in part, a consequence of the intestinal prokinetic properties of the macrolides. Macrolides may also rarely produce cholestatic jaundice after prolonged treatment. QT_c prolongation is a recognized cardiac effect of the macrolides. This may have serious consequences if the syndrome of 'torsades de pointes' is induced.

Chloramphenicol

Structure. Chloramphenicol is the only naturally occurring antibiotic containing nitrobenzene (Fig. 4.9). This structure probably accounts for its toxicity in humans and for its activity against bacteria.

Mechanism of action. Chloramphenicol competes with messenger RNA for ribosomal binding. It also inhibits peptidyl transferase.

Indications for use. Chloramphenicol is rarely used in developed countries. In developing countries it has been invaluable in the treatment of meningitis and severe infections caused by *Salmonella typhi* and *S. paratyphi* (enteric fevers). *H. influenzae* (meningitis and acute epiglottitis) which are still prevalent in countries where Hib vaccination has not been introduced are also susceptible to chloramphenicol. It is also active against *Yersinia pestis* (plague) and is used topically for the treatment of purulent conjunctivitis. Drug resistance is currently eroding the efficacy of chloramphenicol.

Interactions. Chloramphenicol enhances the activity of anticoagulants, phenytoin and some oral hypoglycaemic agents.

Toxicity. Severe irreversible bone marrow suppression is rare but nevertheless now restricts the usage of this drug to only the severely ill patient. Chloramphenicol should not be given to premature infants or neonates because of their inability to conjugate and excrete this drug; high blood levels lead to circulatory collapse and the often fatal 'grey baby syndrome'.

Fusidic acid

Structure. Fusidic acid has a structure resembling that of bile salts (see p. 322).

Mechanism of action. It is a potent inhibitor of bacterial protein synthesis. Its entry into cells is facilitated by the detergent properties inherent in its structure.

Indications for use. Fusidic acid is mainly used for penicillinase-producing *Staph. aureus* infections such as osteomyelitis (it is well concentrated in bone) or endocarditis, and for other staphylococcal infections accompanied by septicaemia. The drug is well absorbed orally but is relatively expensive and should be used in combination with another staphylococcal agent to prevent resistance emerging.

Resistance. Resistance may occur rapidly and is the reason why fusidic acid is given in combination with another antibiotic.

Toxicity. Fusidic acid may occasionally be hepatotoxic but is generally a safe drug and if necessary can be given during pregnancy.

Sulphonamides and trimethoprim

Structure. The sulphonamides are all derivatives of the prototype sulphanilamide. Trimethoprim is a 2,4-diaminopyrimidine.

Mechanism of action. Sulphonamides block thymidine and purine synthesis by inhibiting microbial folic acid synthesis. Trimethoprim prevents the reduction of dihydrofolate to tetrahydrofolate (see Fig. 8.12).

Indications for use. Sulfamethoxazole is mainly used in combination with trimethoprim (as co-trimoxazole). Its use is now largely restricted to the treatment and prevention of *Pneumocystis jiroveci* (formerly *P. carinii*) infection and listeriosis in developed countries, although it is still in widespread use in developing countries. It may also be used for toxoplasmosis and nocardiosis and as a second-line agent in acute exacerbations of chronic bronchitis and in urinary tract infections. Trimethoprim alone is often used for urinary tract infections and acute-on-chronic bronchitis, as the side-effects of co-trimoxazole are most commonly due to the sulphonamide component. Sulfapyridine in combination with 5-aminosalicylic acid (i.e. sulfasalazine) is now less widely used in inflammatory bowel disease.

Resistance. Resistance to sulphonamides is often plasmid-mediated and results from the production of sulphonamide-resistant dihydropteroate synthase from altered bacterial cell permeability to these agents.

Interactions. Sulphonamides potentiate oral anticoagulants and some hypoglycaemic agents.

Toxicity. Sulphonamides cause a variety of skin eruptions, including toxic epidermal necrolysis, the Stevens–Johnson syndrome, thrombocytopenia, folate deficiency and megaloblastic anaemia with prolonged usage. It can provoke haemolysis in individuals with glucose-6-phosphate dehydrogenase deficiency and therefore should not be used in such people. Co-trimoxazole should also be avoided in the elderly if possible, as deaths have been recorded, probably owing to the sulphonamide component.

Quinolones

The quinolone antibiotics, such as ciprofloxacin, norfloxacin, ofloxacin and levofloxacin, are useful oral broad-spectrum antibiotics, related structurally to nalidixic acid. The latter achieves only low serum concentrations after oral administration and its use is limited to the urinary tract where it is concentrated. Newer quinolones, including moxifloxacin, gemifloxacin and gatifloxacin, have greater activity against Gram-positive pathogens. The structure is shown in Figure 4.10.

Mechanism of action. The quinolone group of bactericidal drugs inhibit bacterial DNA synthesis by inhibiting topoisomerase IV and DNA gyrase, the enzyme responsible for maintaining the superhelical twists in DNA.

Indications for use. The extended-spectrum quinolones such as ciprofloxacin have activity against Gram-negative bacteria, including some *Pseudomonas aeruginosa*, and some Gram-positive bacteria (e.g. anthrax, p. 142). They are useful in Gram-negative septicaemia, skin and bone infections, urinary and respiratory tract infections, meningococcal carriage, in some sexually transmitted diseases such as gonorrhoea and non-specific urethritis due to *Chlamydia trachomatis*, and in severe cases of travellers' diarrhoea (see p. 131). The newer oral quinolones provide an alternative to β-lactams in the treatment of community-acquired lower respiratory tract infections.

Resistance. In many countries 30–40% of *E. coli* are resistant. Resistance is also a growing problem among *Salmonella*, *Vibrio cholerae* and *Staphylococcus aureus*, including MRSA strains.

Interactions. Ciprofloxacin can induce toxic concentrations of theophylline.

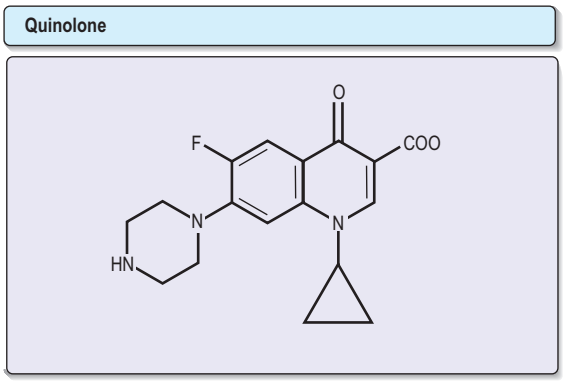


Fig. 4.10 **The structure of a quinolone (ciprofloxacin).**

Toxicity. Gastrointestinal disturbances, photosensitive rashes and occasional neurotoxicity can occur. Avoid in childhood and pregnancy and in patients taking corticosteroids. Tendon damage, including rupture can occur within 48 hours of use and the drug should be stopped immediately; it should not be used in patients with tendonitis. MRSA and *Clostridium difficile* infection in hospitals have been linked to high prescribing rates of quinolones.

Oxazolidinones

Structure. The oxazolidinones are a novel class of antibacterial agents of which linezolid (Fig. 4.11) is the first to become available.

Mechanism of action. The oxazolidinones inhibit protein synthesis by binding to the bacterial 23S ribosomal RNA of the 59S subunit, thereby preventing the formation of a functional 70S complex essential to bacterial translation.

Indications for use. Linezolid is active against a variety of Gram-positive pathogens including vancomycin-resistant *Enterococcus faecium* (unfortunately resistant organisms have already been reported), meticillin-resistant *Staph. aureus* and penicillin-resistant *Strep. pneumoniae*. It is also active against group A and group B streptococci. Clinical experience has demonstrated efficacy in a variety of hospitalized patients with severe to life-threatening infections, such as bacteraemia, hospital-acquired pneumonia and skin and soft tissue infections. It can be given both intravenously and by mouth.

Interactions. Linezolid interacts reversibly as a non-selective inhibitor of monoamine oxidase, and has the potential for interacting with serotoninergic and adrenergic agents.

Toxicity. Side-effects include gastrointestinal disturbances, headache, rash, hypertension, reversible thrombocytopenia and occasional reports of optic neuropathy. Safety has not yet been shown in pregnancy.

Nitroimidazoles

Structure. These agents are active against anaerobic bacteria and some pathogenic protozoa. The most widely used drug is metronidazole (Fig. 4.12). Others include tinidazole and nimorazole.

Mechanism of action. After reduction of their nitro group to a nitrosohydroxyl amino group by microbial enzymes, nitroimidazoles cause strand breaks in microbial DNA.

Indications for use. Metronidazole plays a major role in the treatment of anaerobic bacterial infections, particularly those due to *Bacteroides* spp. It is also used prophylactically in colonic surgery. It may be given orally, by suppository (well absorbed and cheap) or intravenously (very expensive). It is also the treatment of choice for *Clostridium difficile*-associated colitis, amoebiasis, giardiasis and infection with *Trichomonas vaginalis*.

Interactions. Nitroimidazoles can produce a disulfiram-like reaction with ethanol and enhance the anticoagulant effect of warfarin.

Toxicity. Nitroimidazoles are tumorigenic in animals and mutagenic for bacteria, although carcinogenicity has not been described in humans. They cause a metallic taste, and polyneuropathy with prolonged use. They should be avoided in pregnancy.

Glycopeptides

The glycopeptides are antibiotics active against Gram-positive bacteria and act by inhibiting cell wall synthesis.

Vancomycin

Vancomycin is given intravenously for meticillin-resistant *Staph. aureus* and other multiresistant Gram-positive organisms. It is also used for treatment and prophylaxis against Gram-positive infections in penicillin-allergic patients. It is used for *Strep. pneumoniae* meningitis when caused by penicillin-resistant strains. By mouth it is an alternative to metronidazole for *Clostridium difficile*-associated colitis. Glyopeptide-resistant enterococci (GRE) are now seen, as well as vancomycin-resistant *Staphylococcus* (see p. 124).

Toxicity. Vancomycin can cause ototoxicity and nephrotoxicity and thus serum levels should be monitored. Care must be taken to avoid extravasation at the injection site as this causes necrosis and thrombophlebitis. Too rapid infusion can produce symptomatic release of histamine (red man syndrome).

Teicoplanin

This glycopeptide antibiotic is less nephrotoxic than vancomycin. It has more favourable pharmacokinetic properties, allowing once-daily dosage.

Other antibiotics

Clindamycin is not widely used because of its toxic side-effect, antibiotic-associated colitis (pseudomembranous colitis). It is active against Gram-positive cocci including some penicillin-resistant staphylococci. It is increasingly used to treat severe streptococcal or staphylococcal cellulitis. It has the added effect of inhibiting staphylococcal TSST-1 and alpha toxin production. It is also active against anaerobes, especially bacteroides. It is well concentrated in bone and used for osteomyelitis.

Quinupristin and dalfopristin. A combination of these streptogramin antibiotics is used for Gram-positive bacteria which have failed to respond to other antibacterials.

Daptomycin is a lipopeptide with a similar spectrum to vancomycin and is given by i.v. infusion. It is particularly used

Linezolid

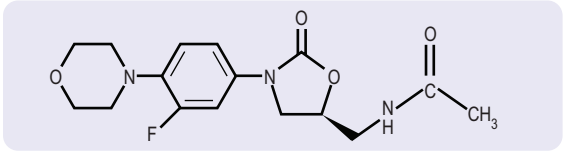


Fig. 4.11 **The structure of linezolid, an oxazolidinone.**

Metronidazole

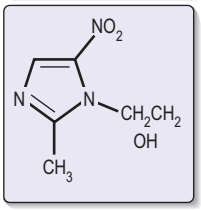


Fig. 4.12
The structure of metronidazole, a nitroimidazole.

for skin and soft tissue infections due to community-associated MRSA. Muscle toxicity occurs.

ANTITUBERCULOSIS DRUGS

These are described on page 866. Rifampicin is also used in other infections apart from tuberculosis.

ANTIFUNGAL DRUGS (Table 4.13)

Polyenes

Polyenes react with the sterols in fungal membranes, increasing permeability and thus damaging the organism. The most potent is amphotericin, which is used intravenously in severe systemic fungal infections. Nephrotoxicity is a major problem and dosage levels must take background renal function into account. Liposomal amphotericin is less toxic but very expensive. Nystatin is not absorbed through mucous membranes and is therefore useful for the treatment of oral and enteric candidiasis and for vaginal infection. It can only be given orally or as pessaries.

Azoles

Imidazoles such as ketoconazole, miconazole and clotrimazole are broad-spectrum antifungal drugs. They are predominantly fungistatic and act by inhibiting fungal sterol synthesis, resulting in damage to the cell wall. Ketoconazole is active orally but can produce liver damage. It is effective in candidiasis and deep mycoses including histoplasmosis and blastomycosis but not in aspergillosis and cryptococcosis.

Clotrimazole and miconazole are used topically for the treatment of ringworm and cutaneous and genital candidiasis. Econazole is used for the topical treatment of cutaneous and vaginal candidiasis and dermatophyte infections while tioconazole is indicated for fungal nail infections.

Triazoles. These include fluconazole, voriconazole and itraconazole. Fluconazole is noted for its ability to enter CSF and is used for candidiasis and for the treatment of central nervous system (CNS) infection with *Cryptococcus neoformans*. Itraconazole fails to penetrate CSF. It is the agent of choice for non-life-threatening blastomycosis and histoplasmosis. It is also moderately effective in invasive aspergillosis. Toxicity is mild. The problem associated with poor absorption of the capsules in the absence of food has been overcome with the liquid formulation. Voriconazole has broad-spectrum activity that includes *Candida*, *Cryptococcus* and *Aspergillus* spp. and other filamentous fungi. It is available for oral and intravenous use. Adverse effects include rash, visual disturbance and abnormalities of liver enzymes. It is indicated for invasive aspergillosis and severe candida infections unresponsive to amphotericin and fluconazole respectively.

Table 4.13 Antifungal agents

Polyenes	**Allylamines**
Amphotericin, nystatin	Terbinafine
Echinocandins	**Other antifungals**
Caspofungin	Amorolfine (topical only)
Anidulafungin	5-Flucytosine
Mycofungin	Griseofulvin
Azoles	
Miconazole, ketoconazole, fluconazole, itraconazole, voriconazole, posaconazole	
Topical clotrimazole, sulconazole, tolnaftate, econazole, tioconazole	

Allylamines

Terbinafine has broad-spectrum antifungal and also anti-inflammatory activity. It is well absorbed orally and accumulates in keratin. It is useful for the treatment of superficial mycoses such as dermatophyte infections, onychomycosis and cutaneous candidiasis. A topical formulation is also available to treat fungal skin infections.

Echinocandins

These act by inhibiting the cell wall polysaccharide, glucan. Caspofungin is active against *Candida* spp. and *Aspergillus* spp. and is indicated for invasive serious aspergillosis unresponsive to other drugs. It is administered intravenously. Other echinocandins include mycofungin and anidulafungin approved for the treatment of disseminated candidiasis.

Other antifungals

Flucytosine. The fluorinated pyridine derivative, flucytosine, is used in combination with amphotericin B for cryptococcal meningitis. Side-effects are uncommon, although it may cause bone marrow suppression. It is active when given orally or parenterally.

Griseofulvin. Griseofulvin, a naturally occurring antifungal, was formerly widely used for the treatment of more extensive superficial mycoses and onychomycosis for which terbinafine is now the preferred agent.

Amorolfine. Amorolfine is available for the topical treatment of fungal skin and nail infections.

Whitfield's ointment. Whitfield's ointment is a traditional remedy for dermatophyte infection and contains benzoic and salicylic acid. Undecylenic acid (undecenoic) is also used in the treatment of tinea pedis.

ANTIVIRAL DRUGS

Drugs for HIV infection are discussed on page 192.

Nucleoside analogues

Aciclovir. Aciclovir (also known as acycloguanosine, Fig. 4.13a) is an acyclic nucleoside analogue which acts as a chain terminator of herpesvirus DNA synthesis. This drug is converted to aciclovir monophosphate by a virus-encoded thymidine kinase produced by alpha herpesviruses, herpes simplex types 1 and 2 and varicella zoster virus (Table 4.14). Conversion to the triphosphate is then achieved by cellular enzymes. Aciclovir triphosphate acts as a potent inhibitor of viral (but not cellular) DNA polymerase, and also competes with deoxyguanine triphosphate for incorporation into the growing chains of herpesvirus DNA, thereby resulting in

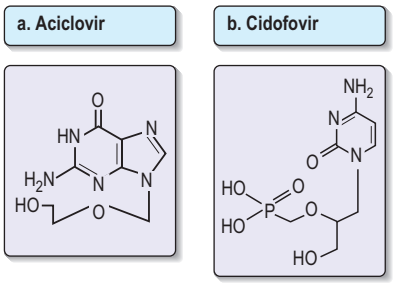


Fig. 4.13 **The structure of antiviral drugs**. (a) Aciclovir; (b) cidofovir.

Table 4.14 **Antiviral agents (for drugs against HIV see Table 4.53)**

Drug	Use
Nucleoside analogues	
Aciclovir	Topical – HSV infection Oral and intravenous – VZV and HSV
Famciclovir	Oral – VZV and HSV
Valaciclovir	Oral – VZV and HSV
Ganciclovir	Intravenous – CMV
Valganciclovir	Oral – CMV
Lamivudine	Oral – HBV infection
Entecavir	Oral – HBV
Nucleotide analogues	
Adefovir	Oral – HBV infection
Cidofovir	Intravenous – CMV
Pyrophosphate analogues	
Foscarnet	Intravenous – CMV
Adamantanes	
Amantadine	Oral – influenza A
Neuraminidase inhibitors	
Zanamivir	Topical (inhalation) – influenza A and B
Oseltamivir	Oral – influenza A and B
Ribavirin	Topical (inhalation) – RSV Oral – Lassa fever, hepatitis C
Palivizumab	Prevention of RSV
Alpha-interferon (INF-α) (Pegylated INF-α)	HBV, HCV, some malignancies (e.g. renal cell carcinoma) (Given once weekly)

chain synthesis termination, due to its acyclic structure. This highly specific mode of activity, targeted only to virus-infected cells, means that aciclovir has very low toxicity. Intravenous, oral and topical preparations are available for the treatment of herpes simplex and varicella zoster virus infections (Table 4.14). Treatment does not eliminate the virus so relapses occur.

A pro-drug of aciclovir, valaciclovir, has been developed. Coupling of the amino acid valine to the acyclic side-chain of aciclovir allows better intestinal absorption. The valine is removed by enzymic action and aciclovir is released into the circulation. A similar pro-drug of a related nucleoside analogue (penciclovir) is the antiherpes drug, famciclovir. The mode of action and efficacy of famciclovir are similar to those of aciclovir.

Ganciclovir. This guanine analogue is structurally similar to aciclovir, with extension of the acyclic side-chain by a carboxymethyl group. It is active against herpes simplex viruses and varicella zoster virus by the same mechanism as aciclovir. In addition, phosphorylation by a protein kinase encoded by the UL97 gene of cytomegalovirus renders it potently active against this virus. Thus ganciclovir is currently the first-line treatment for cytomegalovirus disease. Intravenous and oral preparations are available as is an oral pro-drug, valganciclovir. Unlike aciclovir, ganciclovir has a significant toxicity profile including neutropenia, thrombocytopenia and the likelihood of sterilization by inhibiting spermatogenesis. For this reason, it is reserved for the treatment or prevention of life- or sight-threatening cytomegalovirus infection in immunocompromised patients. Use protective gloves when administering ganciclovir.

Lamivudine is a reverse transcriptase inhibitor active against both HIV and hepatitis B. Drug resistance is a problem (see p. 196 and p. 339).

Entecavir is a more potent inhibitor of hepatitis B and is active against lamivudine-resistant strains. Telbivudine and emtricitabine are also nucleoside analogue inhibitors of HBV DNA polymerase activity.

Nucleotide analogues

Cidofovir (Fig. 4.13b). This is a phosphonate derivative of an acyclic nucleoside which is a DNA polymerase chain inhibitor. It is administered intravenously for the treatment of severe cytomegalovirus (CMV) infections in patients with AIDS for whom other drugs are inappropriate. It is given with probenecid, and as it is nephrotoxic, particular attention should be given to hydration and to monitoring renal function.

Adefovir dipivoxil is used in the treatment of chronic hepatitis B (see p. 339).

Pyrophosphate analogues

Foscarnet (sodium phosphonoformate) is a simple pyrophosphate analogue which inhibits viral DNA polymerases. It is active against herpesviruses and its main roles are as a second-line treatment for severe cytomegalovirus disease and for the treatment of aciclovir-resistant herpes simplex infection. It is given intravenously and the potential for severe side-effects, particularly renal damage, limits its use.

Adamantanes

Amantadine. Amantadine is a synthetic symmetrical amine which is active prophylactically and therapeutically against influenza A virus (it is inactive against influenza B virus). Its prophylactic efficacy is similar to that of influenza vaccine and it is occasionally used to prevent the spread of influenza A in institutions such as nursing homes. Although CNS side-effects such as insomnia, dizziness and headache may occur (it is also used as a treatment for Parkinson's disease), these are not usually produced by the lower doses currently recommended.

Neuraminidase inhibitors

Two drugs that inhibit the action of the neuraminidase of influenza A and B have been introduced. Zanamivir is administered by inhalation and oseltamivir is an oral preparation. Both have been shown to be effective in reducing the

duration of illness in influenza. Oseltamivir is also available for the prophylaxis of influenza among household contacts of an index case.

Ribavirin

This synthetic purine nucleoside derivative which interferes with 5′-capping of messenger RNA, is active against several RNA and DNA viruses, at least in vitro. Its major use is in the treatment of chronic hepatitis C infection in combination with pegylated interferon-α (see p. 344). It is administered orally. Haemolytic anaemia is the most frequent adverse reaction. It is also administered by a small-particle aerosol generator (SPAG) to infants with acute respiratory syncytial virus (RSV) infection. Another indication is in the treatment of Lassa fever virus infection.

www.studentconsult.com
More online

Palivizumab

This monoclonal antibody is specifically indicated to prevent seasonal respiratory syncytial virus (RSV) infection in infants at high risk of this infection. It is administered by intramuscular injection.

Interferons

These are naturally occurring proteins with a multiplicity of actions, including antiviral, immunomodulatory and antiproliferative effects. Interferons are produced by virus-infected cells, macrophages and lymphocytes. Their production is stimulated by a number of factors, including viral nucleic acid. They induce an antiviral state in uninfected cells, through activation of a complex set of biochemical pathways. They have been synthesized commercially by either culture of lymphoblastoid cells or by recombinant DNA technology and are licensed for therapeutic use.

The potency of INF-α has been enhanced by coupling the protein with polyethylene glycol. The resulting PEG interferon given once weekly has been shown to improve the response, and reduce the side-effects, to treatment of hepatitis B and C. When combined with oral ribavirin, selected patients with chronic hepatitis C can be cleared of the virus.

FURTHER READING

De Clercq E. Antiviral drugs in current clinical use. *Journal of Clinical Virology* 2004; **30**: 115–133.

Falagas ME, Giannopoulou KP, Ntziora F et al. Daptomycin for endocarditis and/or bacteraemia: a systematic review of the experimental and clinical evidence. *Journal of Antimicrobial Chemotherapy* 2007; **60**: 7–19.

Falagas ME, Manta KG, Ntziora F et al. Linezolid for the treatment of patients with endocarditis: a systematic review of the published evidence. *Journal of Antimicrobial Chemotherapy* 2006; **58**: 273–280.

Gruchalla RS, Pirmohamed M. Antibiotic allergy. *New England Journal of Medicine* 2006; **354**: 601–609.

Jacoby GA, Munoz-Price LS. The new β-lactamases. *New England Journal of Medicine* 2004; **352**: 380–391.

Shah PM, Isaacs RD. Ertapenem, the first of a new group of carbapenems. *Journal of Antimicrobial Chemotherapy* 2003; **52**: 538–542.

Wenzel R, Del Favero A, Kibler C et al. Economic evaluation of voriconazole compared with conventional amphotericin B for the primary treatment of aspergillosis in immunocompromised patients. *Journal of Antimicrobial Chemotherapy* 2005; **55**: 352–361.

IMMUNIZATION AGAINST INFECTIOUS DISEASES

Although effective antimicrobial chemotherapy is available for many diseases, the ultimate aim of any infectious disease control programme is to prevent infection occurring. This is achieved either by:

- eliminating the source or mode of transmission of an infection (see p. 81)
- reducing host susceptibility to environmental pathogens.

Immunization, immunoprophylaxis and immunotherapy

Immunization has changed the course and natural history of many infectious diseases. Passive immunization by administering preformed antibody, either in the form of immune serum or purified normal immunoglobulin, provides short-term immunity and has been effective in both the prevention (immunoprophylaxis) and treatment (immunotherapy) of a number of bacterial and viral diseases (Table 4.15). The active immunization schedule currently recommended is summarized in Box 4.6. Long-lasting immunity is achieved only by active immunization with a live attenuated or an inactivated organism, or a subunit thereof (Table 4.16). Active immunization may also be performed with microbial toxin (either native or modified) – that is, a toxoid. Immunization should be kept up to date with booster doses throughout life. Travellers to developing countries, especially if visiting rural areas, should in addition enquire about further specific immunizations.

Table 4.15 Examples of passive immunization available

Infection	Antibody	Indication	Efficacy
Bacterial			
Tetanus	Human tetanus immune globulin	Prevention and treatment	+
Diphtheria	Horse serum	Prevention and treatment	+
Botulism	Horse serum	Treatment	+
Viral			
Hepatitis A Measles	Human normal immune globulin	Prevention (rarely required since vaccine introduced)	+
Hepatitis B	Human hepatitis B immune globulin	Prevention	+
Varicella zoster	Human varicella zoster immune globulin	Prevention	+
Rabies	Human rabies immune globulin	Prevention (post-exposure)	+

Box 4.6 Recommended immunization schedules: (i) in the UK; (ii) WHO model schedule for developing countries

Time of immunization	Vaccine
(i) UK	
2 months	DTaP, IPV, Hib + PCV (BCG*)
3 months	DTaP, IPV, Hib + MenC
4 months	DTaP, IPV, Hib + MenC + PCV
12 months	Hib, Men C
13 months	MMR + PCV
By school entry	DTaP, IPV, MMR
6–16 years	(BCG†)
13–18 years	Td, IPV, (HPV‡)
(ii) Developing countries§	
Birth (or first contact)	OPV, BCG
6 weeks	DPT, OPV, HBV
10 weeks	DPT, OPV, HBV
14 weeks	DPT, OPV, HBV
9 months	Measles, YF¶

DPT, adsorbed diphtheria, whole cell pertussis, tetanus triple vaccine; Hib, *Haemophilus influenzae* b vaccine; IPV, inactivated polio vaccine; OPV, oral polio vaccine; MenC, meningococcus group C conjugate vaccine; aP, acellular pertussis; MMR, measles, mumps, rubella triple vaccine; d, adsorbed low-dose diphtheria; HBV, hepatitis B vaccine; HPV, human papillomavirus vaccine; YF, yellow fever vaccine; PCV, pneumococcal conjugate vaccine

*Children at high risk of contact with tuberculosis

†Tuberculin-negative children at low risk of contact with tuberculosis

‡For girls aged 12–13 years

§Model scheme, adapted locally depending on need and availability of vaccines

¶In endemic areas

Mumps vaccine is given in many developing countries

In 1974 the World Health Organization introduced the Expanded Programme on Immunization (EPI). Twenty years later more than 80% of the world's children had been immunized against tuberculosis, diphtheria, tetanus, pertussis, polio and measles. Poliomyelitis should shortly be eradicated world-wide, which will match the past success of global smallpox eradication. Introduction of conjugate vaccines against *Haemophilus influenzae* type b (Hib) has proved highly effective in controlling invasive *H. influenzae* infection, notably meningitis (see p. 1153).

FURTHER READING

Chaves SS, Gargiullo P, Zhang JX et al. Loss of vaccine-induced immunity to varicella over time. *New England Journal of Medicine* 2007; **356**: 1121–1129.

Crowcroft NS, Pebody RG. Recent developments in pertussis. *Lancet* 2006; **367**: 1926–1936.

Gardner P. Prevention of meningococcal disease. *New England Journal of Medicine* 2006; **355**: 1466–1473.

NHS Immunisation Information. http://www.immunisation.nhs.uk.

Pallansch MA, Sandhu HS. The eradication of polio – progress and challenges. *New England Journal of Medicine* 2006; **355**: 2508–2511.

Salisbury D, Ramsay M, Noakes K (eds). *Immunization Against Infectious Diseases.* London: HMSO, 2006.

Whitney CG, Pilishvili T, Farley MM et al. Effectiveness of 7-valent pneumococcal conjugate vaccine against invasive pneumococcal disease: a matched case-control study. *Lancet* 2006; **368**: 1495–1502.

Table 4.16 Preparations available for active immunization

Live attenuated vaccines

Influenza (not in the UK)

Oral polio (Sabin) (not recommended, only used for outbreaks)

Measles

Mumps

Rotavirus

Rubella

Varicella zoster

Yellow fever

BCG

Typhoid

Inactivated and/or conjugate vaccines

Hepatitis A

Pertussis

Typhoid – whole cell and Vi antigen

Polio (Salk) for routine use

Influenza

Human papillomavirus (HPV)

Cholera (oral)

Meningococci (groups A and C)

Meningococcus group C

Rabies

Pneumococcal

Haemophilus influenzae type b

Toxoids

Diphtheria

Tetanus

Recombinant vaccines

Hepatitis B

BCG, bacille Calmette–Guérin.

Protection for travellers to developing/tropical countries

There has been a huge increase in the number of people travelling to developing countries, mainly for recreation and leisure. The risk of infection depends on the area to be visited, the type of activity and the underlying health of the traveller. Advice should therefore always be based on an individual assessment.

Protection for travellers can be divided into three categories:

- Personal protection, e.g.
  - insect repellents
  - impregnated bed netting
  - avoidance of animals
  - care with food and drink
- Chemoprophylaxis, e.g. antimalarials
- Immunization, e.g.
  - yellow fever
  - hepatitis A and B
  - typhoid.

Because situations and risks can change rapidly, websites (which can be regularly updated) are often the best source of advice on travel health.

FURTHER READING

CDC Travel Health Site. http://www.cdc.gov/travel.

Department of Health (UK). Health Information for Overseas Travel. http://www.archive.official-documents.co.uk/document/doh/hinfo/travel02.htm

Lu C, Michaud CM, Gakidou E et al. Effect of the global alliance for vaccines and immunisation on diphtheria, tetanus and pertussis vaccine coverage: an independent assessment. *Lancet* 2006; **368**: 1088–1095.

VIRAL INFECTIONS

Introduction

Viruses are much smaller than other infectious agents (see Tables 4.17 and 4.19) and contain either DNA or RNA, not both as in bacteria and other microorganisms. Since they are metabolically inert, they must live intracellularly, using the host cell for synthesis of viral proteins and nucleic acid. Viruses have a central nucleic acid core surrounded by a protein coat that is antigenically unique for a particular virus. The protein coat (capsid) imparts a helical or icosahedral structure to the virus. Some viruses also possess an envelope consisting of lipid and protein.

Outcomes of virus infection of a cell

Replication of viruses within a cell may result in sufficient distortion of normal cell function so as to result in cell death – a cytocidal or cytolytic infection. However, acute cell death is not an inevitable consequence of virus infection of a cell. In a chronic, or persistent, infection, virus replication continues throughout the lifespan of the cell, but does not interfere with the normal cellular processes necessary for cell survival. Hepatitis B and C viruses may interact with cells in this way. Some viruses, e.g. the herpesvirus family, are able to go latent within a cell – in such a state, the virus genome is present within the cell, but there is very little, if any, production of viral proteins, and no production of mature virus particles. Finally, some viruses are able to transform cells, leading to uncontrolled cell division, e.g. Epstein–Barr virus infection of B lymphocytes, resulting in the generation of an immortal lymphoblastoid cell line.

DNA VIRUSES

Details of the structure, size and classification of human DNA viruses are shown in Table 4.17.

ADENOVIRUSES

Over 50 adenovirus serotypes have been identified as human pathogens, infecting a number of different cell types, and therefore resulting in different clinical syndromes. Adenovirus infection commonly presents as an acute pharyngitis, and extension of infection to the larynx and trachea in infants may lead to croup. By school age the majority of children show serological evidence of previous infection. Certain subtypes produce an acute conjunctivitis associated with pharyngitis. In adults, adenovirus causes acute follicular conjunctivitis and rarely pneumonia that is clinically similar to that produced by *Mycoplasma pneumoniae* (see p. 859). Certain adenoviruses (Ads 40 and 41) cause gastroenteritis (see p. 128) without respiratory disease and adenovirus infection may be responsible for acute mesenteric lymphadenitis in children and young adults. Mesenteric adenitis due to adenoviruses may lead to intussusception in infants. Infection in an immunocompromised host, e.g. a bone marrow transplant recipient, may result in multisystem failure failure and fatal disease.

HERPESVIRUSES

Members of the *Herpesviridae* family are causes of a wide range of human diseases. Details are summarized in Table 4.18. The hallmark of all herpesvirus infections is the ability of the viruses to establish latent (or silent) infections that then persist for the life of the individual.

Herpes simplex virus (HSV) infection

(Fig. 4.14)

Two types of HSV have been identified: HSV-1 is the major cause of herpetic stomatitis, herpes labialis ('cold sore'), keratoconjunctivitis and encephalitis, whereas HSV-2 causes genital herpes and may also be responsible for systemic infection in the immunocompromised host. These divisions, however, are not rigid, for HSV-1 can give rise to genital herpes and HSV-2 can cause infections in the mouth.

Table 4.17 Human DNA viruses

Structure		Approximate size	Family	Viruses
Symmetry	Envelope			
Icosahedral	–	80 nm	*Adenoviridae*	Adenoviruses (>50 serotypes)
Icosahedral	+	100 nm (160 nm with envelope)	*Herpesviridae*	Herpes simplex virus (HSV) types 1 and 2
				Varicella zoster virus
				Cytomegalovirus
				Epstein–Barr virus (EBV)
				Human herpesvirus type 6 (HHV-6)
				Human herpesvirus type 7 (HHV-7)
				Human herpesvirus type 8 (HHV-8)
Icosahedral	+	42 nm	*Hepadnaviridae*	Hepatitis B virus (HBV)
Icosahedral	–	50 nm	*Papovaviridae*	Human papillomaviruses (>100 types)
				Polyomaviruses JC and BK
Icosahedral	–	23 nm	*Parvoviridae*	Erythrovirus B19, bocavirus
Complex	+	300 nm × 200 nm	*Poxviridae*	Variola virus
				Vaccinia virus
				Monkeypox
				Cowpox
				Orf
				Molluscum contagiosum

Table 4.18 Major diseases caused by human herpesviruses

Subfamily	Virus	Children	Adults	Immunocompromised
α-Herpesvirus	Herpes simplex type 1	Stomatitis*	Cold sores Herpes simplex encephalitis Keratitis Erythema multiforme	Dissemination
	Herpes simplex type 2	Neonatal herpes*	Primary genital herpes* Recurrent genital herpes	Dissemination
	Varicella zoster virus	Chickenpox*	Shingles	Dissemination
β-Herpesvirus	Cytomegalovirus	Congenital*		Pneumonitis Retinitis Gastrointestinal
	Human herpesvirus type 6	Roseola infantum*		Pneumonitis
	Human herpesvirus type 7	Roseola infantum*		
γ-Herpesvirus	Epstein–Barr virus		Infectious mononucleosis* Burkitt's lymphoma Nasopharyngeal carcinoma	Lymphoma
	Human herpesvirus type 8		Kaposi's sarcoma	Kaposi's sarcoma

*Signifies primary infection.

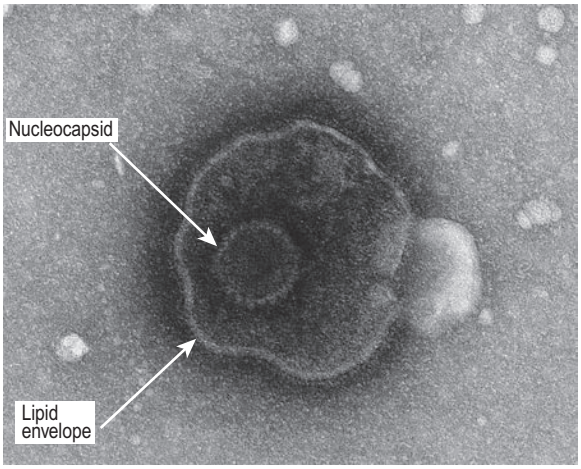


Fig. 4.14 **Electronmicrograph of herpes simplex virus.**

Fig. 4.15 **Primary herpes simplex type 1 (gingivostomatitis).**

HSV-1

The portal of entry of HSV-1 infection is usually via the mouth or occasionally the skin. The primary infection may go unnoticed or may produce a severe inflammatory reaction with vesicle formation leading to painful ulcers (gingivostomatitis; see Fig. 4.15). The virus then remains latent, most commonly in the trigeminal ganglia, but may be reactivated by stress, trauma, febrile illnesses and ultraviolet radiation, producing the recurrent form of the disease known as herpes labialis ('cold sore'). Approximately 70% of the population are infected with HSV-1 and recurrent infections occur in one-third of individuals. Reactivation often produces localized paraesthesiae in the lip before the appearance of a cold sore.

Complications of HSV-1 infection include transfer to the eye (dendritic ulceration, keratitis), acute encephalitis (see p. 1156), skin infections such as herpetic whitlow and erythema multiforme (see p. 1250).

HSV-2

The clinical features, diagnosis and management of genital herpes are described on pages 180 and 1231. The virus remains latent in the sacral ganglia and during recurrence can produce a radiculomyelopathy, with pain in the groin, buttocks and upper thighs. Primary anorectal herpes infection is common in men having sex with men (see p. 180).

Neonates may develop primary HSV infection following vaginal delivery in the presence of active genital HSV infection in the mother, particularly if the maternal disease is a primary, rather than a recurrent infection. The disease in the baby varies from localized skin lesions (about 10–15%) to widespread visceral disease often with encephalitis, with a poor prognosis. Caesarean section should therefore be performed if active genital HSV infection is present during labour.

Immunocompromised patients such as those receiving intensive cancer chemotherapy or those with the acquired immunodeficiency syndrome (AIDS) may develop disseminated HSV infection involving many of the viscera. In severe cases death may result from hepatitis and encephalitis.

Humoral antibody develops following primary infection, but mononuclear cell responses probably prevent dissemination of disease.

The clinical picture, diagnosis and treatment are described on page 180.

Varicella zoster virus (VZV) infection

VZV produces two distinct diseases, varicella (chickenpox) and herpes zoster (shingles). The *primary infection* is chickenpox. It usually occurs in childhood, the virus entering through the mucosa of the upper respiratory tract. In some countries (e.g. the Indian subcontinent) a different epidemiological pattern exists with most infections occurring in adulthood. Chickenpox rarely occurs twice in the same individual. Infectious virus is spread from fresh skin lesions by direct contact or airborne transmission and the period of infectivity in chickenpox extends from 2 days before the appearance of the rash until the skin lesions are all at the crusting stage. Following recovery from chickenpox the virus remains latent in dorsal root and cranial nerve ganglia. *Reactivation* of infection then results in shingles.

Clinical features of chickenpox

Fourteen to twenty-one days after exposure to VZV, a brief prodromal illness of fever, headache and malaise heralds the eruption of chickenpox, characterized by the rapid progression of macules to papules to vesicles to pustules in a matter of hours (Fig. 4.16). In young children the prodromal illness may be very mild or absent. The illness tends to be more severe in older children and can be debilitating in adults. The lesions occur on the face, scalp and trunk, and to a lesser extent on the extremities. It is characteristic to see skin lesions at all stages of development on the same area of skin. Fever subsides as soon as new lesions cease to appear. Eventually the pustules crust and heal without scarring.

Complications of chickenpox include pneumonia, which generally begins 1–6 days after the skin eruption, and bacterial superinfection of skin lesions. Pneumonia is more common in adults than in children and cigarette smokers are at particular risk. Pulmonary symptoms are usually more striking than the physical findings, although a chest radiograph usually shows diffuse changes throughout both lung fields. CNS involvement occurs in about 1 per 1000 cases and most commonly presents as an acute truncal cerebellar ataxia. The immunocompromised are susceptible to disseminated infection with multiorgan involvement. Women in pregnancy are prone to severe chickenpox and, in addition, there is a risk of intrauterine infection with structural damage to the fetus (if maternal infection is within the first 20 weeks of pregnancy, the risk of varicella embryopathy is 1–2%).

Clinical features of shingles

Shingles (see p. 1231) arises from the reactivation of virus latent within the dorsal root or cranial nerve ganglia. It may occur at all ages but is most common in the elderly, producing skin lesions similar to chickenpox, although classically they are unilateral and restricted to a sensory nerve (i.e. dermatomal) distribution (Fig. 4.17). The onset of the rash of shingles is usually preceded by severe dermatomal pain, indicating the involvement of sensory nerves in its pathogenesis. Virus is disseminated from freshly formed vesicles and may cause chickenpox in susceptible contacts.

The commonest complication of shingles is post-herpetic neuralgia (PHN) (see p. 1157).

Diagnosis

The diseases are usually recognized clinically but can be confirmed by detection of VZV DNA within vesicular fluid using PCR, electronmicroscopy, immunofluorescence or culture of vesicular fluid, and by serology.

Prophylaxis and treatment

Chickenpox usually requires no treatment in healthy children and infection results in lifelong immunity. Aciclovir and derivatives are, however, licensed for this indication in the USA, where the argument for treatment is one of health economics, *viz.* the sooner the child recovers, the sooner the carer can return to work. However, the disease may be fatal in the immunocompromised, who should therefore be offered protection, after exposure to the virus, with zoster immune immunoglobulin (ZIG).

Anyone with chickenpox who is over the age of 16 years should be given antiviral therapy with aciclovir, or a similar drug, if they present within 72 hours of onset. Prophylactic ZIG is recommended for women in pregnancy exposed to varicella zoster virus and, if chickenpox develops, aciclovir treatment should be given (NB: aciclovir has not been licensed for use in pregnant women). If a woman has chickenpox at term, her baby should be protected by ZIG if delivery occurs within 7 days of the onset of the mother's rash. An effective live attenuated varicella vaccine is licensed as a routine vaccination of childhood in the USA; it is available on a named-patient basis in the UK, and also for susceptible healthcare workers.

Shingles involving motor nerves, e.g. 7th cranial nerve leading to facial palsy, is also treated with aciclovir (or deriva-

Fig. 4.16 **Chickenpox in an adult.** Generalized VZV.

Fig. 4.17 **Shingles – VZV affecting a dermatome.** Reproduced with kind permission of Imperial College School of Medicine.

www.studentconsult.com
Ask the authors

tives thereof) as the duration of lesion formation and time to healing can be reduced by early treatment. Aciclovir, valaciclovir and famciclovir have all been shown to reduce the burden of zoster-associated PHN when treatment is given in the acute phase. Shingles involving the ophthalmic division of the trigeminal nerve has an associated incidence of acute and chronic ophthalmic complications of 50%. Early treatment with aciclovir reduces this to 20% or less. As for chickenpox, all immunocompromised individuals should be given aciclovir at the onset of shingles, no matter how mild the attack appears when it first presents.

Recent evidence suggests that vaccination of all adults over the age of 60, with a dose higher than that used for chickenpox prophylaxis in childhood, reduces shingles-related morbidity and post-herpetic neuralgia.

Cytomegalovirus (CMV) infection

Infection with CMV is found world-wide. Clinically significant CMV infection arises particularly in two patient groups – fetuses who acquire the infection transplacentally and are born congenitally infected, and patients who are immunosuppressed, e.g. transplant recipients, or patients with HIV infection. Over 50% of the adult population have serological evidence of latent infection with the virus. As with all herpesviruses, the virus persists for life, usually as a latent infection in which the naked DNA is situated extrachromosomally in the nuclei of the cells in the endothelium of the arterial wall and in T lymphocytes.

Clinical features

In healthy adults CMV infection is usually asymptomatic but may cause an illness similar to infectious mononucleosis, with fever, occasionally lymphocytosis with atypical lymphocytes, and hepatitis with or without jaundice. The Paul–Bunnell test for heterophile antibody is negative. Infection may be spread by kissing, sexual intercourse or blood transfusion, and transplacentally to the fetus. Disseminated fatal infection with widespread visceral involvement occurs in the immunocompromised (see p. 202) and may cause encephalitis, retinitis, pneumonitis and diffuse involvement of the gastrointestinal tract.

Intrauterine infection may arise from either primary or reactivated maternal infection. CMV is, by far, the commonest congenital infection – in developed countries, such as the UK, 0.3–1% of all babies are born congenitally infected with CMV. Around 5–10% of such babies have severe disease evident at birth, with a poor prognosis. CNS involvement may cause microcephaly and motor disorders. Jaundice and hepatosplenomegaly are common, and thrombocytopenia and haemolytic anaemia also occur. Periventricular calcification is seen on skull X-ray. A further 5–10% of infected babies are normal at birth, but developmental abnormalities become apparent later, e.g. sensorineural nerve deafness. The remaining 80–85% of infected babies are normal at birth and develop normally.

Diagnosis

Most infections are now diagnosed by detection and quantification of CMV DNA or RNA using molecular amplification techniques, in blood or other body fluid samples. The virus can also be identified in tissues by the presence of characteristic intranuclear 'owl's eye' inclusions (Fig. 4.18) on histological staining and by direct immunofluorescence. Culture in human embryo fibroblasts is usually slow but diagnosis can be accelerated by immunofluorescent detection of antigen in the cultures. Serological tests can identify latent (IgG) or primary (IgM) infection.

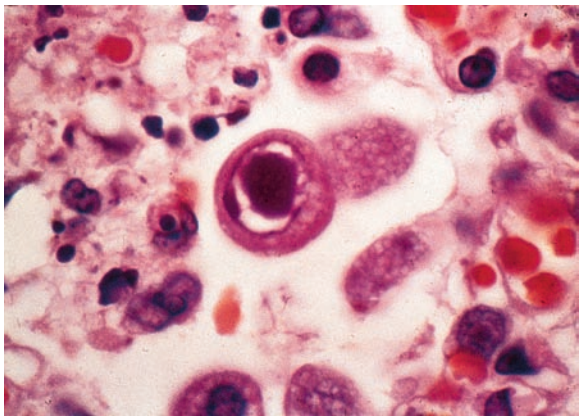


Fig. 4.18 **Typical 'owl-eye' inclusion-bearing cell infected with cytomegalovirus.**

Treatment

In the immunocompetent, infection is usually self-limiting and no specific treatment is required. In the immunosuppressed, ganciclovir (5 mg/kg daily for 14–21 days) reduces retinitis and gastrointestinal damage and can eliminate CMV from blood, urine and respiratory secretions. It is less effective against pneumonitis. In patients who are continually immunocompromised, particularly those with AIDS, maintenance therapy may be necessary. Drug resistance has been reported in AIDS patients and transplant recipients Bone marrow toxicity is common. Valganciclovir, foscarnet and cidofovir are available (see p. 101). Treatment of CMV in neonates is difficult, but ganciclovir therapy of infected babies with evident CNS involvement has been shown to improve long-term hearing outcome.

Epstein–Barr virus (EBV) infection

Globally, most individuals are infected with this virus at an early age (0–5 years), at which time clinical symptoms are unusual. Infection at an older age is associated with an acute febrile illness known as infectious mononucleosis (glandular fever), which occurs world-wide in adolescents and young adults. EBV is probably transmitted in saliva and by aerosol.

Clinical features

The predominant symptoms of infectious mononucleosis are fever, headache, malaise and sore throat. Palatal petechiae and a transient macular rash are common, the latter occurring in 90% of patients who have received ampicillin (inappropriately) for the sore throat. Cervical lymphadenopathy, particularly of the posterior cervical nodes, and splenomegaly are characteristic. Mild hepatitis is common, but other complications such as myocarditis, meningitis, encephalitis, mesenteric adenitis and splenic rupture are rare.

Although some young adults remain debilitated and depressed for some months after infection, the evidence for reactivation of latent virus in healthy individuals is

controversial, although this is thought to occur in immunocompromised patients.

Following primary infection, EBV remains latent in resting memory B lymphocytes. It has been shown in vitro that of nearly 100 viral genes expressed during replication, approximately only 10 are expressed in the latently infected B cells.

Severe, often fatal infectious mononucleosis may result from a rare X-linked lymphoproliferative syndrome affecting young boys. Those who survive have an increased risk of hypogammaglobulinaemia and/or lymphoma.

EBV is the cause of oral hairy leucoplakia in AIDS patients and is the major aetiological agent responsible for Burkitt's lymphoma, undifferentiated nasopharyngeal carcinoma, post-transplant lymphoma, the immunoblastic lymphoma of AIDS patients and some forms of Hodgkin's lymphoma and gastric cancer. Different levels of expression of EBV latency genes occur in these proliferative conditions caused by the virus, and various co-factors are also involved in their pathogenesis, e.g. in Burkitt's lymphoma, the commonest tumour of childhood in sub-Saharan Africa, the interplay between EBV infection and the presence of hyperendemic (i.e. present all year-round) malaria appears to be relevant.

Diagnosis

EBV infection should be strongly suspected if atypical mononuclear cells (activated CD8 positive T lymphocytes, also known as glandular fever cells) are found in the peripheral blood. It can be confirmed during the second week of infection by a positive Paul–Bunnell reaction, which detects heterophile antibodies (IgM) that agglutinate sheep erythrocytes. False-positives can occur in other conditions such as viral hepatitis, Hodgkin's lymphoma and acute leukaemia. The Monospot test is a sensitive and easily performed screening test for heterophile antibodies. Specific EBV IgM antibodies indicate recent infection by the virus. Clinically similar illnesses are produced by CMV, toxoplasmosis and acute HIV infection (the so-called seroconversion illness) but these can be distinguished serologically.

Treatment

The majority of cases require no specific treatment and recovery is rapid. Corticosteroid therapy is advised when there is neurological involvement (e.g. encephalitis, meningitis, Guillain–Barré syndrome), when there is marked thrombocytopenia or haemolysis, or when the tonsillar enlargement is so marked as to cause respiratory obstruction.

Human herpesvirus type 6 (HHV-6)

This human herpesvirus infects $CD4^+$ T lymphocytes, occurs world-wide, and exists as a latent infection in over 85% of the adult population. It is spread by contact with oral secretions. The virus causes roseola infantum (exanthem subitum) which presents as a high fever followed by generalized macular rash in infants. HHV-6 is a common cause of febrile convulsions, and aseptic meningitis or encephalitis occur as rare complications. Reactivation in the immunocompromised may lead to severe pneumonia.

Treatment

Supportive management only is recommended for the common infantile disease. Ganciclovir can be used in the immunocompromised.

Human herpesvirus type 7 (HHV-7)

This virus is similar to HHV-6 in being a T lymphotropic herpesvirus. It is also present as a latent infection in over 85% of the adult population and it is known to infect CD4+ helper T cells by using the CD4 antigen (the main receptor employed by HIV). The full spectrum of disease due to HHV-7 has not yet been fully characterized, but, like HHV-6, it may cause roseola infantum in infants.

Human herpesvirus type 8 (Kaposi's sarcoma-associated herpesvirus)

This human herpesvirus is strongly associated with the aetiology of classical and AIDS-related Kaposi's sarcoma. Antibody prevalence is high in those with tumours but relatively low in the general population of most industrialized countries. High rates of infection (>50% population) have been described in central and southern Africa and this matches the geographic distribution of Kaposi's sarcoma before the era of AIDS. HHV-8 can be sexually transmitted among men having sex with men, through heterosexual sex and through exposure to blood from needle sharing. It is thought that salivary transmission may be the predominant route in Africa. HHV-8 RNA transcripts have been detected in Kaposi's sarcoma cells and in circulating mononuclear cells from patients with the tumour. This virus also has an aetiological role in two rare lymphoproliferative diseases – multicentric *Castleman's disease* (a disorder of the plasma cell type) and *primary effusion lymphoma* (body-cavity-based lymphoma), which is characterized by pleural, pericardial or peritoneal lymphomatous effusions in the absence of a solid tumour mass.

PAPOVAVIRUSES

This virus family comprises the *pa*pilloma, *po*lyoma and *va*cuolar viruses. These viruses tend to produce chronic infections, often with evidence of latency. They are capable of inducing neoplasia in some animal species and were among the first viruses to be implicated in tumorigenesis. Human papillomaviruses, of which there are at least 100 types, are responsible for the common skin and genital warts and certain types (mainly 16 and 18) are the cause of carcinoma of the cervix and some oral cancers (type 16). The realization that cancer of the cervix is an infectious disease has led to the development of papillomavirus vaccines, which have been shown to prevent disease associated with the high-risk HPV types 16 and 18, and have recently been licensed for use in the USA, Australia and Europe. The current recommendations are for vaccination of all girls at age 9–14 years.

For *genital warts* see page 181.

The *human BK* virus, a polyomavirus, is generally found in immunocompromised individuals and may be detected in the urine of 15–40% of renal transplant patients. Rarely, this is associated with impairment of function of the transplanted kidney – BK nephropathy. A related virus, JC, is the cause of progressive multifocal leucoencephalopathy (PML) which presents as dementia in the immunocompromised and is due to progressive cerebral destruction resulting from accumulation of the virus in brain tissue. Two further human polyomaviruses, WU and KI, have been recently identified. These may

DNA viruses
Papovaviruses
Erythroviruses
Poxviruses
Hepadna viruses
RNA viruses
Picornaviruses

be associated with respiratory tract infections in young children.

ERYTHROVIRUSES

Human erythrovirus B19 (also known as parvovirus) causes erythema infectiosum (fifth disease), a common infection in schoolchildren. The rash is typically on the face (the 'slapped-cheek' appearance). The patient is well and the rash can recur over weeks or months. Asymptomatic infection occurs in 20% of children. Non-specific respiratory tract illness is another common manifestation of infection. Moderately severe self-limiting arthropathy (see p. 535) is common if infection occurs in adulthood, especially in women. Aplastic crisis may occur in patients with chronic haemolysis (e.g. sickle cell disease). Chronic infection with anaemia occurs in immunocompromised subjects. Hydrops fetalis (3% risk) and spontaneous abortion (9% risk) may result from infection during the first and second trimesters of pregnancy.

Bocavirus is a recently identified erythrovirus which accounts for around 3–5% of respiratory tract infections in young children.

POXVIRUSES

Smallpox (variola)

This disease was eradicated in 1977 following an aggressive vaccination policy and careful detection of new cases coordinated by the World Health Organization. Its possible use in bioterrorism has resulted in the reintroduction of smallpox vaccination in some countries (see p. 959).

Monkeypox

This is a rare zoonosis that occurs in small villages in the tropical rainforests in several countries of western and central Africa. Its clinical effects, including a generalized vesicular rash, are indistinguishable from smallpox, but person-to-person transmission is unusual. Most infections occur in children who have not been vaccinated against smallpox. Disease can be severe, with mortality rates of 10–15% in unvaccinated individuals. Serological surveys indicate that several species of squirrel are likely to represent the animal reservoir. The virus was introduced into the USA in 2003 via West African small mammals illegally imported as pets. Widespread infection of prairie dogs resulted, and there were 37 laboratory confirmed cases in humans, only two of which suffered complications (keratitis, encephalopathy).

Cowpox

Cowpox produces large vesicles which are classically on the hands in those in contact with infected cows. The lesions are associated with regional lymphadenitis and fever. Cowpox virus has been found in a range of species including domestic and wild cats, and the reservoir is thought to exist in a range of rodents.

Vaccinia virus

This is a laboratory virus and does not occur in nature in either humans or animals. Its origins are uncertain but it has been invaluable in its use as the vaccine to prevent smallpox. Vaccination is now not recommended except for laboratory personnel handling certain poxviruses for experimental purposes or in contingency planning to manage a deliberate release of smallpox virus. It is being assessed experimentally as a possible carrier for new vaccines.

Orf

This poxvirus causes contagious pustular dermatitis in sheep and hand lesions in humans (see p. 1232).

Molluscum contagiosum

This is discussed on page 1232.

HEPADNA VIRUSES

These partially double-stranded DNA viruses infect a number of species. The representative virus from this family which infects humans, hepatitis B virus, is discussed, along with other hepatitis viruses, on page 333.

FURTHER READING

Frey SE, Belshe RB. Poxvirus zoonoses. *New England Journal of Medicine* 2004; **350**: 324–327.

Heininger U, Seward JF. Varicella. *Lancet* 2006; **368**: 1365–1376.

Kahn JA, Burk RD. Papillomavirus vaccines in perspective. *Lancet* 2007; **369**: 2135–2137.

Young NS, Brown KE. Parvovirus B19. *New England Journal of Medicine* 2004; **350**: 586–598.

RNA VIRUSES (Table 4.19)

PICORNAVIRUSES (PICO = SMALL)

This is a large family of small RNA viruses, which includes the enteroviruses and rhinoviruses which infect humans. The term enterovirus refers to the enteric means of spread of these viruses, i.e. via the faeco-oral route. The enteroviruses include poliovirus types 1–3, coxsackie A and B viruses, echoviruses and enteroviruses (EV) 68–71.

Poliovirus infection (poliomyelitis)

Poliomyelitis occurs when a susceptible individual is infected with poliovirus type 1, 2 or 3. These viruses have a propensity for the nervous system, especially the anterior horn cells of the spinal cord and cranial nerve motor neurones. Poliomyelitis is found world-wide but its incidence has decreased dramatically following improvements in sanitation, hygiene and the widespread use of polio vaccines. Spread is usually via the faeco-oral route, as the virus is excreted in the faeces.

Clinical features

The incubation period is 7–14 days. Although polio is essentially a disease of childhood, no age is exempt. The clinical manifestations vary considerably.

Inapparent infection

Inapparent infection is common and occurs in 95% of infected individuals.

Abortive poliomyelitis

Abortive poliomyelitis occurs in approximately 4–5% of cases and is characterized by the presence of fever, sore throat and myalgia. The illness is self-limiting and of short duration.

Table 4.19 Human RNA viruses

Structure		Approximate size	Family	Viruses
Symmetry	Envelope			
Icosahedral	–	30 nm	Picornavirus	Poliovirus Coxsackievirus Echovirus Enterovirus 68–71 Hepatitis A virus Rhinovirus
Icosahedral	–	80 nm	Reovirus	Reovirus Rotavirus
Icosahedral	+	50–80 nm	Togavirus	Rubella virus Alphaviruses
Icosahedral	+	50–80 nm	Flavivirus	Yellow fever virus Dengue virus Japanese encephalitis virus West Nile virus Hepatitis C virus Tick-borne encephalitis virus
Spherical	+	80–100 nm	Bunyavirus	Congo–Crimean haemorrhagic fever Hantavirus Rift valley fever virus
Spherical	–	35–40 nm	Calicivirus	Norovirus Sapovirus
Spherical	–	28–30 nm	Astrovirus	Astrovirus
Spherical	–	30–34 nm	Hepevirus	Hepatitis E virus
Helical	+	80–120 nm	Orthomyxovirus	Influenza viruses A, B and C
Helical	+	100–300 nm	Paramyxovirus	Parainfluenza viruses 1–4 Measles virus Mumps virus Respiratory syncytial virus Human metapneumovirus Nipah virus Hendra virus
Helical	+	80–220 nm	Coronavirus	229E, OC43, NL63, HK1 SARS
Helical	+	60–175 nm	Rhabdovirus	Lyssavirus – rabies
Helical	+	100 nm	Retrovirus	Human immunodeficiency viruses (HIV 1 and 2) Human T cell lymphotropic viruses (HTLV 1 and 2)
Helical	+	100–300 nm	Arenavirus	Lassa virus Lymphocytic choriomeningitis virus
Pleomorphic	+	Filaments or circular forms; 100 × 130–2600 nm	Filovirus	Marburg virus Ebola virus

Non-paralytic poliomyelitis (poliovirus meningitis)

Non-paralytic poliomyelitis has features of abortive poliomyelitis as well as signs of meningeal irritation, but recovery is complete.

Paralytic poliomyelitis

Paralytic poliomyelitis occurs in approximately 0.1% of infected children (1.3% of adults). Several factors predispose to the development of paralysis:

- male sex
- exercise early in the illness
- trauma, surgery or intramuscular injection, which localize the paralysis
- recent tonsillectomy (bulbar poliomyelitis).

This form of the disease is characterized initially by features simulating abortive poliomyelitis. Symptoms subside for 4–5 days, only to recur in greater severity with signs of meningeal irritation and muscle pain, which is most prominent in the neck and lumbar region. These symptoms persist for a few days and are followed by the onset of asymmetric flaccid paralysis without sensory involvement. The paralysis is usually confined to the lower limbs in children under 5 years of age and the upper limbs in older children, whereas in adults it manifests as paraplegia or quadriplegia.

Bulbar poliomyelitis

Bulbar poliomyelitis is characterized by the presence of cranial nerve involvement and respiratory muscle paralysis.

Soft palate, pharyngeal and laryngeal muscle palsies are common.

Aspiration pneumonia, myocarditis, paralytic ileus and urinary calculi are late complications of poliomyelitis.

Diagnosis

The diagnosis is a clinical one. Distinction from Guillain–Barré syndrome is easily made by the absence of sensory involvement and the asymmetrical nature of the paralysis in poliomyelitis. Laboratory confirmation and distinction between the wild virus and vaccine strains is achieved by genome detection techniques, virus culture, neutralization and temperature marker tests.

Treatment

Treatment is supportive. Bed rest is essential during the early course of the illness. Respiratory support with intermittent positive-pressure respiration is required if the muscles of respiration are involved. Once the acute phase of the illness has subsided, occupational therapy, physiotherapy and occasionally surgery have roles in patient rehabilitation.

Prevention and control

Immunization (Box 4.6) has dramatically decreased the prevalence of this disease world-wide and global eradication of the virus, coordinated by the World Health Organization, is expected within the next 1–2 years. The virus remains endemic in only four countries – Nigeria, India, Pakistan and Afghanistan. However, there have been outbreaks in West Africa where cultural taboos have disrupted the polio vaccination campaign and also in the Sudan. Recent studies using inactivated IM poliovirus vaccine (IPV) have revealed greater potency than the original Salk IPV. The greater reliability of IPV in hot climates and the scientific and ethical problems of continuing to use oral polio vaccine (OPV) in countries free from poliomyelitis, mean that IPV has replaced OPV in the routine immunization schedules in many countries.

Coxsackievirus, echovirus and other enterovirus infections

These viruses are also spread by the faeco-oral route. They each have a number of different types and are responsible for a broad spectrum of disease involving the skin and mucous membranes, muscles, nerves, the heart (Table 4.20) and, rarely, other organs, such as the liver and pancreas. They are frequently associated with pyrexial illnesses and are the most common cause of aseptic meningitis.

Herpangina

This disease is mainly caused by Coxsackie A viruses and presents with a vesicular eruption on the fauces, palate and uvula. The lesions evolve into ulcers. The illness is usually associated with fever and headache but is short-lived, recovery occurring within a few days.

Hand, foot and mouth disease

This disease is mainly caused by Coxsackievirus A16 or A10. It is also the main feature of infection with enterovirus (EV)71. Oral lesions are similar to those seen in herpangina but may be more extensive in the oropharynx. Vesicles and a maculopapular eruption also appear, typically on the palms of the hands and the soles of the feet, but also on other parts of the body. This infection commonly affects children. Recovery occurs within a week.

Neurological disease

Other enteroviruses in addition to poliovirus can cause a broad range of neurological disease, including meningitis, encephalitis, and a paralytic disease similar to poliomyelitis. EV71 has particular predilection for neuroinvasion. Thus, in epidemics of EV71 infection, a variety of serious neuromotor syndromes arise, albeit in a minority of those infected.

Heart and muscle disease

Enterovirus infection is a cause of acute myocarditis and pericarditis, from which, in general, there is complete

Table 4.20 Picornavirus infections (excluding poliovirus and rhinovirus)

Disease	Coxsackievirus		Echovirus (types 1–9, 11–17, 29–33)	Enterovirus (types 68–71)
	A (types A1–A22, A24)	B (types B1–B6)		
Cutaneous and mucosal				
Herpangina	+++	+	+	
Hand, foot and mouth	+++	+		+
Erythematous rashes	+	+	+++	
Conjunctivitis			+	
Neurological				
Paralytic	+		±	+
Meningitis	++	++	+++	+
Encephalitis	++	++	±	+
Cardiac				
Myocarditis and pericarditis	+	+++	+	
Muscle				
Myositis (Bornholm disease)	+	+++	+	

+++, often causes; ++, sometimes causes; +, rarely causes; ±, possibly causes.

recovery. However, these viruses can also cause chronic congestive cardiomyopathy and, rarely, constrictive pericarditis.

Skeletal muscle involvement, particularly of the intercostal muscles, is a feature of Bornholm disease, a febrile illness usually due to Coxsackievirus B. The pain may be of such an intensity as to mimic pleurisy or an acute abdomen. The infection affects both children and adults and may be complicated by meningitis or cardiac involvement.

Rhinovirus infection

Rhinoviruses are responsible for the common cold (see p. 831). Chimpanzees and humans are the only species to develop the common cold. ICAM-1 is the cellular receptor for rhinovirus and it is only in these two species that the specific binding domain is present. Peak incidence rates occur in the colder months, especially spring and autumn. There are multiple rhinovirus immunotypes (>100), which makes vaccine control impracticable. In contrast to enteroviruses, which replicate at 37°C, rhinoviruses grow best at 33°C (the temperature of the upper respiratory tract), which explains the localized disease characteristic of common colds.

REOVIRUSES

Reovirus infection

Reovirus infection occurs mainly in children, causing mild respiratory symptoms and diarrhoea. A few deaths have been reported following disseminated infection of brain, liver, heart and lungs.

Rotavirus infection

Rotavirus (Latin *rota* = wheel) is so named because of its electronmicroscopic appearance with a characteristic circular outline with radiating spokes (Fig. 4.19). It is responsible world-wide for both sporadic cases and epidemics of diarrhoea, and is currently one of the most common causes of childhood diarrhoea. More than 600 000 infected children under the age of 5 years are estimated to die annually in resource-deprived countries, compared with 75–150 in the USA. The prevalence is higher during the winter months in non-tropical areas. Asymptomatic infections are common, and bottle-fed babies are more likely to be symptomatic than those that are breast-fed.

Adults may become infected with rotavirus but symptoms are usually mild or absent. The virus may, however, cause diarrhoea in immunosuppressed adults, or outbreaks in patients on care of the elderly wards.

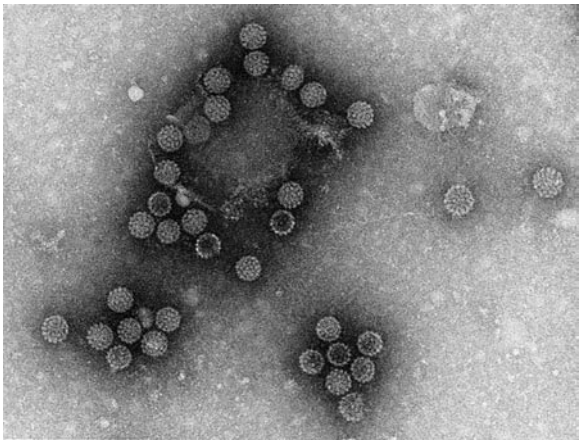


Fig. 4.19 **Electronmicrograph of human rotavirus.**

Clinical features

The illness is characterized by vomiting, fever, diarrhoea, and the metabolic consequences of water and electrolyte loss.

Diagnosis and differential diagnosis

The diagnosis can be established by PCR for genome detection, or ELISA for the detection of rotavirus antigen in faeces and by electronmicroscopy of faeces. Histology of the jejunal mucosa in children shows shortening of the villi, with crypt hyperplasia and mononuclear cell infiltration of the lamina propria.

Treatment and prevention

Treatment is directed at overcoming the effects of water and electrolyte imbalance with adequate oral rehydration therapy and, when indicated, intravenous fluids (see Box 4.9, p. 132). Antibiotics should not be prescribed. A controlled trial in Egypt in children with rotavirus diarrhoea demonstrated faster recovery (31 h versus 75 h) in those given nitazoxanide, a broad-spectrum anti-infective agent, for 3 days compared with placebo.

Rotavirus vaccines

Despite a major setback when the first licensed rotavirus vaccine was rapidly withdrawn from the market in 1999 following reports of increased rates of intususception, two new vaccines have been developed, and are now licensed for use in many countries. Both are live vaccines. One vaccine contains an attenuated human strain with the relevant antigens being P[8] and G1, while another is based on a bovine parent strain, and comprises five single-gene reassortants each containing a human-strain outer capsid gene encoding the most common human antigenic types (P[8] and G1-4).

CALICIVIRUSES

This extensive virus family, named after the cup-shaped (Latin *calyx* = cup) indentations on their viral surface seen by electron microscopy, contains four genera, two of which, the noroviruses and sapoviruses, infect humans and cause gastroenteritis. Norovirus is the major cause of acute non-bacterial gastroenteritis, causing outbreaks in nursing homes, hospitals, schools, leisure centres, restaurants and cruise ships. Transmission is mostly faeco-oral with outbreaks suggesting a common source, such as food and water and fomites. Aerosol transmission also occurs and noroviruses can be detected in vomit. Illness is usually self-limiting (12–48 hours) and mild, consisting of nausea, headache and abdominal cramps, followed by diarrhoea and vomiting, which may be the only feature (winter vomiting). Diagnosis is by demonstration of viral nucleic acid in diarrhoeal faeces. Treatment is with oral rehydration solutions (ORS). Prevention can be difficult but hand washing and good hygienic food preparation is required.

Sapovirus causes a gastroenteritis, mainly in children.

Other viruses associated with gastroenteritis are shown in Table 4.21.

Table 4.21 Viruses associated with gastroenteritis

Rotaviruses (groups A, B, C, D and E)
Enteric adenoviruses (types 40 and 41)
Caliciviruses (includes noroviruses and sapoviruses)
Astroviruses

TOGAVIRUSES

This family comprises two genera: the rubiviruses, which include rubella virus; and the alphaviruses, which include some of the arthropod-borne viruses.

Rubella

Rubella ('German measles') is caused by a spherical, enveloped RNA virus which is easily killed by heat and ultraviolet light. While the disease can occur sporadically, epidemics are not uncommon. It has a world-wide distribution. Spread of the virus is via droplets; maximum infectivity occurs before and during the time the rash is present.

Clinical features

The incubation period is 14–21 days, averaging 18 days. The clinical features are largely determined by age, with symptoms being mild or absent in children under 5 years.

During the prodrome the patient may develop malaise and fever and mild conjunctivitis and lymphadenopathy. The distribution of the lymphadenopathy is characteristic and involves particularly the suboccipital, postauricular and posterior cervical groups of lymph nodes. Small petechial lesions on the soft palate (Forchheimer spots) are suggestive but not diagnostic. Splenomegaly may be present.

The eruptive or exanthematous phase usually occurs within the first 7 days of the initial symptoms. The rash first appears on the forehead and then spreads to involve the trunk and the limbs. It is pinkish red, macular and discrete, although some of these lesions may coalesce (Fig. 4.20). It usually fades by the second day and rarely persists beyond the third day after its appearance.

Complications

Complications are rare. They include superadded pulmonary bacterial infection, arthralgia (commoner in females), haemorrhagic manifestations due to thrombocytopenia, encephalitis and the congenital rubella syndrome. Rubella affects the fetuses of up to 80% of all women who contract the infection during the first trimester of pregnancy. The incidence of congenital abnormalities diminishes in the second trimester and no ill-effects result from infection in the third trimester.

Congenital rubella syndrome is characterized by the presence of fetal cardiac malformations, especially patent ductus arteriosus and ventricular septal defect, eye lesions (especially cataracts), microcephaly, mental retardation and deafness.

The expanded rubella syndrome consists of the manifestations of the congenital rubella syndrome plus other effects including hepatosplenomegaly, myocarditis, interstitial pneumonia and metaphyseal bone lesions.

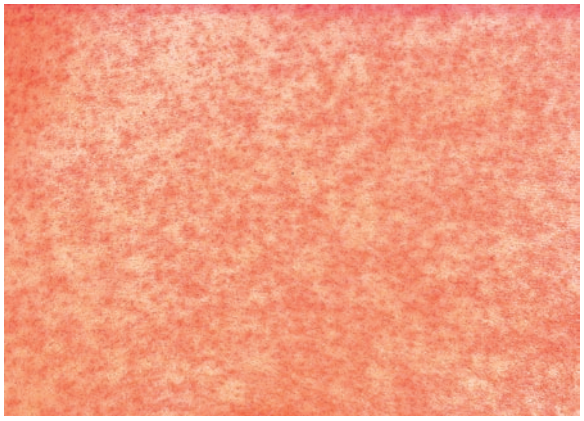


Fig. 4.20 **Rubella rash.**

Diagnosis and treatment

The diagnosis may be suspected clinically, but laboratory diagnosis is essential to distinguish the illness from other virus infections (e.g. erythrovirus B19, echovirus) and drug rashes. This is achieved by the detection of rubella-specific IgM by ELISA in an acute serum sample, preferably confirmed by the demonstration of IgG seroconversion (or a rising titre of IgG) in a subsequent sample taken 14 days later. Viral genome can be detected in throat swabs (or oral fluid samples), urine and, in the case of intrauterine infection, the products of conception.

Treatment is supportive.

Prevention

Human immunoglobulin can decrease the symptoms of this already mild illness, but does not prevent the teratogenic effects. Several live attenuated rubella vaccines have been used with great success in preventing this illness and these have been successfully combined with the measles and mumps (MMR) vaccine. The side-effects of vaccination have been dramatically decreased by using vaccines prepared in human embryonic fibroblast cultures (RA 27/3 vaccine). Use of the vaccine is contraindicated during pregnancy or if there is a likelihood of pregnancy within 3 months of immunization. Inadvertent use of the vaccine during pregnancy has not, however, revealed a risk of teratogenicity.

Arbovirus (arthropod-borne) infection

Arboviruses are zoonotic viruses, with the possible exception of the O'nyong-nyong fever virus of which humans are the only known vertebrate hosts. They are transmitted through the bites of insects, especially mosquitoes and ticks. Over 385 viruses are classified as arboviruses. *Culex*, *Aedes* and *Anopheles* mosquitoes account for the transmission of the majority of these viruses.

Although most arbovirus diseases are generally mild, epidemics are frequent and when these occur the mortality is high. In general, the incubation period is less than 10 days. The illness tends to be biphasic and, as in other viral fevers, pyrexia, conjunctival suffusion, a rash, retro-orbital pain, myalgia and arthralgia are common. Lymphadenopathy is seen in dengue. Lifelong immunity to a particular virus is usual. In some of these viral fevers, haemorrhage is a feature (Table 4.22). Increased vascular permeability, capillary fragility and consumptive coagulopathy have been implicated as causes of the haemorrhage. Encephalitis resulting from cerebral invasion may be prominent in some fevers.

Table 4.22 Viral infections associated with haemorrhagic manifestations*

Flavivirus
Yellow fever (urban and sylvan)
Dengue haemorrhagic fever
Kysanur Forest disease
Omsk haemorrhagic fever

Bunyavirus
Rift Valley fever
Congo–Crimean haemorrhagic fever
Hantavirus infections

Arenavirus
Argentinian haemorrhagic fever
Bolivian haemorrhagic fever
Lassa fever
Epidemic haemorrhagic fever

Filovirus
Marburg
Ebola

*Most of these are arboviruses. Some (e.g. hantavirus, Lassa fever) have a rodent vector. The source and transmission route of filoviruses is not known.

Alphaviruses

The 24 viruses of this group are all transmitted by mosquitoes; eight result in human disease. These viruses are globally distributed and tend to acquire their names from the location where they were first isolated (such as *Ross River*, *Eastern Venezuelan* and *Western encephalitis viruses*) or by the local expression for a major symptom caused by the virus (such as chikungunya, meaning 'doubled up'). Infection is characterized by fever, headache, maculopapular skin rash, arthralgia, myalgia and sometimes encephalitis.

Major epidemics of chikungunya were reported in India, Sri Lanka and islands in the Indian Ocean (including Reunion, Mauritius and the Seychelles) in 2005 and 2006, with at least 1 million cases and several hundred deaths. The severity of these epidemics is possibly due to a viral strain with mutations resulting in a higher neurovirulence. Several European countries reported cases of chikungunya in returning travellers – including 133 cases in the UK in 2006. The virus may now populate the local *Aedes* mosquito via increased numbers of travellers who may import virus into countries where it has not previously been described. This has resulted in outbreaks in Italy with over 150 cases and 1 death reported in 2007.

www.studentconsult.com
More online

FLAVIVIRUSES

There are 60 viruses in this group, some of which are transmitted by ticks and others by mosquitoes.

Yellow fever

Yellow fever, caused by a flavivirus, results in an illness of widely varying severity so that the disease is under-reported. It is a disease confined to Africa (90% of cases) and South America between latitudes 15°N and 15°S. For poorly understood reasons, yellow fever has not been reported from Asia, despite the fact that climatic conditions are suitable and the vector, *Aedes aegypti*, is common. The infection is transmitted in the wild by *A. africanus* in Africa and the *Haemagogus* species in South and Central America. Extension of infection to humans (via the mosquito or from monkeys) leads to the occurrence of 'jungle' yellow fever. *A. aegypti*, a domestic mosquito which lives in close relationship to humans, is responsible for human-to-human transmission in urban areas (urban yellow fever). Once infected, a mosquito remains so for its whole life.

Clinical features

The incubation period is 3–6 days. When the infection is mild, the disease is indistinguishable from other viral fevers such as influenza or dengue.

Three phases in the severe (classical) illness are recognized. Initially the patient presents with a high fever of acute onset, usually 39–40°C, which then returns to normal in 4–5 days. During this time, headache is prominent. Retrobulbar pain, myalgia, arthralgia, a flushed face and suffused conjunctivae are common. Epigastric discomfort and vomiting are present when the illness is severe. Relative bradycardia (Faget's sign) is present from the second day of illness. The patient then makes an apparent recovery and feels well for several days. Following this 'phase of calm' the patient again develops increasing fever, deepening jaundice and hepatomegaly. Ecchymosis, bleeding from the gums, haematemesis and melaena may occur. Coma, which is usually a result of uraemia or haemorrhagic shock, occurs for a few hours preceding death. The mortality rate is up to 40% in severe cases. The pathology of the liver shows mid-zone necrosis, and eosinophilic degeneration of hepatocytes (Councilman bodies) (see p. 333).

Diagnosis and treatment

The diagnosis is established by a history of travel and vaccination status, and by isolation of the virus (when possible) from blood during the first 3 days of illness. Serodiagnosis is possible, but in endemic areas cross-reactivity with other flaviviruses is a problem.

Treatment is supportive. Bed rest (under mosquito nets), analgesics, and maintenance of fluid and electrolyte balance are required.

Prevention and control

Yellow fever is an internationally notifiable disease. It is easily prevented using the attenuated 17D chick embryo vaccine but concerns over safety have arisen because of infection with the 17D virus. Vaccination is not recommended for children under 9 months and immunosuppressed patients unless there are compelling reasons. For the purposes of international certification, immunization is valid for 10 years, but protection lasts much longer than this and probably for life. The WHO Expanded Programme of Immunization includes yellow fever vaccination in endemic areas.

Dengue

This is the commonest arthropod-borne viral infection in humans: over 100 million cases occur every year in the tropics, with over 10 000 deaths from dengue haemorrhagic fever. Dengue is caused by a flavivirus and is found mainly in Asia, South America and Africa, although it has been reported from the USA and, more recently, in Italy.

Four different antigenic varieties of dengue virus are recognized and all are transmitted by the daytime-biting *A. aegypti* which breed in standing water in refuse dumps in

inner cities. *A. albopictus* is a less common transmitter. Humans are infective during the first 3 days of the illness (the viraemic stage). Mosquitoes become infective about 2 weeks after feeding on an infected individual, and remain so for the rest of their lives. The disease is usually endemic. Heterotopic immunity between serotypes after the illness is partial and lasts only a few months, although homotype immunity is life long.

Clinical features

The incubation period is 5–6 days following the mosquito bite. Asymptomatic or mild infections are common. Two clinical forms are recognized (Fig. 4.21).

Classic dengue fever

Classic dengue fever is characterized by the abrupt onset of fever, malaise, headache, facial flushing, retrobulbar pain which worsens on eye movements, conjunctival suffusion and severe backache, which is a prominent symptom. Lymphadenopathy, petechiae on the soft palate and skin rashes may also occur. The rash is transient and morbilliform. It appears on the limbs and then spreads to involve the trunk. Desquamation occurs subsequently. Cough is uncommon. The fever subsides after 3–4 days, the temperature returns to normal for a couple of days, and then the fever returns, together with the features already mentioned, but milder. This biphasic or 'saddleback' pattern is considered characteristic. Severe fatigue, a feeling of being unwell and depression are common for several weeks after the fever has subsided.

Dengue haemorrhagic fever (DHF)

Dengue haemorrhagic fever is a severe form of dengue fever and is believed to be the result of two or more sequential infections with different dengue serotypes. It is characterized by the capillary leak syndrome, thrombocytopenia, haemorrhage, hypotension and shock. It is a disease of children and has been described almost exclusively in South East Asia. The disease has a mild start, often with symptoms of an upper respiratory tract infection. This is then followed by the abrupt onset of shock and haemorrhage into the skin and ear, epistaxis, haematemesis and melaena known as the dengue shock syndrome. This has a mortality of up to 44%. Serum complement levels are depressed and there is laboratory evidence of a consumptive coagulopathy.

Diagnosis and treatment

- Isolation of dengue virus by tissue culture in sera obtained during the first few days of illness is diagnostic.
- Rising antibody titres on neutralization (most specific), haemagglutination inhibition 'ELISA' or complement-fixing antibodies in sequential serum samples.
- Blood tests show leucopenia and thrombocytopenia.

Treatment is supportive with analgesics and adequate fluid replacement. Corticosteroids are of no benefit and convalescence can be slow. In DHF, blood transfusion may be necessary.

Prevention

Travellers should be advised to sleep under impregnated nets but this is not very effective as the mosquito bites in daytime. Topical insect repellents should be used. Adult mosquitoes should be destroyed by sprays, and breeding sites, e.g. small stagnant water pools, should be eradicated. There is no effective vaccine yet although some are being trialled.

Japanese encephalitis

Japanese encephalitis is a mosquito-borne encephalitis caused by a flavivirus. It has been reported most frequently from the rice-growing countries of South East Asia and the Far East. *Culex tritaeniorhynchus* is the usual vector and this feeds mainly on pigs as well as birds such as herons and sparrows. Humans are accidental hosts.

As with other viral infections, the clinical manifestations are variable. The incubation period is 5–15 days. Most infections are asymptomatic. When disease arises, the onset is heralded by severe rigors. Fever, headache and malaise last 1–6 days. Weight loss is prominent. In the acute encephalitic stage the fever is high (38–41°C), neck rigidity occurs and neurological signs such as altered consciousness, hemipa-

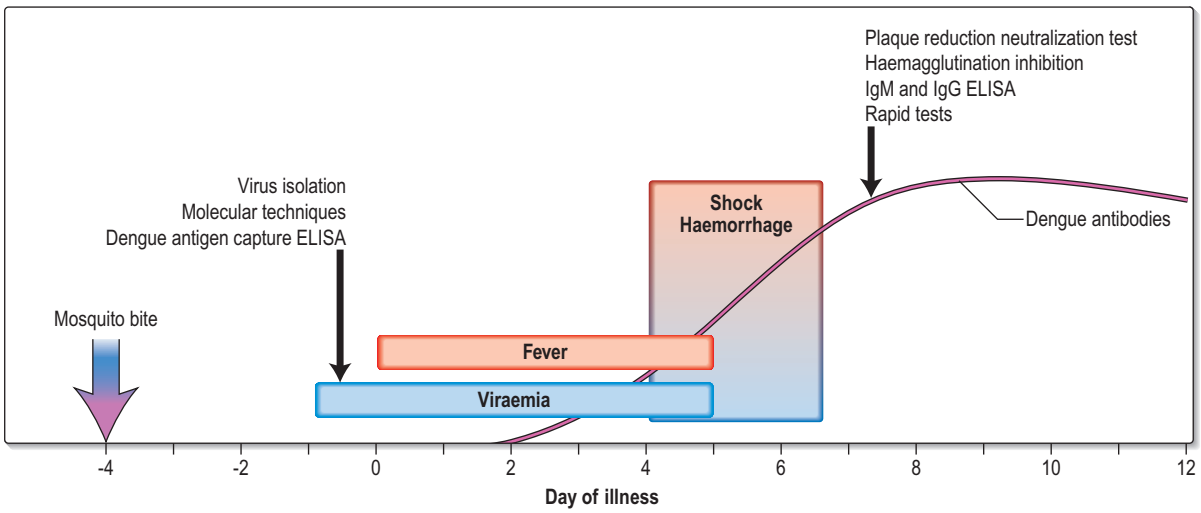
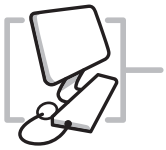


Fig. 4.21 **Dengue infection – course and timing of diagnosis.** Reprinted from Halstead SB. Dengue. *Lancet* 2007; **370:** 1644–1652, with permission from Elsevier.

resis and convulsions develop. Mental deterioration occurs over a period of 3–4 days and culminates in coma. Mortality varies from 7% to 40% and is higher in children. Residual neurological defects such as deafness, emotional lability and hemiparesis occur in about 70% of patients who have had CNS involvement. Convalescence is prolonged. Antibody detection in serum and CSF by IgM capture ELISA is a useful rapid diagnostic test. Vaccines containing formalin-inactivated viruses derived from mouse brain are effective and available. Treatment is supportive.

www.studentconsult.com
More online


West Nile virus

In 1999, the West Nile virus was first recognized in the western hemisphere (New York, USA) and it had previously been reported in Africa, Asia and parts of Europe. The outbreak in the USA produced thousands of symptomless infections. In a minority of cases, infection presents as a febrile illness, with a maculopapular rash, with 1% resulting in severe encephalitis. Disease severity and mortality is age-related. The primary hosts of infection are birds. It is spread by mosquitoes and may also infect humans and horses. It can also be transmitted by blood transfusions, breast-feeding and organ donation. Diagnosis is by genome detection in appropriate samples, or specialized serology for the detection of IgM virus-specific antibodies.

Tick-borne encephalitis virus (TBEV)

This flavivirus is transmitted by *Ixodes* spp. ticks. It occurs in an area extending from Western Europe to Japan. The tick is the main reservoir for the virus which is transmitted when it feeds on mice and other rodents.

The disease starts 4–28 days after a bite from an infected tick and is biphasic in 80% of patients. Fever, malaise, headache and fatigue are followed, after a symptom free period of about 7 days, by encephalitis. There may be associated limb paralysis which is due to anterior horn cell involvement mainly of the cervical region. Cranial nerve involvement also occurs. TBEV IgM and IgG antibodies are present and the virus can be detected in blood by RT-PCR. Mortality is about 1% but 30% have impairment in neurological function with persistent paralysis in 6%. A preventative vaccine is available.

BUNYAVIRUSES

Bunyaviruses belong to a large family of more than 200 viruses, most of which are arthropod-borne.

Congo–Crimean haemorrhagic fever

This widespread disease, caused by a virus of the nairovirus genus of Bunyaviruses, is found mainly in Asia and Africa. The primary hosts are cattle and hares and the vectors are the *Hyalomma* ticks. Following an incubation period of 3–6 days there is an influenza-like illness with fever and haemorrhagic manifestations. The mortality is 10–50%.

Hantaviruses

Hantaviruses are enzootic viruses of wild rodents which are spread by aerosolized excreta and not by insect vectors. The most severe form of this infection is Korean haemorrhagic fever (or haemorrhagic fever with renal syndrome – HFRS). This condition has a mortality of 5–10% and is characterized by fever, shock and haemorrhage followed by an oliguric phase. Milder forms of the disease are associated with related viruses (e.g. Puumala virus) and may present as nephropathia epidemica, an acute fever with renal involvement. It is seen in Scandinavia and in other European countries in people who have been in contact with bank voles. In the USA, a hantavirus (transmitted by the deer mouse) termed Sin Nombre was identified as the cause of outbreaks of acute respiratory disease in adults, referred to as hantavirus pulmonary syndrome (HPS). Other hantavirus types and rodent vector systems have been associated with this syndrome.

Diagnosis of hantavirus infection is made by an ELISA technique for specific antibodies.

Rift Valley fever

Rift Valley fever, caused by a bunyavirus, is primarily an acute febrile illness of livestock – sheep, goats and camels. It is found in southern and eastern Africa. The vector in East Africa is *Culex pipiens* and in southern Africa, *Aedes caballus,* but can be transmitted by the bite of an infected animal. Following an incubation period of 3–6 days, the patient has an acute febrile illness that is difficult to distinguish clinically from other viral fevers. The temperature pattern is usually biphasic. The initial febrile illness lasts 2–4 days and is followed by a remission and a second febrile episode. Complications are indicative of severe infection and include retinopathy, meningoencephalitis, haemorrhagic manifestations and hepatic necrosis. Mortality approaches 50% in severe forms of the illness. Treatment is supportive. Animals can be vaccinated.

ORTHOMYXOVIRUSES

Influenza

Three types of influenza virus are recognized: A, B and C, distinguishable by the nature of their internal proteins. The influenza virus is a spherical or filamentous enveloped virus. Haemagglutinin (H), a surface glycoprotein, aids attachment of the virus to the wall of susceptible host cells at specific receptor sites. Cell penetration, probably by pinocytosis, and release of replicated viruses from the cell surface effected by budding through the cell membrane, are facilitated by the action of the enzyme neuraminidase (N) which is also present on the viral envelope. Sixteen H subtypes (H1–H16) and nine N subtypes (N1–N9) have been identified for influenza A viruses but only H1, H2, H3 and N1 and N2 have established stable lineages in the human population since 1918.

- Influenza A is generally responsible for pandemics and epidemics.
- Influenza B often causes smaller or localized and milder outbreaks, such as in camps or schools.
- Influenza C rarely produces disease in humans.

Antigenic shift generates new influenza A subtypes, which emerge at irregular intervals and give rise to influenza pandemics. Possible mechanisms include:

1 Genetic reassortment of the RNA of the virus (which is arranged in eight segments) with that of an avian influenza virus; this requires co-infection of a host with

both human and avian viruses. The pig is one animal in which this may occur. Alternatively, humans may act as the mixing vessel.

2 Trans-species transmission of an avian influenza virus to humans. Viruses transmitted in this way are usually not well adapted to growth in their new host, but adaptation may occur as a result of spontaneous mutations, leading to the emergence of a pandemic strain.

Antigenic drift (minor changes in influenza A and B viruses) results from point mutations leading to amino acid changes in the two surface glycoproteins, haemagglutinin (H) and neuraminidase (N), which induce humoral immunity. This enables the virus to evade previously induced immune responses, and is the process whereby annual influenza epidemics arise.

Thus, changes due to antigenic shift or drift render the individual's immune response less able to combat the new variant.

The most serious pandemic of influenza, caused by influenza A/H1N1, occurred in 1918, and was associated with more than 20 million deaths world-wide. In 1957, antigenic shift led to the appearance of influenza A/H2N2, which caused a world-wide pandemic. A further pandemic occurred in 1968 owing to the emergence of Hong Kong influenza A/H3N2, and minor antigenic drifts have caused outbreaks around the world ever since. In 1997, avian influenza A/H5N1 viruses were first isolated from humans, raising the spectre of another pandemic. Over 300 sporadic human A/H5N1 infections have been reported since, mostly from South East Asia, and almost always arising from direct contact with infected chickens, with a mortality of >50%. However, whilst this virus is highly pathogenic to humans, due to the induction of a cytokine storm within the lungs, it still has not evolved to replicate well in human cells. Infected individuals do not therefore shed large quantities of virus from the respiratory tract, and human-to-human spread is unusual. However, anxieties remain that either genetic reassortment will occur in a human co-infected with human A/H1N1 or A/H3N2 viruses, or adaptive mutations will occur within infected human hosts, such that a truly pandemic strain will emerge.

Purified haemagglutinin and neuraminidase from recently circulating strains of influenza A and B viruses are incorporated in current vaccines.

Sporadic cases of influenza and outbreaks among groups of people living in a confined environment are frequent. The incidence increases during the winter months. Spread is mainly by droplet infection but fomites and direct contact have also been implicated.

The clinical features, diagnosis, treatment and prophylaxis of influenza are discussed on page 834.

PARAMYXOVIRUSES

These are a heterogeneous group of enveloped viruses with negative single-stranded RNA genomes of varying size that are responsible for a variety of infections.

Parainfluenza

Parainfluenza is caused by the parainfluenza viruses types I–IV; these have a world-wide distribution and cause acute respiratory disease. Type IV appears to be less virulent than the other types and has been linked only to mild upper respiratory diseases in children and adults.

Parainfluenza is essentially a disease of children and presents with features similar to the common cold. When severe, a brassy cough with inspiratory stridor and features of laryngotracheitis (croup) are present. Fever is usually present for 2–3 days and may be more prolonged if pneumonia develops. The development of croup is due to submucosal oedema and consequent airway obstruction in the subglottic region. This may lead to cyanosis, subcostal and intercostal recession and progressive airway obstruction. Infection in the immunocompromised is usually prolonged and may be severe. Treatment is supportive with oxygen, humidification and sedation when required. The role of steroids and the antiviral agent ribavirin, is controversial.

Measles (rubeola)

Measles is a highly communicable disease that occurs world-wide. With the introduction of aggressive immunization policies, the incidence of measles has fallen dramatically in the West but there are an estimated 0.5 million deaths annually due to measles infection world-wide, mostly in Africa and South East Asia, with mortality being highest in children less than 12 months of age. It is spread by droplet infection and the period of infectivity is from 4 days before until 2 days after the onset of the rash.

More online
www.studentconsult.com

Clinical features

The incubation period is 8–14 days. Two distinct phases of the disease can be recognized.

Typical measles

- The *pre-eruptive and catarrhal stage*. This is the stage of viraemia and viral dissemination. Malaise, fever, rhinorrhoea, cough, conjunctival suffusion and the pathognomonic Koplik's spots are present during this stage. Koplik's spots are small, greyish, irregular lesions surrounded by an erythematous base and are found in greatest numbers on the buccal mucous membrane opposite the second molar tooth. They occur a day or two before the onset of the rash.
- The *eruptive or exanthematous stage*. This is characterized by the presence of a maculopapular rash that initially occurs on the face, chiefly the forehead, and then spreads rapidly to involve the rest of the body (Fig. 4.22). At first the rash is discrete but later it may

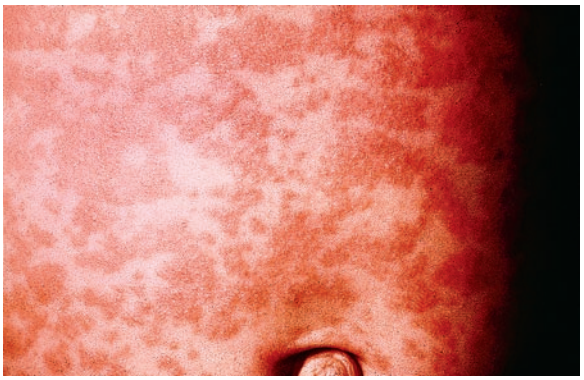


Fig. 4.22 **Measles.** Courtesy of Dr MW McKendrick, Royal Hallamshire Hospital, Sheffield.

become confluent and patchy, especially on the face and neck. It fades in about 1 week and leaves behind a brownish discoloration.

The most feared complication in an immunocompetent child is acute measles encephalitis, with an incidence of 1/1000 to 1/5000 cases of measles. This is post-infectious, i.e. virus is not present in the brain, and the encephalitis presumably arises through an aberrant cross-reaction of the host immune response to infection. Prognosis is poor, with a high mortality (30%), and severe residual damage in survivors.

Measles carries a high mortality in the malnourished and in those who have other diseases. Complications are common in such individuals and include bacterial pneumonia, bronchitis, otitis media and gastroenteritis. Less commonly, myocarditis, hepatitis and encephalomyelitis may occur. In those who are malnourished or those with defective cell-mediated immunity, the classical maculopapular rash may not develop and widespread desquamation may occur. The virus also causes the rare condition subacute sclerosing panencephalitis, which may follow measles infection occurring early in life (<18 months of age). Persistence of the virus with reactivation pre-puberty results in accumulation of virus in the brain, progressive mental deterioration and a fatal outcome (see p. 1158).

Maternal measles, unlike rubella, does not cause fetal abnormalities. It is, however, associated with spontaneous abortions and premature delivery.

Diagnosis and treatment

Most cases of measles are diagnosed clinically but, if necessary, detection of measles-specific IgM in blood or oral fluid, or genome or antigen detection from nasopharyngeal aspirates or throat swabs, can be used to confirm the diagnosis.

Treatment is supportive. Antibiotics are indicated only if secondary bacterial infection occurs.

Prevention

A previous attack of measles confers a high degree of immunity and second attacks are uncommon. Normal human immunoglobulin given within 5 days of exposure effectively aborts an attack of measles. It is indicated for previously unimmunized children below 3 years of age, during pregnancy, and in those with debilitating disease.

Active immunization. Children in the UK are immunized with the combined mumps–measles–rubella (MMR) vaccine (Box 4.6). In developing countries, measles vaccination is given at 9 months.

Mumps

Mumps is the result of infection with a paramyxovirus. It is spread by droplet infection, by direct contact or through fomites. Humans are the only known natural hosts. The peak period of infectivity is 2–3 days before the onset of the parotitis and for 3 days afterwards.

Clinical features

The incubation period averages 18 days. Although no age is exempt, it is primarily a disease of school-aged children and young adults; it is uncommon before the age of 2 years. The prodromal symptoms are non-specific and include fever, malaise, headache and anorexia. This is usually followed by severe pain over the parotid glands, with either unilateral or bilateral parotid swelling (Fig. 4.23). The enlarged parotid glands obscure the angle of the mandible and may elevate the ear lobe, which does not occur in cervical lymph node enlargement. Trismus due to pain is common at this stage. Submandibular gland involvement occurs less frequently.

Complications

CNS involvement is the most common extrasalivary gland manifestation of mumps. Clinical meningitis occurs in 5% of all infected patients, and 30% of patients with CNS involvement have no evidence of parotid gland involvement.

Epididymo-orchitis develops in about one-third of patients who develop mumps after puberty. Bilateral testicular involvement results in sterility in only a small percentage of these patients. Pancreatitis, oophoritis, myocarditis, mastitis, hepatitis and polyarthritis may also occur.

Diagnosis and treatment

The diagnosis of mumps is on the basis of the clinical features. In doubtful cases, serological demonstration of a mumps-specific IgM response in an acute blood or oral fluid sample is diagnostic. Virus can be isolated in cell culture from saliva, throat swab, urine and CSF and identified by genome detection or antigen detection by immunofluorescence.

Treatment is supportive. Attention should be given to adequate nutrition and mouth care. Analgesics should be used to relieve pain.

Prevention

Active immunization Children in the UK are immunized with the MMR vaccine (Box 4.6) and the mumps vaccine is given in most developing countries. Vaccination is contraindicated in immunosuppressed individuals, during pregnancy, or in those with severe febrile illnesses.

Fig. 4.23 **Mumps – a child with a swollen face.** From Chaudhry B, Harvey DR. *Mosby's Color Atlas and Text of Paediatrics and Child Health*. Edinburgh: Mosby, 2001. Copyright © Elsevier 2001.

Respiratory syncytial virus infection

Respiratory syncytial virus (RSV) is a paramyxovirus that causes many respiratory infections in epidemics each winter. It is a common cause of bronchiolitis in infants, which is complicated by pneumonia in approximately 10% of cases. The infection normally starts with upper respiratory symptoms. After an interval of 1–3 days a cough and low-grade fever may develop. The onset of bronchiolitis is characterized by dyspnoea and hyperexpansion of the chest with subcostal and intercostal recession. The disease may be severe and potentially fatal in babies with underlying cardiac, respiratory (including prematurity) or immunodeficiency disease. RSV infection has been associated with the occurrence of sudden infant death syndrome (SIDS). The virus undergoes antigenic drift, and, consequently, reinfection can occur throughout life. RSV is occasionally the cause of outbreaks of influenza-like illness or pneumonia in the elderly and in the immunocompromised.

Transfer of infection between children in hospital commonly occurs unless infected patients are isolated or cohorted. Meticulous attention to handwashing and other infection control measures reduces the risk of transmission by staff members.

Diagnosis and treatment

Genome detection or immunofluorescence on nasopharyngeal aspirates, virus culture and serology are the usual ways of confirming the diagnosis.

Treatment is generally supportive, but aerosolized ribavirin can be given to severe cases, particularly those with underlying cardiac or respiratory disease.

Prevention

No vaccine is currently available for RSV but high-risk children (including those with bronchopulmonary dysplasia and congenital heart disease) can be protected against severe disease by monthly administration of either a hyperimmune globulin against RSV, or a humanized monoclonal antibody (palivizumab) during the winter months (see p. 101).

Metapneumovirus

Human metapneumovirus (hMPV) is a recently discovered virus causing approximately 10% of lower respiratory tract infections in infants and young children. hMPV infection is clinically indistinguishable from that caused by RSV.

Hendra and Nipah viruses

Hendra virus (formerly called equine morbillivirus) and Nipah virus are zoonotic viruses that have caused disease in humans who have been in contact with infected animals (horses and pigs respectively). The viruses are named after the locations where they were first isolated, Hendra in Australia and Nipah in Malaysia, and both are classified as paramyxoviruses. Hendra virus has caused severe respiratory distress in horses and humans and Nipah virus caused a major outbreak of viral encephalitis (265 cases and 105 deaths) in Malaysia between September 1998 and April 1999. Treatment of these conditions is largely supportive, although there is some evidence that early treatment with ribavirin may reduce the severity of the diseases.

CORONAVIRUSES

Human coronaviruses were first isolated in the mid 1960s and the majority of isolates (related to the reference strains 229E and OC43) have been associated with common colds. In November 2002, an apparently new viral disease occurred in China (Guangdong province) and this spread rapidly in other parts of the Far East and also in Canada (Toronto and Vancouver).

This disease, known as 'severe acute respiratory syndrome' (SARS) of which bronchopneumonia has been a major feature, is caused by a previously unknown coronavirus (SARS-CoV). Similarity of this virus to coronaviruses isolated from civet cats, raccoons and ferret badgers indicates the likelihood that SARS is a zoonotic disease. Bats are the likely host species for this virus.

The epidemic was finally brought under control in the summer of 2003 and by then there had been >8000 cases with approximately 800 deaths. In 2004 and 2005, two new coronavirus infections of humans were described – NL63 and HKU1. These are associated with upper respiratory tract symptomatology, such as the common cold.

More online
www.studentconsult.com

RHABDOVIRUSES

Rabies

Rabies is a major problem in some countries, with an estimated 55 000 deaths per year world-wide. Established infection is almost invariably fatal. It is caused by a genotype 1, single-stranded RNA virus of the Lyssavirus genus. The rabies virus is bullet-shaped and has spike-like structures arising from its surface containing glycoproteins that cause the host to produce neutralizing, haemagglutination-inhibiting antibodies. The virus has a marked affinity for nervous tissue and the salivary glands. It exists in two major epidemiological settings:

- *Urban rabies* is most frequently transmitted to humans through rabid dogs and, less frequently, cats.
- *Sylvan (wild) rabies* is maintained in the wild by a host of animal reservoirs such as foxes, skunks, jackals, mongooses and bats.

With the exception of Australia, New Zealand and the Antarctic, human rabies has been reported from all continents. Transmission is usually through the bite of an infected animal. However, the percentage of rabid bites leading to clinical disease ranges from 10% (on the legs) to 80% (on the head). Other forms of transmission, if they occur, are rare.

Having entered the human body, the virus replicates in the muscle cells near the entry wound. It penetrates the nerve endings and travels in the axoplasm to the spinal cord and brain. In the CNS the virus again proliferates before spreading to the salivary glands, lungs, kidneys and other organs via the autonomic nerves.

There have been only six recorded cases of survival from clinical rabies.

Clinical features

The incubation period is variable and may range from a few weeks to several years; on average it is 1–3 months. In general, bites on the head, face and neck have a shorter incubation period than those elsewhere. In humans, two distinct clinical varieties of rabies are recognized:

More online
www.studentconsult.com

- furious rabies – the classic variety
- dumb rabies – the paralytic variety.

Furious rabies

The only characteristic feature in the prodromal period is the presence of pain and tingling at the site of the initial wound. Fever, malaise and headache are also present. About 10 days later, marked anxiety and agitation or depressive features develop. Hallucinations, bizarre behaviour and paralysis may also occur. Hyperexcitability, the hallmark of this form of rabies, is precipitated by auditory or visual stimuli. Hydrophobia (fear of water) is present in 50% of patients and is due to severe pharyngeal spasms on attempting to eat or drink. Aerophobia (fear of air) is considered pathognomonic of rabies. Examination reveals hyperreflexia, spasticity, and evidence of sympathetic overactivity indicated by pupillary dilatation and diaphoresis.

The patient goes on to develop convulsions, respiratory paralysis and cardiac arrhythmias. Death usually occurs in 10–14 days.

Dumb rabies

Dumb rabies, or paralytic rabies, presents with a symmetrical ascending paralysis resembling the Guillain–Barré syndrome. This variety of rabies commonly occurs after bites from rabid bats.

Diagnosis

The diagnosis of rabies is generally made clinically. Skin-punch biopsies are used to detect antigen with an immunofluorescent antibody test on frozen section. Viral RNA can be isolated using the reverse transcription polymerase chain reaction (RT-PCR). Isolation of viruses from saliva or the presence of antibodies in blood or CSF may establish the diagnosis. The corneal smear test is not recommended as it is unreliable. The classic Negri bodies are detected at post-mortem in 90% of all patients with rabies; these are eosinophilic, cytoplasmic, ovoid bodies, 2–10 nm in diameter, seen in greatest numbers in the neurones of the hippocampus and the cerebellum. The diagnosis should be made pathologically on the biting animal using RT-PCR, immunofluorescence assay (IFA) or tissue culture of the brain.

Treatment

Once the CNS disease is established, therapy is symptomatic as death is virtually inevitable. The patient should be nursed in a quiet, darkened room. Nutritional, respiratory and cardiovascular support may be necessary.

Drugs such as morphine, diazepam and chlorpromazine should be used liberally in patients who are excitable.

Prevention

Pre-exposure prophylaxis. This is given to individuals with a high risk of contracting rabies, such as laboratory workers, animal handlers and veterinarians. Three doses (1.0 mL) of human diploid (HDCV) or chick embryo cell vaccine given by deep subcutaneous or intramuscular injection on days 0, 7 and 28 provide effective immunity. A reinforcing dose is given after 12 months and additional reinforcing doses are given every 3–5 years depending on the risk of exposure. Vaccines of nervous-tissue origin are still used in some parts of the world. These, however, are associated with significant side-effects and are best avoided if HDCV is available.

Postexposure prophylaxis. The wound should be cleaned carefully and thoroughly with soap and water and left open. Human rabies immunoglobulin should be given immediately (20 IU/kg); half should be injected around the area of the wound and the other half should be given intramuscularly. Five 1.0 mL doses of HDCV should be given intramuscularly: the first dose is given on day 0 and is followed by injections on days 3, 7, 14 and 28. Reaction to the vaccine is uncommon.

Control of rabies

Domestic animals should be vaccinated if there is any risk of rabies in the country. In the UK, control has been by quarantine of imported animals for 6 months and no indigenous case of rabies has been reported for many years. The Pet Travel Scheme (PETS) recently introduced enables certain pet animals to enter or re-enter Great Britain without quarantine if they come from qualifying countries via designated routes, are carried by authorized transport companies, and meet the conditions of the scheme. Wild animals in 'at risk' countries must be handled with great care.

RETROVIRUSES

Retroviruses (Table 4.23) are distinguished from other RNA viruses by their ability to replicate through a DNA intermediate using an enzyme, reverse transcriptase.

HIV-1 and the related virus, HIV-2, are further classified as lentiviruses ('slow' viruses) because of their slowly progressive clinical effects.

HIV-1 and HIV-2 are discussed on page 184.

HTLV-1 causes adult T cell leukaemia/lymphoma and tropical spastic paraparesis.

ARENAVIRUSES

Arenaviruses are pleomorphic, round or oval viruses with diameters ranging from 50 to 300 nm. The virion surface has club-shaped projections, and the virus itself contains a variable number of characteristic electron-dense granules that represent residual, non-functional host ribosomes. Arenaviruses are responsible for Lassa fever and also for lymphocytic choriomeningitis, Argentinian and Bolivian haemorrhagic fevers.

Lassa fever

This illness was first documented in the town of Lassa, Nigeria, in 1969 and is confined to sub-Saharan West Africa

Table 4.23 Human lymphotropic retroviruses

Subfamily	Virus	Disease
Lentivirus	HIV-1	AIDS
	HIV-2	AIDS
Oncovirus	HTLV-1*	Adult T cell leukaemia/lymphoma
		Tropical spastic paraparesis
	HTLV-2	Myelopathy

*HTLV, human T cell lymphotropic virus.

(Nigeria, Liberia and Sierra Leone). The multimammate rat, *Mastomys natalensis*, is known to be the reservoir. Humans are infected by ingesting foods contaminated by rat urine or saliva containing the virus. Person-to-person spread by body fluids also occurs. Only 10–30% of infections are symptomatic.

Clinical features

The incubation period is 7–18 days. The disease is insidious in onset and is characterized by fever, myalgia, severe backache, malaise and headache. A transient maculopapular rash may be present. A sore throat, pharyngitis and lymphadenopathy occur in over 50% of patients. In severe cases epistaxis and gastrointestinal bleeding may occur – hence the classification of Lassa fever as a viral haemorrhagic fever. The fever usually lasts 1–3 weeks and recovery within a month of the onset of illness is usual. However, death occurs in 15–20% of hospitalized patients, usually from irreversible hypovolaemic shock.

Diagnosis

The diagnosis is established by serial serological tests (including the Lassa virus-specific IgM titre) or by genome detection by means of the reverse transcriptase polymerase chain reaction in throat swab, serum or urine.

Treatment

Treatment is supportive. In addition, clinical benefit and reduction in mortality can be achieved with ribavirin therapy, if given in the first week.

In non-endemic countries, strict isolation procedures should be used, the patient ideally being nursed in a flexible-film isolator. Specialized units for the management of Lassa fever and other haemorrhagic fevers have been established in the UK. As Lassa fever virus and other causes of haemorrhagic fever (Marburg/Ebola and Congo–Crimean haemorrhagic fever viruses) have been transmitted from patients to staff in healthcare situations, great care should be taken in handling specimens and clinical material from these patients.

Lymphocytic choriomeningitis (LCM)

This infection is a zoonosis, the natural reservoir of LCM virus being the house mouse. Infection is characterized by:

- non-nervous-system illness, with fever, malaise, myalgia, headache, arthralgia and vomiting
- aseptic meningitis in addition to the above symptoms.

Occasionally, a more severe form occurs, with encephalitis leading to disturbance of consciousness.

This illness is generally self-limiting and requires no specific treatment.

FILOVIRUSES

MARBURG VIRUS DISEASE AND EBOLA VIRUS DISEASE

These severe, haemorrhagic, febrile illnesses are discussed together because their clinical manifestations are similar. The diseases are named after Marburg in Germany and the Ebola River region in the Sudan and Zaire where these viruses first appeared. The natural reservoir for these viruses has not been identified and the precise mode of spread from one individual to another has not been elucidated.

Epidemics have occurred periodically in recent years, mainly in sub-Saharan Africa. The mortality from Marburg and Ebola has ranged from 25% to 90% and recovery is slow in those who survive.

The illness is characterized by the acute onset of severe headache, severe myalgia and high fever, followed by prostration. On about the fifth day of illness a non-pruritic maculopapular rash develops on the face and then spreads to the rest of the body. Diarrhoea is profuse and is associated with abdominal cramps and vomiting. Haematemesis, melaena or haemoptysis may occur between the seventh and sixteenth day. Hepatosplenomegaly and facial oedema are usually present. In Ebola virus disease, chest pain and a dry cough are prominent symptoms.

Treatment is symptomatic. Convalescent human serum appears to decrease the severity of the attack.

POSTVIRAL/CHRONIC FATIGUE SYNDROME (see also p. 1192)

Viral illnesses have been implicated aetiologically, including those due to EBV, Coxsackie B viruses, echoviruses, CMV and hepatitis A virus. Non-viral causes such as allergy to *Candida* spp. have also been proposed. Only a minority of patients have an identifiable precipitating infectious illness.

TRANSMISSIBLE SPONGIFORM ENCEPHALOPATHIES (TSE OR PRION DISEASE)

Transmissible spongiform encephalopathies are caused by the accumulation in the nervous system of a protein, termed a 'prion', which is an abnormal isoform PrP^{res} of a normal, host protein (PrP^{c}).

Although familial forms of prion disease are known to exist, these conditions can be transmissible, particularly if brain tissue enters another host. There is no convincing evidence for the presence of nucleic acid in association with prions; thus these agents cannot be considered orthodox viruses and it is the abnormal prion protein itself that is infectious and can trigger a conversion of the normal protein into the atypical isoform. After infection, a long incubation period is followed by CNS degeneration associated with dementia or ataxia which invariably leads to death. Histology of the brain reveals spongiform change with an accumulation of the abnormal prion protein in the form of amyloid plaques.

The human prion diseases are Creutzfeldt–Jakob disease, including the sporadic, familial, iatrogenic and variant forms of the disease, Gerstmann–Straussler–Scheinker syndrome, fatal familial insomnia and kuru.

- *Creutzfeldt–Jakob disease (CJD)* usually occurs sporadically world-wide with an annual incidence of one per million of the population. Although, in most cases, the epidemiology remains obscure, transmission to others has occurred as a result of administration of human cadaveric growth hormone or gonadotrophin,

Ask the authors
www.studentconsult.com

from dura mater and corneal grafting, and in neurosurgery from reuse of contaminated instruments and electrodes (iatrogenic CJD).

- *Variant CJD*. In the UK, knowledge that large numbers of cattle with the prion disease, bovine spongiform encephalopathy (BSE) had gone into the human food chain, led to enhanced surveillance for emergence of the disease in humans. The evidence is convincing, based on transmission studies in mice and on glycosylation patterns of prion proteins, that this has occurred and, to date, there have been approximately 160 confirmed and suspected cases of variant CJD (human BSE) in the UK and 40 in the rest of the world. In contrast to sporadic CJD, which presents with dementia at a mean age of onset of 60 years, variant CJD presents with ataxia, dementia, myoclonus and chorea at a mean age of onset of 29 years. The epidemic curve of vCJD in the UK is shown in Figure 4.24.
- *Gerstmann–Straussler–Scheinker syndrome* and *fatal familial insomnia* are rare prion diseases usually occurring in families with a positive history. The pattern of inheritance is autosomal dominant with some degree of variable penetrance.
- *Kuru* was described and characterized in the Fore highlanders in NE New Guinea. Transmission was associated with ritualistic cannibalism of deceased relatives. With the cessation of cannibalism by 1960, the disease has gradually diminished and recent cases had all been exposed to the agent before 1960.

The infectious agents of prion disease have remarkable characteristics. In the infected host there is no evidence of inflammatory, cytokine or immune reactions. The agent is highly resistant to decontamination, and infectivity is not reliably destroyed by autoclaving or by treatment with formaldehyde and most other gas or liquid disinfectants. It is very resistant to γ irradiation. Autoclaving at a high temperature (134–137°C for 18 minutes) is used for decontamination of instruments, and hypochlorite (20 000 p.p.m. available chlorine) or 1 molar sodium hydroxide are used for liquid disinfection. Uncertainty about the reliability of any methods for safe decontamination of surgical instruments has necessitated the introduction of guidelines for patient management.

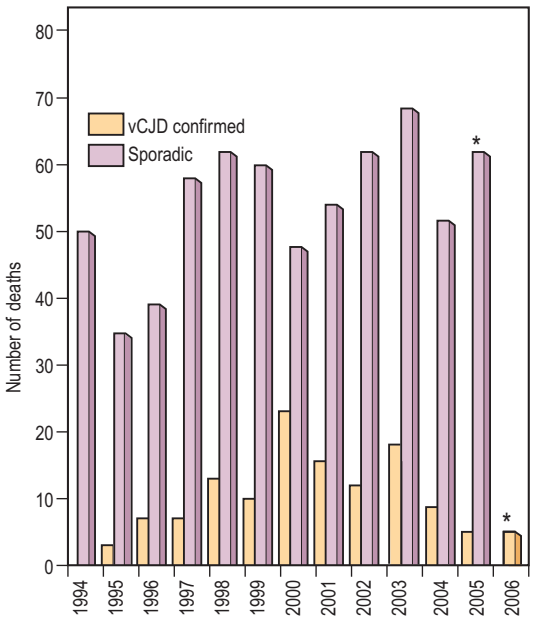


Fig. 4.24 **Creutzfeldt–Jakob disease.** Deaths of confirmed and sporadic cases of CJD (* to early 2006) in the UK. Courtesy of Professor JW Ironside, Director National CJD Surveillance Unit, University of Edinburgh.

FURTHER READING

Asaria P, MacMahon E. Measles in the United Kingdom: can we eradicate it by 2010? *British Medical Journal* 2006; **333**: 890–895.

Banatvala JE, Brown DWG. Rubella. *Lancet* 2004; **363**: 1127–1136.

Bodenmann P, Genton B. Chikungunya: an epidemic in real time. *Lancet* 2006; **368**: 258.

Collins ST, Lawson VA, Masters CL. Transmissible spongiform encephalopathies. *Lancet* 2004; **363**: 51–61.

Grimwood K, Buttery JP. Clinical update: rotavirus gastroenteritis and its prevention. *Lancet* 2007; **370**: 302–304.

Halstead SB. Dengue. *Lancet* 2007; **370**: 1644–1652.

Hviid A, Rubin S, Muhlemann K. Mumps. *Lancet* 2008; **371**: 932–944.

Lindquist L, Vapalahti O. Tick-borne encephalitis. *Lancet* 2008; **371**: 1861–1871.

Peiris JSM, Yuen KY, Osterhuis ADME et al. The severe acute respiratory syndrome. *New England Journal of Medicine* 2003; **349**: 2431–2441.

Warrell MJ, Warrell DA. Rabies and other lyssavirus diseases. *Lancet* 2004; **363**: 959–969.

Writing Committee of the Second World Health Organization Consultation on Clinical Aspects of Human Infection with Avian Influenza A (H5N1) Virus. *New England Journal of Medicine* 2008; **358**: 261–273.

BACTERIAL INFECTIONS

Classification of bacteria

Bacteria are unicellular organisms (prokaryotes), of which a small fraction are of medical relevance. Unusual infections may result from exposure under circumstances of altered host defences, notably in the severely immunocompromised patient.

Bacteria have traditionally been classified according to the Gram stain which distinguishes Gram-positive from Gram-negative organisms. Using light microscopy, these can then largely be divided into cocci and bacilli (rods). Some have a spiral appearance (spirochaetes) while others, such as *Clostridium* spp., may contain spores (Table 4.24). The cell wall arrangement of Gram-positive cocci contains a phospholipid bilayer surrounded by peptidoglycan made up of repeating units of *N*-acetylglucosamine and *N*-acetylmuramic acid. By contrast, Gram-negative bacilli possess a second outer lipid bilayer containing protein and lipopolysaccharide (endotoxin). Some pathogens are encapsulated, which is an antiphagocytic virulence factor.

Bacteria can often be cultured in broth or on solid agar. Those growing in the absence of oxygen are strict anaerobes (e.g. *Bacteroides* spp.), whilst oxygen-dependent bacteria are known as aerobes (e.g. *Pseudomonas* spp.). Many pathogens can tolerate reduced concentrations of oxygen (e.g. *E. coli*). Some organisms are more demanding in their growth requirements and require special laboratory media (e.g. *Mycoplasma* spp. and *Mycobacterium* spp.); others require more prolonged incubation (e.g. *Brucella* spp.).

Table 4.24 Classification of bacteria affecting humans

Bacteria	Cocci	Bacilli	Others
Aerobic			
Gram-positive	*Staphylococcus aureus* *Staphylococcus epidermidis* *Streptococcus pneumoniae* *Strep. pyogenes* (group A) *Strep. agalactiae* (group B) Enterococci *Viridans streptococci*	*Listeria monocytogenes* *Corynebacterium diphtheriae* *Bacillus anthracis* *B. cereus*	
Gram-negative	*Neisseria gonorrhoeae* *N. meningitidis* *Moraxella catarrhalis* *Bordetella pertussis*	*Escherichia coli* *Klebsiella* spp. *Proteus* spp. *Haemophilus influenzae* *Legionella* spp. *Salmonella* spp. *Shigella* spp. *Campylobacter jejuni* *Helicobacter pylori* *Pseudomonas* spp. *Brucella* spp. *Acinetobacter* spp. *Burkholderia* spp. *Vibrio cholerae* *Yersinia pestis*	
Anaerobic			
Gram-positive	Peptococci Peptostreptococci	*Actinomyces* spp. *Clostridium perfringens* *C. difficile* *C. botulinum* *C. tetani*	
Gram-negative		*Bacteroides fragilis* group *Fusobacterium* spp.	
Spirochaetes			*Treponema pallidum* *Leptospira* spp. *Borrelia* spp.
Others			*Mycobacterium* spp. *Mycoplasma moniae* *Ureaplasma* spp. *Chlamydia* spp.

Genetic classification is defining bacteria in terms of DNA sequence information and has led to the reclassification of several bacteria. DNA fingerprinting is also being increasingly applied to distinguishing similar isolates, which has applications in defining the epidemiology of infection.

Diagnosis and management of bacterial infections

The history and examination usually localizes the infection to a specific organ or body site. A systemic response may accompany such localized disease or, in the case of bloodstream infections, be the primary mode of presentation. The microbiological diagnosis is difficult to establish in most community-managed infections, and even in hospital where there is ready access to diagnostic laboratories, only a minority of infections are documented. For these reasons, a clinical approach to bacterial diseases has been adopted.

SKIN AND SOFT TISSUE INFECTION

Superficial infections

Infections of the skin and the soft tissues beneath are common. These are usually fungal (see p. 1232) or bacterial. Although a wide range of bacteria have been recovered from skin and soft tissue infections (Table 4.25), the majority are caused by the Gram-positive cocci *Staphylococcus aureus* and *Streptococcus pyogenes*.

Staphylococcus aureus

Staphylococci are part of the normal microflora of the human skin and nasopharynx; up to 25% of people are carriers of *S. aureus*, which is the species responsible for the majority of staphylococcal infections. Although soft tissue infections are the most common manifestation of *S. aureus* disease, numerous other sites can be affected (Table 4.6).

The classification of soft tissue infections is complex, because imprecise and overlapping terms are in use. The

Table 4.25 Bacterial causes of superficial skin and soft tissue infection

Specific risk factors	Likely organisms
None	*Staphylococcus aureus* *Streptococcus pyogenes*
Diabetes, peripheral vascular disease	Group B streptococci
Animal bite	*Pasteurella multocida*, *Capnocytophaga canimorsus*
Fresh water exposure	*Aeromonas hydrophila*
Sea water exposure	*Vibrio vulnificans*
Lymphoedema, stasis dermatitis	Groups A, C and G streptococci
Hot tub exposure	*Pseudomonas aeruginosa*
Malignant otitis externa	*Pseudomonas aeruginosa*
Human bite	*Fusobacterium* spp.

Table 4.26 Predisposing factors for skin and soft tissue infection

Diabetes mellitus
Chronic lymphoedema
Peripheral vascular disease
Venous ulcers
Steroid treatment
Malnutrition
Some immunodeficiency states (e.g. Job's syndrome)
Nasal carriage of *Staphylococcus aureus*

commonly encountered infections are described in more detail on page 1228.

The majority of skin and superficial soft tissue infections are due to bacteria on the skin surface penetrating the dermis or the subcutaneous tissues. Infection can take place via hair follicles, insect bites, cuts and abrasions, or skin damaged by superficial fungal infection. Sometimes infection is introduced by an animal bite or a penetrating foreign body: in these cases more unusual organisms may be found. A number of factors predispose to cellulitis and other soft tissue infections (Table 4.26).

Invasive staphylococcal infection is often associated with breaches in the skin, for example due to injecting drug use, iatrogenic cannulation, surgery or trauma. In clinical situations scrupulous attention to disinfection and hygiene when performing invasive procedures can minimize the risk of infection. Although *S. aureus* is the most common species of staphylococci implicated in catheter-related infections, other normally non-pathogenic species such as *S. epidermidis* (which is often intrinsically resistant to flucloxacillin) may be found. Flucloxacillin remains the first choice in staphylococcal infection when the organism is known to be sensitive, but with the increasing prevalence of meticillin-resistant *S. aureus* (MRSA), other agents are often needed.

S. aureus can produce a variety of toxins and virulence factors which affect the type and severity of infection. These include staphylococcal enterotoxin A, superantigenic staphylococcal exotoxins, toxic shock toxin 1 and Panton Valentine leucocidin (PVL). The last of these has been found mainly in community-associated strains of *S. aureus* (both MRSA and MSSA), rather than in hospital-acquired or epidemic strains. Community-associated PVL-producing MRSA and MSSA are becoming an increasingly common cause of invasive soft tissue and lung infections in some countries (notably the USA), although it is unclear whether PVL itself is directly responsible for the increased virulence.

Meticillin-resistant Staphylococcus aureus *(MRSA)*

S. aureus are commonly resistant to penicillin, and isolated resistance to other β-lactam antibiotics such as meticillin (now rarely used) and flucloxacillin has been recognized since the development of the first semisynthetic penicillins in the early 1960s. However, in the last 40 years, strains of MRSA with resistance to a much wider range of antibiotics have emerged. In some cases only the glycopeptide antibiotics, vancomycin and teicoplanin, are effective (along with the new agents discussed below), and a few organisms have been isolated with decreased sensitivity even to these.

Vancomycin-insensitive *S. aureus* (VISA) develop because the organism produces a thick cell wall by changing the synthesis of cell wall material. Vancomycin-resistant *S. aureus* (VRSA) acquires resistance by receiving the van A gene from vancomycin-resistant enterococci (see Fig. 4.4).

Apart from the glycopeptides, four *newer classes of antibiotics* are effective against Gram-positive bacteria, including MRSA: these are the streptogramins (e.g. quinupristin with dalfopristin), the oxazolidinones (e.g. linezolid), tigecycline and daptomycin. They should usually be reserved for multiresistant organisms. Control of the use of antibiotics in hospitals and good infection control policies are vital to prevent the spread of multiresistant organisms. Scrupulous hygiene on the part of healthcare workers is especially necessary (e.g. hand washing).

MRSA is usually found as a skin commensal, especially in hospitalized patients or nursing home residents. However it can cause a variety of infections in soft tissues and elsewhere, and can cause death. It is particularly associated with surgical wound infections. Eradication of the organism is difficult, and people who are known to be colonized should be isolated from those at risk of significant infection. Topical treatment with antibiotics is often used, but is of limited efficacy. Hand-washing is more effective at controlling spread.

Although MRSA is generally regarded as a hospital-associated organism it is now commonly seen in people away from the hospital setting, both as a colonizer and as a cause of disease. Often the organisms are the typical 'hospital' strains of MRSA, and have been acquired directly or indirectly from a healthcare setting (e.g. in workers in care homes). However there is an increasing prevalence in some countries of true community-associated MRSA (CA-MRSA), with no discernible links to hospital or residential care.These CA-MRSA have different resistance profiles to typical hospital strains (often retaining sensitivity to tetracyclines, clindamycin and co-trimoxazole), and are more likely to produce PVL.

Detection

Culture takes 24 hours. A new rapid (2 h) real time PCR assay is now available.

Pasteurellosis

Pasteurella multocida is found in the oropharynx of up to 90% of cats and 70% of dogs. It can cause soft tissue infec-

tions following animal bites. Although the infection initially resembles other forms of cellulitis, there is a much higher incidence of spread to deeper tissues, resulting in osteomyelitis, tenosynovitis or septic arthritis. The organism is sensitive to penicillin, but as infections following animal bites are often polymicrobial, co-amoxiclav is used.

Cat scratch disease

Cat scratch disease is a zoonosis caused by *Bartonella henselae*. Asymptomatic bacteraemia is relatively common in domestic and especially feral cats, and human infection is probably due to direct inoculation from the claws or via cat flea bites. Regional lymphadenopathy appears 1–2 weeks after infection; the nodes become tender and may suppurate. Histology of the nodes shows granuloma formation, and the illness may be mistaken for mycobacterial infection or lymphoma. There are usually few systemic symptoms in immunocompetent patients, although more severe disease may be seen in the immunocompromised. In these patients tender cutaneous or subcutaneous nodules are seen (*bacillary angiomatosis*) which may ulcerate. The lymphadenopathy resolves spontaneously over weeks or months, although surgical drainage of very large suppurating nodes may be necessary. *B. henselae* is sensitive to azithromycin, doxycycline and ciprofloxacin, but drug selection and clinical benefit of treatment is variable according to the primary site of the infection.

Toxin-mediated skin disease

A number of skin conditions, although caused by bacteria, are mediated by exotoxins rather than direct local tissue damage.

Staphylococcal scalded skin syndrome

The scalded skin syndrome is caused by a toxin-secreting strain of *S. aureus*. It principally affects children under the age of 5. The toxin, exfoliatin, causes intra-epidermal cleavage at the level of the stratum corneum leading to the formation of large flaccid blisters that shear readily. It is a relatively benign condition, and responds to treatment with flucloxacillin.

Toxic shock syndrome (TSS)

TSS is usually due to toxin-secreting staphylococci, but toxin-secreting streptococci have also been implicated. Although historically associated with vaginal colonization and tampon use in women, this is not always the case. The exotoxin (normally toxic shock syndrome toxin 1, TSST-1) causes cytokine release with abrupt onset of fever and shock, with a diffuse macular rash and desquamation of the palms and soles. Many patients are severely ill and mortality is about 5%. Treatment is mainly supportive, although the organism should be eradicated.

Scarlet fever

See page 126.

Deep soft tissue infections

Infections of the deeper soft tissues are much less common than superficial infections, and tend to be more serious. Usually they are related to penetrating injuries or to surgery, and the causative organisms relate to the nature of the wound.

Necrotizing fasciitis

Necrotizing fasciitis is a fulminant, rapidly spreading infection associated with widespread tissue destruction (through all tissue planes) and a high mortality. Historically two forms are described. Type 1, caused by a mixture of aerobic and anaerobic bacteria, is usually seen following abdominal surgery or in diabetics. Type 2, caused by group A streptococci, arises spontaneously in previously healthy people. Other organisms are now also recognized as causing necrotising fasciitis, most notably *Vibrio* species (*V. vulnificans*), associated with sea water in the tropics. All types are characterized by severe pain at the site of initial infection, rapidly followed by tissue necrosis. Infection tracks rapidly along the tissue planes, causing spreading erythema, pain and sometimes crepitus. In patients with fever, toxicity and pain, which is out of proportion to the skin findings, necrotizing fasciitis should be suspected and must be treated aggressively and promptly with antibiotics and urgent surgical exploration with extensive debridement or amputation if necessary. Laboratory investigations show a high CRP, a very raised white count (often over 25×10^9/L) and hyponatraemia. Imaging with US/CT may be helpful but should not deter urgent surgical exploration. Multiorgan failure is common and mortality is high. Type 2 necrotizing fasciitis is treated with high doses of benzylpenicillin and clindamycin; type 1 is treated with a broad-spectrum combination, which should include metronidazole.

Gas gangrene

Gas gangrene is caused by deep tissue infection with *Clostridium* spp., especially *C. perfringens*, and follows contaminated penetrating injuries. It is particularly associated with battlefield wounds, but is also seen in intravenous drug users, and following abdominal surgery. The initial infection develops in an area of necrotic tissue caused by the original injury; toxins secreted by the bacteria kill surrounding tissue and enable the anaerobic organism to spread rapidly. Toxins are also responsible for the severe systemic features of gas gangrene. Treatment consists of urgent surgical removal of necrotic tissue, and treatment with benzylpenicillin and clindamycin.

RESPIRATORY TRACT INFECTIONS

Infections of the respiratory tract are divided into infections of the upper and lower respiratory tract, which are separated by the larynx. In health, the lower respiratory tract is normally sterile owing to a highly efficient defence system (see p. 818). Infections of the upper respiratory tract are particularly common in childhood when they are usually the result of virus infection. The paranasal sinuses and middle ear are contiguous structures and can be involved secondary to viral infections of the nasopharynx. The lower respiratory tract is frequently compromised by smoking, air pollution, aspiration of upper respiratory tract secretions and chronic lung disease, notably chronic obstructive pulmonary disease. Infections of the respiratory tract are defined clinically, sometimes radiologically, as in the case of pneumonia, and by appropriate microbiological sampling.

Upper respiratory tract infections

- The common cold (acute coryza) (see p. 831).
- Sinusitis (see p. 1081).

FURTHER READING

Anany DA, Patchen Dellinger E. Necrotizing soft tissue infection: diagnosis and management. *Clinical Infectious Diseases* 2007; **44**: 705–710.

Diep BA, Carleton HA, Chang RF et al. Roles of 34 virulence genes in the evolution of hospital and community-associated strains of MRSA. *Journal of Infectious Diseases* 2006; **193**: 1495–1503.

Huskins WC, Goldmann DA. Controlling meticillin-resistant *Staph aureus*. *Lancet* 2005; **365**: 273–275.

- Rhinitis (see p. 831).
- Pharyngitis (see p. 833).

Scarlet fever

Scarlet fever occurs when the infectious organism (usually a group A streptococcus) produces erythrogenic toxin in an individual who does not possess neutralizing antitoxin antibodies. Infections may be sporadic or epidemic occuring in residential institutions such as schools, prisons and military establishments.

Clinical features

The incubation period of this relatively mild disease, which mainly affects children, is 2–4 days following a streptococcal infection, usually in the pharynx. Regional lymphadenopathy, fever, rigors, headache and vomiting are present. The rash, which blanches on pressure, usually appears on the second day of illness; it initially occurs on the neck but rapidly becomes punctate, erythematous and generalized. It is typically absent from the face, palms and soles, and is prominent in the flexures. The rash usually lasts about 5 days and is followed by extensive desquamation of the skin (Fig. 4.25). The face is flushed, with characteristic circumoral pallor. Early in the disease the tongue has a white coating through which prominent bright red papillae can be seen ('strawberry tongue'). Later the white coating disappears, leaving a raw-looking, bright red colour ('raspberry tongue'). The patient is infective for 10–21 days after the onset of the rash, unless treated with penicillin.

Scarlet fever may be complicated by peritonsillar or retropharyngeal abscesses and otitis media.

Diagnosis

The diagnosis is established by the typical clinical features and culture of a throat swab. Elevated antistreptolysin O and anti-DNase B levels in convalescent serum are indicative of streptococcal infection.

Treatment

Penicillin is the drug of choice and is given orally as phenoxymethylpenicillin 500 mg four times daily for 10 days. Individuals allergic to penicillin can be treated effectively with erythromycin 250 mg four times daily for 10 days. Treatment is usually effective in preventing rheumatic fever (see p. 136) and acute glomerulonephritis (see p. 592), which are non-suppurative complications of streptococcal pharyngitis. Unlike acute rheumatic fever, streptococcal nephritis may also complicate streptococcal skin infection.

Fig. 4.25 **Scarlet fever rash, showing desquamation.**

Prevention

Chemoprophylaxis with penicillin or erythromycin should be given in epidemics.

Diphtheria

Diphtheria caused by *Corynebacterium diphtheriae* occurs world-wide. Its incidence in the West has fallen dramatically following widespread active immunization, but has re-emerged in the independent states of the former Soviet Union. Transmission is mainly through airborne droplet infection and rarely through fomites.

C. diphtheriae is a Gram-positive bacillus. Only strains which carry the tox+ gene, are capable of toxin production. The toxin has two subunits, A and B. Subunit A is responsible for clinical toxicity. Subunit B serves only to transport the toxin component to specific receptors present chiefly on the myocardium and in the peripheral nervous system. Humans are the only natural hosts.

Clinical features

Diphtheria was formerly a disease of childhood but may affect adults in countries where childhood immunization has been interrupted or uptake is poor. The incubation period is 2–7 days. The manifestations may be regarded as local (due to the membrane) or systemic (due to exotoxin). The presence of a membrane, however, is not essential to the diagnosis. The illness is insidious in onset and is associated with tachycardia but only low-grade fever. If complicated by infection with other bacteria such as *Strep. pyogenes*, fever is high and spiking.

Nasal diphtheria is characterized by the presence of a unilateral, serosanguineous nasal discharge that crusts around the external nares.

Pharyngeal diphtheria is associated with the greatest toxicity and is characterized by marked tonsillar and pharyngeal inflammation and the presence of a membrane. This tough greyish yellow membrane is formed by fibrin, bacteria, epithelial cells, mononuclear cells and polymorphs, and is firmly adherent to the underlying tissue. Regional lymphadenopathy, often tender, is prominent and produces the so-called 'bull-neck'.

Laryngeal diphtheria is usually a result of extension of the membrane from the pharynx. A husky voice, a brassy cough, and later dyspnoea and cyanosis due to respiratory obstruction are common features.

Clinically evident myocarditis occurs, often weeks later, in patients with pharyngeal or laryngeal diphtheria. Acute circulatory failure due to myocarditis may occur in convalescent individuals around the tenth day of illness and is usually fatal. Neurological manifestations occur either early in the disease (palatal and pharyngeal wall paralysis) or several weeks after its onset (cranial nerve palsies, paraesthesiae, polyneuropathy or, rarely, encephalitis).

Cutaneous diphtheria is uncommon but seen in association with burns and in individuals with poor personal hygiene. Typically the ulcer is punched-out with undermined edges and is covered with a greyish white to brownish adherent membrane. Constitutional symptoms are uncommon.

Diagnosis

This must be made on clinical grounds since therapy is usually urgent and bacteriological results of culture studies and toxin production cannot be awaited.

Treatment

The patient should be isolated and bed rest advised. Antitoxin therapy is the only specific treatment. It must be given promptly to prevent further fixation of toxin to tissue receptors, since fixed toxin is not neutralized by antitoxin. Depending on the severity, 20 000–100 000 units of horse-serum antitoxin should be administered intramuscularly after an initial test dose to exclude any allergic reaction. Intravenous therapy may be required in a very severe case. There is a risk of acute anaphylaxis after antitoxin administration and of serum sickness 2–3 weeks later (Box 4.7). However, the risk of death outweighs the problems of anaphylaxis. Antibiotics should be administered concurrently to eliminate the organisms and thereby remove the source of toxin production. Benzylpenicillin 1.2 g four times daily is given for 1 week.

The cardiac and neurological complications need intensive therapy. Recovery and rehabilitation can take many weeks.

Prevention

Diphtheria is prevented by active immunization in childhood (see p. 102). Booster doses should be given to those travelling to endemic areas if more than 10 years has elapsed following their primary course of immunization. All contacts of the patient should have throat swabs sent for culture; those with a positive result should be treated with penicillin or erythromycin and active immunization or a booster dose of diphtheria toxoid given.

Pertussis (whooping cough)

Pertussis occurs world-wide. Humans are both the natural hosts and reservoirs of infection. The disease is caused by *Bordetella pertussis* which is a Gram-negative coccobacillus. *B. parapertussis* and *B. bronchiseptica* produce milder infections. Pertussis is highly contagious and is spread by droplet infection. In its early stages it is indistinguishable from other types of upper respiratory tract infection. Epidemic disease occurred in the UK when the safety of the whooping cough vaccine was questioned. Currently, uptake exceeds 95% and the disease is uncommon.

Clinical features

The incubation period is 7–10 days. It is a disease of childhood, with 90% of cases occurring below 5 years of age. However, no age is exempt although in adults it may be unsuspected.

During the catarrhal stage the patient is highly infectious, and cultures from respiratory secretions are positive in over 90% of patients. Malaise, anorexia, mucoid rhinorrhoea and conjunctivitis are present. The paroxysmal stage, so called because of the characteristic paroxysms of coughing, begins about a week later. Paroxysms with the classic inspiratory whoop are seen only in younger individuals in whom the lumen of the respiratory tract is compromised by mucus secretion and mucosal oedema. The whoop results from air being forcefully drawn through the narrowed tract. These paroxysms usually terminate in vomiting. Conjunctival suffusion and petechiae and ulceration of the frenulum of the tongue are usual. Lymphocytosis due to the elaboration of a lymphocyte-promoting factor by *B. pertussis* is characteristic; lymphocytes may account for over 90% of the total white blood cell count. This stage lasts approximately 2 weeks and may be associated with several complications, including pneumonia, atelectasis, rectal prolapse and inguinal hernia. Cerebral anoxia may occur, especially in younger children, resulting in convulsions. Bronchiectasis is a rare sequel.

Box 4.7 Antitoxin administration

- Many antitoxins are heterologous and therefore dangerous
- Hypersensitivity reactions are common

Prior to treatment

Question patient about:
(a) allergic conditions (e.g. asthma, hay fever)
(b) previous antitoxin administration.

Read instructions on antitoxin package carefully.
Always give a subcutaneous test dose.
Have adrenaline (epinephrine) available.

Diagnosis

The diagnosis is suggested clinically by the characteristic whoop and a history of contact with an infected individual. It is confirmed by isolation of the organism. Cultures of swabs (dacron or calcium alginate) of nasopharyngeal secretions result in a higher positive yield than cultures of 'cough plates'; PCR assays are available.

Treatment

If the disease is recognized in the catarrhal stage, erythromycin will abort or decrease the severity of the infection. In the paroxysmal stage, antibiotics have little role to play in altering the course of the illness.

Prevention and control

Affected individuals should be isolated to prevent contact with others, e.g. in hostels and boarding schools. Pertussis is an easily preventable disease and effective active immunization is available (Box 4.6). Convulsions and encephalopathy have been reported as rare complications of vaccination but they are probably less frequent than after whooping cough itself. Any exposed susceptible infant should receive prophylactic erythromycin.

Acute epiglottitis (see p. 834)

This has been virtually eliminated among children in those countries which have introduced *Haemophilus influenzae* vaccine, as in the UK. Occasionally, infections are being recognized in adults. The clinical features are described on page 834.

Acute laryngotracheobronchitis (CROUP)

See page 833.

Influenza

See pages 116 and 834.

Lower respiratory tract infections

Pneumonia: community-acquired (see p. 859); hospital-acquired (see p. 862); in immunocompromised persons (see p. 189).

Psittacosis (ornithosis)

Although originally thought to be limited to the psittacine birds (parrots, parakeets and macaws), it is known that the disease is widely spread amongst many species of birds,

including pigeons, turkeys, ducks and chickens (hence the broader term 'ornithosis'). Human infection is related to exposure to infected birds and is therefore a true zoonosis. The causative organism, *Chlamydia psittaci*, is excreted in avian secretions; it can be isolated for prolonged periods from birds who have apparently recovered from infection. The organism gains entry to the human host by inhalation. For *clinical features and treatment* – see page 860.

Acinetobacter infection

This Gram-negative coccobacillus is becoming increasingly prominent in nosocomical infections, particularly as a cause of ventilator-associated pneumonia (see p. 914) and vascular catheter infections. It is a cause of community acquired infections in tropical countries and is associated with wars and natural disasters. The organism is resistant to many antibiotics including carbapenems. Polymyxin and tigecycline are being used but resistance is still a problem.

Other respiratory infections

(see also p. 860)

Chlamydia pneumoniae causes a relatively mild pneumonia in young adults, clinically resembling infection caused by *Mycoplasma pneumoniae*. Diagnosis can be confirmed by specific IgM serology. Treatment is with erythromycin 500 mg 6-hourly, tetracycline 500 mg every 6–8 hours, or a fluoroquinolone, e.g. ciprofloxacin 500 mg twice daily.

Other chlamydial infections include trachoma (see p. 143), lymphogranuloma venereum (see p. 177) and other genital infections.

Legionnaires' disease. This is caused by *Legionella pneumophila* and other *Legionella* spp. It is described on page 860.

Lung abscess. See page 836.

Tuberculosis. See pages 145 and 863.

FURTHER READING

Hewlet EL, Edwards KM. Clinical practice. Pertussis – not just for kids. *New England Journal of Medicine* 2005; **352**: 1215–1220.

Munoz-Price LS, Weinstein RA. *Acinetobacter* infection. *New England Journal of Medicine* 2008; **358**: 1271–1281.

Stevens DL, Bisno AL, Chambers HF et al. Practice guidelines for the diagnosis and management of skin and soft tissue infections. *Clinical Infectious Diseases* 2005; **41**: 1373–1406.

GASTROINTESTINAL INFECTIONS

Gastroenteritis

The most common form of acute gastrointestinal infection is gastroenteritis, causing diarrhoea with or without vomiting. Children in the developing world can expect, on average, three to six bouts of severe diarrhoea every year. Although oral rehydration programmes have cut the death toll significantly, at least 2.25 million people die every year as a direct result of diarrhoeal disease. In the western world diarrhoea is both less common and less likely to cause death. However, it remains a major cause of morbidity, especially in the elderly. Other groups who are at increased risk of infectious diarrhoea include travellers to developing countries, homosexual men, and infants in day care facilities. Viral gastroenteritis (see p. 144) is a common cause of diarrhoea and vomiting in young children, but is rarely seen in adults, other than in the context of common source outbreaks, usually due to noroviruses. Protozoal and helminthic gut infections (see p. 160) are rare in the West but relatively common in developing countries. The most common cause of significant adult gastroenteritis world-wide is bacterial infection.

Mechanisms

Bacteria can cause diarrhoea in three different ways (Table 4.27). Some species may employ more than one of these methods.

Mucosal adherence

Most bacteria causing diarrhoea must first adhere to specific receptors on the gut mucosa. A number of different molecular adhesion mechanisms have been elaborated; for example, adhesions at the tip of the pili or fimbriae which protrude from the bacterial surface aid adhesion. For some pathogens this is merely the prelude to invasion or toxin production but others such as enteropathogenic *Escherichia coli (EPEC)* cause attachment-effacement mucosal lesions on electron microscopy (EM) and produce a secretory diarrhoea directly as a result of adherence. *Enteroaggregative E. coli (EAggEC)*

Table 4.27 Pathogenic mechanisms of bacterial gastroenteritis

Pathogenesis	Mode of action	Clinical presentation	Examples
Mucosal adherence	Effacement of intestinal mucosa	Moderate watery diarrhoea	Enteropathogenic *E. coli* (EPEC) Enteroaggregative *E. coli* (EAggEC) Diffusely adhering *E. coli* (DAEC)
Mucosal invasion	Penetration and destruction of mucosa	Dysentery	*Shigella* spp. *Campylobacter* spp. Enteroinvasive *E. coli* (EIEC)
Toxin production			
Enterotoxin	Fluid secretion without mucosal damage	Profuse watery diarrhoea	*Vibrio cholerae* *Salmonella* spp. *Campylobacter* spp. Enterotoxigenic *E. coli* (ETEC) *Bacillus cereus* *Staphylococcus aureus* producing enterotoxin B *Clostridium perfringens* type A
Cytotoxin	Damage to mucosa	Dysentery	*Salmonella* spp. *Campylobacter* spp. Enterohaemorrhagic *E. coli* 0157 (EHEC)

adhere in an aggregative pattern with the bacteria clumping on the cell surface and its toxin causes persistent diarrhoea in people in developing countries. Diffusely adhering *E. coli* (DAEC) adheres in a uniform manner and may also cause diarrhoea seen in children and in developing countries.

Mucosal invasion

Invasive pathogens such as *Shigella* spp., enteroinvasive *E. coli* (EIEC) and *Campylobacter* spp. penetrate into the intestinal mucosa. Initial entry into the mucosal cells is facilitated by the production of 'invasins', which disrupt the host cell cytoskeleton. Subsequent destruction of the epithelial cells allows further bacterial entry, which also causes the typical symptoms of dysentery: low-volume bloody diarrhoea, with abdominal pain.

Toxin production (see Fig. 4.3)

Gastroenteritis can be caused by different types of bacterial toxins:

- *Enterotoxins*, produced by the bacteria adhering to the intestinal epithelium, induce excessive fluid secretion into the bowel lumen, leading to watery diarrhoea, without physically damaging the mucosa, e.g. cholera, enterotoxigenic *E. coli* (ETEC). Some enterotoxins preformed in the food primarily cause vomiting, e.g. *Staph. aureus* and *Bacillus cereus*. A typical example of this is 'fried rice poisoning', in which *B. cereus* toxin is present in cooked rice left standing overnight at room temperature.
- *Cytotoxins* damage the intestinal mucosa and, in some cases, vascular endothelium as well (e.g. *E. coli* O157).

Clinical syndromes

Bacterial gastroenteritis can be divided on clinical grounds into two broad syndromes: *watery diarrhoea* (usually due to enterotoxins, or adherence), and dysentery (usually due to mucosal invasion and damage) (Box 4.8). With some pathogens such as *Campylobacter jejuni* there may be overlap between the two syndromes.

Management (see p. 131)

Box 4.8 Bacterial causes of watery diarrhoea and dysentery

Watery diarrhoea

Bacillus cereus, *Staphylococcus aureus*, *Vibrio cholerae* — plus profuse vomiting

Enterotoxigenic *Escherichia coli* (ETEC)
Enteropathogenic *Escherichia coli* (EPEC)
Salmonella spp.
Campylobacter jejuni
Clostridium perfringens
Clostridium difficile

Dysentery

Shigella spp.
Salmonella spp.
Campylobacter spp.
Enteroinvasive *Escherichia coli* (EIEC)
Enterohaemorrhagic *Escherichia coli* (EHEC)
Yersinia enterocolitica
Vibrio parahaemolyticus
Clostridium difficile

Salmonella

Gastroenteritis can be caused by many of the numerous serotypes of salmonella (all of which are members of a single species, *S. choleraesuis*), but the most commonly implicated are *S. enteritidis* and *S. typhimurium*. These organisms, which are found all over the world, are commensals in the bowels of livestock (especially poultry) and in the oviducts of chicken (where the eggs can become infected). They are usually transmitted to man in contaminated foodstuffs and water.

Salmonellae can affect both the large and small bowel, and induce diarrhoea both by production of enterotoxins and by epithelial invasion. The typical symptoms commence abruptly 12–48 hours after infection and consist of nausea, cramping abdominal pain, diarrhoea, and sometimes fever. The diarrhoea can vary from profuse and watery to a bloody dysentery syndrome. Spontaneous resolution usually occurs in 3–6 days, although the organism may persist in the faeces for several weeks. Bacteraemia occurs in 1–4% of cases and is more common in the elderly and the immunosuppressed. Occasionally bacteraemia is complicated by metastatic infection, especially of atheroma on vascular endothelium, with potentially devastating consequences. In healthy adults salmonella gastroenteritis is usually a relatively minor illness, but young children and the elderly are at risk of significant dehydration.

Specific diagnosis is made by culturing the organism from blood or faeces, but management is usually empirical. Antibiotic therapy (ciprofloxacin 500 mg twice daily) may decrease the duration and severity of symptoms, but is rarely warranted (see Box 4.10).

More online
www.studentconsult.com

Campylobacter jejuni

C. jejuni is also a zoonotic infection, existing as a bowel commensal in many species of livestock, e.g. poultry and cattle. It is found world-wide, and is a common cause of childhood gastroenteritis in developing countries. Adults in these countries may be tolerant of the organism, excreting it asymptomatically. In the West it is a common cause of sporadic food-borne outbreaks of diarrhoea, the commonest causes being undercooked meat, e.g. chicken, and contaminated milk products.

Like salmonella, campylobacter can affect large and small bowel and can cause a wide variety of symptoms. The incubation period is usually 2–4 days, after which there is an abrupt onset of nausea, diarrhoea and abdominal cramps. The diarrhoea is usually profuse and watery, but an invasive haemorrhagic colitis is sometimes seen. Bacteraemia is very rare, and infection is usually self-limiting in 3–5 days. Diagnosis is made from stool cultures. If symptoms are severe, azithromycin 500 mg once daily is the drug of choice (see Box 4.10).

Shigella

Shigellae are enteroinvasive bacteria which cause classical bacillary dysentery. The principal species causing gastroenteritis are *S. dysenteriae*, *S. flexneri* and *S. sonnei*, which are found with varying prevalence in different parts of the world. All cause a similar syndrome, as a result of damage to the intestinal mucosa. Some strains of *S. dysenteriae* also secrete a cytotoxin affecting vascular endothelium. Although shigellae are found world-wide, transmission is strongly associated with poor hygiene. The organism is spread from

person to person, and only small numbers need to be ingested to cause illness (<200, compared to 10^4 for campylobacter and $>10^5$ for salmonella). Bacillary dysentery is far more prevalent in the developing world, where the main burden falls on children.

Symptoms start 24–48 hours after ingestion and typically consist of frequent small-volume stools containing blood and mucus. Dehydration is not as significant as in the secretory diarrhoeas, but systemic symptoms and intestinal complications are worse. The illness is usually self-limiting in 7–10 days, but in children in developing countries the mortality may be as high as 20%. Antibiotic treatment decreases the severity and duration of diarrhoea, and possibly reduces the risk of further transmission (see Box 4.10). Resistance to antibiotics is widespread: in some areas amoxicillin or co-trimoxazole may still be effective, but in many places nalidixic acid or ciprofloxacin is needed.

Enteroinvasive Escherichia coli *(EIEC)*

This causes an illness indistinguishable from shigellosis. Definitive diagnosis is made by stool culture.

Enterohaemorrhagic Escherichia coli *(EHEC)*

EHEC (usually serotype O157:H7, and also known as verotoxin-producing *E. coli*, or VTEC) is a well recognized cause of gastroenteritis in man. It is a zoonosis usually associated with cattle, and there have been a number of major outbreaks (notably in Scotland and Japan) associated with contaminated food. A variety of modes of transmission have been reported, and EHEC is a paradigm for all enteric livestock-associated zoonoses (Fig. 4.26). EHEC secretes a toxin (Shiga-like toxin 1) which affects vascular endothelial cells in the gut and in the kidney. After an incubation period of 12–48 hours it causes diarrhoea (frequently bloody), associated with abdominal pain and nausea. Some days after the onset of symptoms the patient may develop thrombotic thrombocytopenic purpura (see p. 137) or haemolytic uraemic syndrome (HUS, p. 598). This is more common in children, and may lead to permanent renal damage or death. Treatment is mainly supportive: there is evidence that antibiotic therapy might precipitate HUS by causing increased toxin release.

Enterotoxigenic Escherichia coli *(ETEC)*

ETEC produce both heat-labile and heat-stable enterotoxins which stimulate secretion of fluid into the intestinal lumen. The result is watery diarrhoea of varying intensity, which usually resolves within a few days. Transmission is normally from person to person via contaminated food and water. The organism is common in developing countries, and is a major cause of travellers' diarrhoea (see below).

Vibrio

Cholera, due to *Vibrio cholerae*, is the prototypic pure enterotoxigenic diarrhoea: it is described on page 143.

Vibrio parahaemolyticus causes acute watery diarrhoea after eating raw fish or shellfish that has been kept for several hours without refrigeration. Explosive diarrhoea, abdominal cramps and vomiting occurs with a fever in 50%. It is self-limiting, lasting up to 10 days.

Yersiniosis

Yersinia enterocolitica infection is a zoonosis of a variety of domestic and wild mammals. Human disease can arise either via contaminated food products, e.g. pork, or from direct animal contact. *Y. enterocolitica* can cause a range of gastroenteric symptoms including watery diarrhoea, dysentery and mesenteric adenitis. The illness is usually self-limiting, but ciprofloxacin may shorten the duration. *Y. pseudotuberculosis* is a much less common human pathogen: it causes mesenteric adenitis and terminal ileitis.

Staphylococcus aureus

Some strains of *S. aureus* can produce a heat-stable toxin (enterotoxin B) which causes massive secretion of fluid into the intestinal lumen. It is a common cause of food-borne

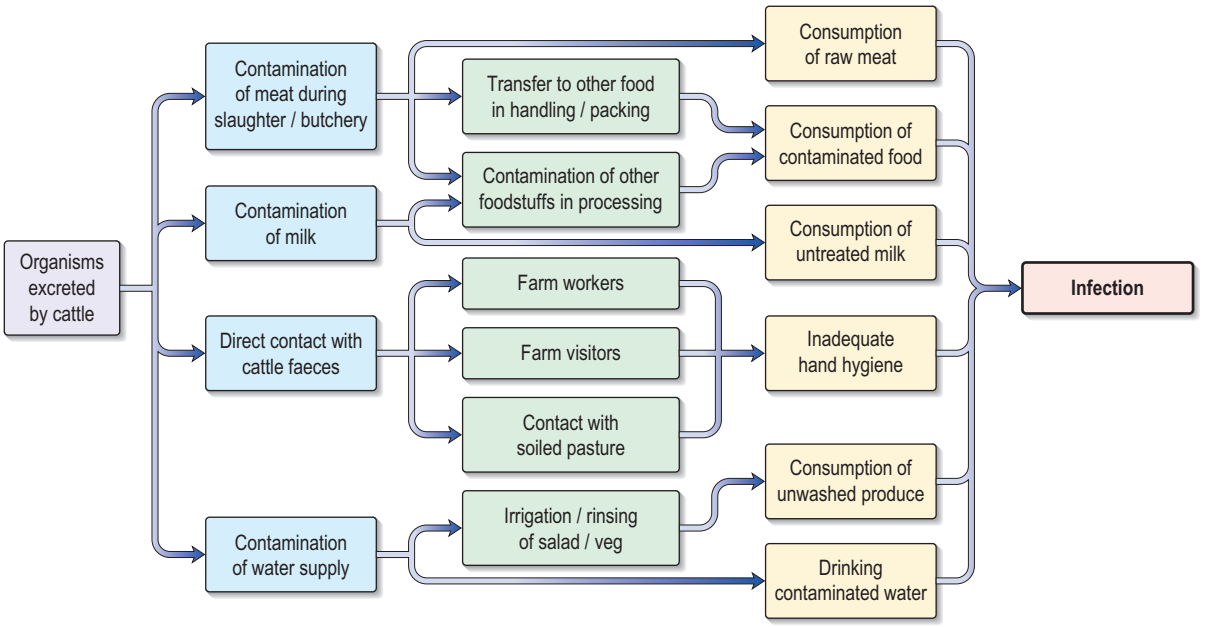


Fig. 4.26 **Routes of human infection with *E. coli* O157.**

gastroenteritis in Europe and the USA, outbreaks usually occurring as a result of poor food hygiene. Because the toxin is preformed in the contaminated food, onset of symptoms is rapid, often within 2–4 hours of consumption. There is violent vomiting, followed within hours by profuse watery diarrhoea. Symptoms have usually subsided within 24 hours.

Bacillus cereus

B. cereus produces two toxins. One produces watery diarrhoea up to 12 hours after ingesting the organism. The other toxin is preformed in food and causes severe vomiting, e.g. 'fried rice poisoning' (see p. 132).

Clostridial infections

Clostridium difficile causes antibiotic-associated diarrhoea, colitis and pseudomembranous colitis. It is a Gram-positive, anaerobic, spore-forming bacillus and is found as part of the normal bowel flora in 3–5% of the population and even more commonly (up to 20%) in hospitalized people.

Pathogenesis. *C. difficile* produces two toxins: toxin A is an enterotoxin while toxin B is cytotoxic and causes bloody diarrhoea. It causes illness either after other bowel commensals have been eliminated by antibiotic therapy or in debilitated patients who have not been on antibiotics. Almost all antibiotics have been implicated but the present increase has been attributed to the overuse of quinolones (e.g. ciprofloxacin). Hospital-acquired infections remain high, partly due to increased person-to-person spread and from fomites (see below). In recent years new strains of *C. difficile* with greater capacity for toxin production have been reported. There have been a number of hospital outbreaks with a high mortality.

Clinical features. *C. difficile* diarrhoea can begin anything from 2 days to a month after taking antibiotics. Elderly hospitalized patients are most frequently affected. It is unclear as to why some carriers remain asymptomatic. Symptoms can range from mild diarrhoea to profuse, watery, haemorrhagic colitis, along with lower abdominal pain. The colonic mucosa is inflamed and ulcerated and can be covered by an adherent membrane-like material (pseudomembranous colitis). The disease is usually more severe in the elderly and can cause intractable diarrhoea, leading to death.

Diagnosis is made by detecting A or B toxins in the stools by ELISA techniques.

Treatment is with metronidazole 400 mg three times daily or oral vancomycin 125 mg four times daily. Causative antibiotics should be discontinued if possible.

Prevention. Infection control relies on:

- Responsible use of antibiotics.
- Hygiene, which should involve all health workers, as well as patients and relatives. Washing hands thoroughly using soap and water is essential as alcohol disinfectants do not kill spores.
- Hospital cleaning of surfaces should be performed regularly to try and reduce transmission from formites.
- Isolation of patients with *C. difficile*.

Clostridium perfringens infection is due to inadequately cooked food, usually meat or poultry allowed to cool for a long time, during which the spores germinate. The ingested organism produces an enterotoxin causing watery diarrhoea with severe abdominal pain, usually without vomiting.

Travellers' diarrhoea

Travellers' diarrhoea is defined as the passage of three or more unformed stools per day in a resident of an industrialized country travelling in a developing nation. Infection is usually food- or water-borne, and younger travellers are most often affected (probably reflecting behaviour patterns). Reported attack rates vary from country to country, but approach 50% for a 2-week stay in many tropical countries. The disease is usually benign and self-limiting: treatment with quinolone antibiotics may hasten recovery but is not normally necessary. Prophylactic antibiotic therapy may also be effective for short stays, but should not be used routinely. The common causative organisms are listed in Table 4.28.

Management of acute gastroenteritis

In children, untreated diarrhoea has a high mortality due to dehydration, especially in hot climates. Death and serious morbidity are less common in adults but still occur, particularly in developing countries and in the elderly. The mainstay of treatment for all types of gastroenteritis is oral rehydration solutions (ORS): antibiotics have a subsidiary role in some cases (Fig. 4.27; Boxes 4.9 and 4.10 and p. 132). The use and formulation of ORS are discussed under cholera on page 143. It should also be remembered that other diseases, notably urinary tract infections and chest infections in the elderly, and malaria at any age, can present with acute diarrhoea.

Food poisoning

Food poisoning is a legally notifiable disease in England and Wales, and is defined as 'any disease of an infective or toxic nature caused by or thought to be caused by the consumption of food and water'. Not all cases of gastroenteritis are food poisoning, as the pathogens are not always food- or water-borne. Common bacterial causes of food poisoning are listed in Table 4.29. Food poisoning may also be caused by a number of non-infectious organic and inorganic toxins (Table 4.30). Illnesses such as botulism (see p. 134) are also classified as food poisoning, even though they do not primarily cause gastroenteritis. Listeriosis is described on page 138.

The increase in reported food poisoning in developed countries is at least in part due to changes in the production and distribution of food. Livestock raised and slaughtered

Table 4.28 Common causes of travellers' diarrhoea (TD)

Organism	Frequency (varies from country to country)
ETEC	30–70%
Shigella spp.	0–15%
Salmonella spp.	0–10%
Campylobacter spp.	0–15%
Viral pathogens	0–10%
Giardia intestinalis	0–3%

ETEC, enterotoxigenic *Escherichia coli*.

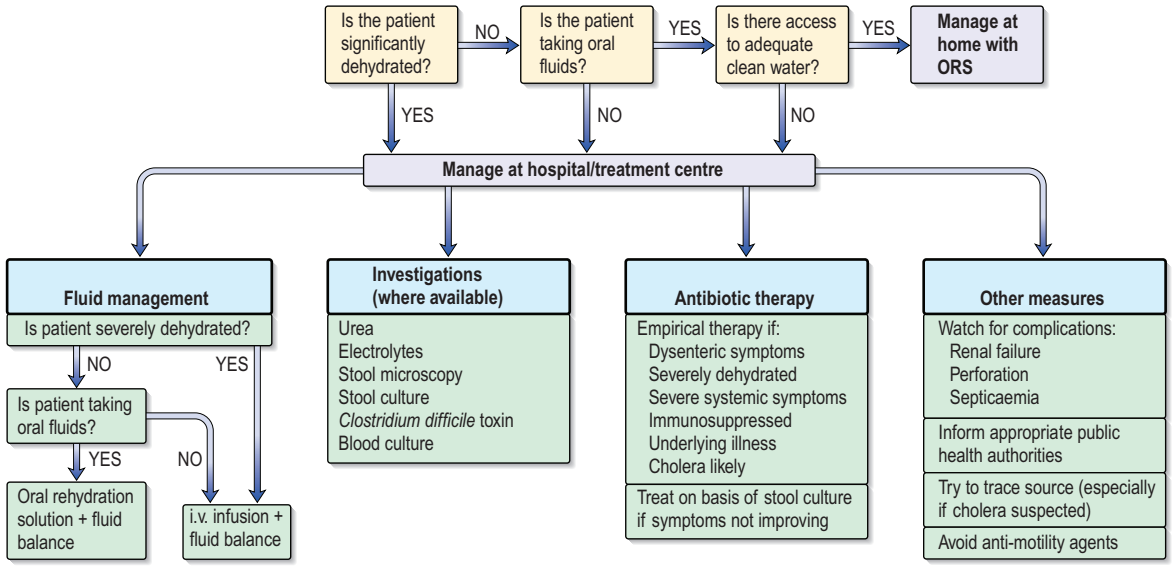


Fig. 4.27 **Gastroenteritis – management plan.** ORS, oral rehydration solution.

Box 4.9 Oral rehydration solutions (ORS) and intravenous solutions used in moderate and severe diarrhoea

	Salts (mmol per litre)				Substance added (per litre of water)	
	Na^+	K^+	Cl^-	Glucose		
Oral						
WHO	75	20	65	75	NaCl	2.6 g
New reduced osmolality formulation					Na citrate	2.9 g
					KCl	1.5 g
					Glucose	13.5 g
Cereal-based	85	–	80	–	80 g cooked rice 5 g salt (NaCl)	
Household	85	–	80	111	20 g glucose 5 g salt (NaCl)	
UK/Europe	35–60	20	37	90–200	Pre-prepared solutions	
Intravenous						
Ringer's lactate	131	4	109	0	Pre-prepared	

Box 4.10 Antibiotics in adult acute bacterial gastroenteritis

Condition	Indications	Drug of choice	Other drugs	Benefits
Dysentery	Most patients	Ciprofloxacin 500 mg twice daily	Nalidixic acid 1 g four times daily Ampicillin 500 mg four times daily Co-trimoxazole 960 mg twice daily	Relieve symptoms Shorten illness Decrease transmission
Cholera	All patients	Ciprofloxacin	Tetracycline 250 mg four times daily Azithromycin Nalidixic acid Co-trimoxazole	Relieve symptoms Shorten illness Decrease transmission
Empirical therapy of watery diarrhoea	Severe symptoms Prolonged illness Elderly patients Immunosuppressed	Ciprofloxacin	Azithromycin 500 mg once daily Co-trimoxazole	Relieve symptoms Shorten illness May decrease complications
Travellers' diarrhoea	Rarely used	Ciprofloxacin	Co-trimoxazole	Relieve symptoms Shorten illness
Treatment of confirmed *Salmonella, Shigella, Campylobacter*	Symptoms not improving (rarely needed)	Ciprofloxacin Azithromycin	Erythromycin Co-trimoxazole	May shorten illness
Clostridium difficile	Most cases (unless symptoms resolved)	Metronidazole 400 mg three times daily	Vancomycin 125 mg four times daily	Relieve symptoms Shorten illness

Table 4.29 Bacterial causes of food poisoning

Organism	Source/vehicles	Incubation period	Symptoms	Diagnosis	Recovery
Staphylococcus aureus	Man – contaminated food and water	2–4 h	Diarrhoea, vomiting and dehydration	Culture organism in vomitus or remaining food	<24 h
E. coli (ETEC)	Salads, water, ice	24 h	Watery diarrhoea	Stool culture	1–4 days
E. coli O157:H7	Cattle – contaminated foodstuffs	12–48 h	Watery diarrhoea ± haemorrhagic colitis, HUS	Stool culture	10–12 days
Yersinia enterocolitica	Milk, pork	2–14 h	Abdominal pain, vomiting, diarrhoea	Stool culture	2–30 days
Bacillus cereus	Environment – rice, ice-cream, chicken	1–6 h	Vomiting	Culture organism in faeces and food	Rapid
		6–14 h	Diarrhoea		
Clostridium perfringens	Environment – contaminated food	8–22 h	Watery diarrhoea and cramping pain	Culture organism in faeces and food	2–3 days
Listeria monocytogenes	Environment – milk, raw vegetables dairy products, unpasteurized cheese	Variable	Colic, diarrhoea and vomiting	Stool culture	Variable
Vibrio parahaemolyticus	Seafood	2–48 h	Diarrhoea, vomiting	Stool, food	2–10 days
Clostridium botulinum	Environment – bottled or canned food	18–24 h	Brief diarrhoea and paralysis due to neuromuscular blockade	Demonstrate toxin in food or faeces	10–14 days
Salmonella spp.	Cattle and poultry – eggs, meat	12–48 h	Abrupt diarrhoea, fever and vomiting	Stool culture	Usually 3–6 days, but may be up to 2 weeks
Campylobacter jejuni	Cattle and poultry – meat, milk	48–96 h	Diarrhoea ± blood, fever, malaise and abdominal pain	Stool culture	3–5 days
Shigella spp.	Man – contaminated food and water	24–48 h	Acute watery, bloody diarrhoea	Stool culture	7–10 days

HUS, haemolytic uraemic syndrome.

Table 4.30 Organic toxins causing food poisoning (see p. 950)

Toxin	Source	Illness
Scombrotoxin	Tuna, mackerel	Histamine fish poisoning
Ciguatoxin	Barracuda, snapper	Ciguatera (diarrhoea and paraesthesia)
Dinoflagellate plankton toxin	Shellfish	Neurotoxic shellfish poisoning
Haemagglutinins	Inadequately prepared dried kidney beans	Diarrhoea and vomiting
Unknown	Buffalo fish	Haff disease (toxic rhabdomyolysis)
Phallotoxins, amatoxins	Mushrooms	Various

under modern intensive farming conditions is frequently contaminated with salmonella or campylobacter. However, the main problem is not at this stage. Only 0.02–0.1% of the eggs from a flock of chickens infected with *S. enteritidis* will be affected, and then only at a level of less than 20 cells per egg – harmless to most healthy individuals. It is flaws in the processing, storage and distribution of food products which allow massive amplification of the infection, resulting in extensive contamination. The internationalization of the food supply encourages widespread and distant transmission of the resulting infections. Kitchen hygiene with careful separation of meat products from salads, along with using the correct temperature for cooking meat are necessary.

Other preventative measures include culling and vaccination of chicken.

Enteric fever (see p. 144)

Other gastrointestinal infections

Gastric infection with *Helicobacter pylori* is discussed on page 258, Whipple's disease on page 281, and bacterial peritonitis on page 316.

INFECTIONS OF THE CARDIOVASCULAR SYSTEM

- Infective endocarditis (see p. 768).

INFECTIONS OF THE NERVOUS SYSTEM

The central and peripheral nervous systems can be affected by a variety of microorganisms including bacteria, viruses (see p. 1156) and protozoa (see p. 157) which cause disease by direct invasion or via toxins. The nervous system is also vulnerable to prion disease (see pp. 121 and 1158).

FURTHER READING

Kuipers ES, Surawicz CM. *Clostridium difficile* infection. *Lancet* 2008; **371**: 1486–1488.

Musher DM, Musher BL. Contagious acute gastrointestinal infections. *New England Journal of Medicine* 2004; **351**: 2417–2427.

Bacterial meningitis (see p. 1153)

The most common bacterial disease affecting the central nervous system is acute meningitis, which causes about 175000 deaths per year, predominantly in the developing world. Epidemic meningitis due to *Neisseria meningitidis* (usually group A) is common in a broad belt across sub-Saharan Africa and is also seen in parts of Asia. In Europe and North America bacterial meningitis is usually sporadic, with B and C strains predominating. A conjugate vaccine for serogroup C meningococcus has resulted in a fall in the number of cases of C meningitis in those countries where it is now part of the childhood immunization schedule, such as the UK.

Streptococcus pneumoniae is the other major cause throughout the world, while tuberculous meningitis (see p. 1155) is common in sub-Saharan Africa and parts of Asia.

Haemophilus influenzae b (Hib) was once a common cause of meningitis in children, but since an effective vaccine has been available, serious *H. influenzae* infections have become rare in western countries. Many developing countries have also instituted immunization programmes, but invasive *H. influenzae* infection remains common in some parts of the world.

Other less common causes of meningitis in adults include group B streptococci, *Listeria monocytogenes* (see p. 138), *Staph. aureus* and Gram-negative bacilli. These organisms are usually associated with an underlying illness or immunocompromising condition, or with a cerebrospinal fluid leak.

FURTHER READING

Straus S, Thorpe K, Holroyd L. How do I perform a lumbar puncture and analyse the results to diagnose bacterial meningitis? *Journal of the American Medical Association* 2006; **296**: 2012–2022.

Van de Beek D, de Gans J, Tunkel AR et al. Community acquired bacterial meningitis in adults. *New England Journal of Medicine* 2006; **354**: 44–53.

Cerebral abscess

This is covered on page 1159.

Toxin-mediated infections

Botulism

Clostridium botulinum is a common environmental organism, and produces spores which can survive heating to 100°C. It causes botulism, an illness caused by contamination of canned or bottled foodstuff, in which the anaerobic organism can multiply and elaborate a neurotoxin. After ingestion the toxin causes profound neuromuscular blockade, leading to autonomic and motor paralysis. The first symptoms, occurring 18–24 hours after ingestion, are nausea and diarrhoea. These are followed by cranial nerve palsies and then progressive symmetrical paralysis, leading to respiratory failure.

The diagnosis is usually clinical, and is confirmed by detection of toxin in faeces or in the contaminated food. Treatment is mainly supportive, with mechanical ventilation if necessary. Antitoxin is available in some countries (including the UK); the risk of anaphylaxis is relatively high, and it should only be used in severe cases. A subcutaneous test dose should be given before intravenous or intramuscular injection. Antibiotics have no proven role. The overall mortality from botulism is high, but patients who survive the acute paralysis can make a full recovery.

Botulism may also follow the contamination of wounds, street heroin injection contaminated with *C. botulinum*, and in infants botulism may be related to bowel colonization by the organism.

Tetanus

Tetanus is also due to a toxin-secreting clostridium: *C. tetani*. The organism is found in soil, and illness usually results from a contaminated wound. The injury itself may be trivial and disregarded by the individual. It has also complicated intravenous drug use. In developing countries neonatal tetanus follows contamination of the umbilical stump, often after dressing the area with dung.

The organism is not invasive, and clinical manifestations of the disease are due to the potent neurotoxin, tetanospasmin. Tetanospasmin acts on both the α and δ motor systems at synapses, resulting in disinhibition. It also produces neuromuscular blockade and skeletal muscle spasm, and acts on the sympathetic nervous system. The end result is marked flexor muscle spasm and autonomic dysfunction.

Clinical features

The incubation period varies from a few days to several weeks. The most common form of the disease is generalized tetanus. General malaise is rapidly followed by trismus (lockjaw) due to masseter muscle spasm. Spasm of the facial muscles produces the characteristic grinning expression known as risus sardonicus. If the disease is severe, painful reflex spasms develop, usually within 24–72 hours of the initial symptoms. The interval between the first symptom and the first spasm is referred to as the 'onset time'. The spasms may occur spontaneously but are easily precipitated by noise, handling of the patient, or by light. Respiration may be impaired because of laryngeal spasm; oesophageal and urethral spasm lead to dysphagia and urinary retention, respectively, and there is arching of the neck and back muscles (opisthotonus). Autonomic dysfunction produces tachycardia, a labile blood pressure, sweating and cardiac arrhythmias. Patients with tetanus are mentally alert.

Death results from aspiration, hypoxia, respiratory failure, cardiac arrest or exhaustion. Mild cases with rigidity usually recover. Poor prognostic indicators include short incubation period, short onset time and extremes of age.

Localized tetanus is a milder form of the disease. Pain and stiffness are confined to the site of the wound, with increased tone in the surrounding muscles. Recovery usually occurs.

Cephalic tetanus is uncommon but invariably fatal. It usually occurs when the portal of entry of *C. tetani* is the middle ear. Cranial nerve abnormalities, particularly of the seventh nerve, are usual. Generalized tetanus may or may not develop.

Neonatal tetanus is usually due to infection of the umbilical stump. Failure to thrive, poor sucking, grimacing and irritability are followed by the rapid development of intense rigidity and spasms. Mortality approaches 100%. One aim of the WHO Expanded Programme on Immunization (EPI) is to eliminate this condition by immunizing all women of childbearing age, providing clean delivery facilities and strengthening surveillance in high-risk areas.

Diagnosis

Few diseases resemble tetanus in its fully developed form, and the diagnosis is therefore usually clinical. Rarely, *C. tetani* is isolated from wounds. Phenothiazine overdosage, strychnine poisoning, meningitis and tetany can occasionally mimic tetanus.

Management

Suspected tetanus. Any wound must be cleaned and debrided if necessary, to remove the source of toxin. Human

tetanus immunoglobin 250 units should be given along with an intramuscular injection of tetanus toxoid. If the patient is already protected a single booster dose of the toxoid is given; otherwise the full three-dose course of adsorbed vaccine is given (see below).

Established tetanus. Management is supportive medical and nursing care. Improvement in this area has contributed more than any other single measure to the decrease in the mortality rate from 60% to nearer 20%. Patients are nursed in a quiet, isolated, well-ventilated, darkened room. Benzodiazepines are used to control spasms and sedate the patient; if the airway is compromised intubation and mechanical ventilation may be necessary. Magnesium sulphate infusion decreases the need for antispasmodics.

Antibiotics and antitoxin should be administered, even in the absence of an obvious wound. Intravenous metronidazole is the drug of choice, although penicillin and cephalosporins are also effective. Human tetanus immunoglobulin (HTIG) 500 IU should be given by intramuscular injection to neutralize any circulating toxin. If HTIG is not available, immune equine tetanus immunoglobulin 10 000 IU should be given intramuscularly: this is probably as effective as HTIG, but there is a high incidence of severe allergic reactions. If the patient recovers, active immunization should be instituted, as immunity following tetanus is incomplete.

Prevention

Tetanus is an eminently preventable disease and all persons should be immunized regardless of age. Those who work in a contaminated environment, such as farmers, are particularly at risk and should have regular booster injections. Active immunization with the alum-adsorbed toxoid should be given. Initially two doses of 0.5 mL of the toxoid are given intramuscularly with an 8-week interval. The third dose is given 6–12 months later as a booster. Subsequent boosters are required at 5-year intervals. Infant immunization schedules in all countries include tetanus (Box 4.6). Protection by passive immunization with either the equine or human antitetanus immunoglobulin is short-lived, lasting only about 2 weeks.

BONE AND JOINT INFECTIONS

- Infective arthritis (see p. 539)
- Osteomyelitis (see p. 541).

URINARY TRACT INFECTIONS

- Complicated versus uncomplicated infections (see p. 600)
- Acute pyelonephritis (see p. 600)
- Reflux nephropathy (see p. 600)
- Perinephric abscess (renal carbuncle) (see p. 603)
- Bacterial prostatitis (see p. 603)
- Tuberculosis of the urinary tract (see p. 603)

SYSTEMIC/MULTISYSTEM INFECTIONS

Many infections are confined to a particular body organ or system, owing to the metabolic requirements of the organism, the route of infection or the response of host defences. Other infections can potentially affect several systems or the entire body. Under unusual circumstances such as altered host immunity, infections which are normally circumscribed may become systemic. This section describes those infections which commonly cause multisystem disease in an immunocompetent host.

Bacteraemia and sepsis syndrome

Bacteraemia, the transient presence of organisms in the blood, can occur in healthy people without causing symptoms. It can follow surgery, dental treatment and even toothbrushing. Bacteraemia can also occur from the bowel or bladder, especially in the presence of local inflammation. Unless a site of metastatic infection is established (such as the heart valves), most organisms are rapidly cleared from the blood.

Sepsis is the term used to describe the signs and symptoms of a systemic inflammatory response syndrome (SIRS) to a localized primary site of infection. Viral, bacterial, fungal and parasitic disease can all trigger the sepsis syndrome.

SIRS is not unique to infection and may complicate a variety of events and conditions such as trauma, chronic inflammatory diseases and malignancy (e.g. lymphoma).

Severe sepsis. This is the presence of the sepsis syndrome (presence of either a positive blood culture or clinical features of fever, tachypnoea, tachycardia, suspected infection), complicated by organ dysfunction, hypotension or hypoperfusion and manifested by low blood pressure, oliguria, hypoxia, acute confusion and lactic acidosis.

Septic shock is defined as the sepsis syndrome plus organ dysfunction and hypotension unresponsive to adequate fluid replacement. Mortality rates often exceed 50%.

Patients presenting with symptoms and signs suggesting sepsis syndrome should be examined for evidence of a source: common sites of infection and responsible organisms are listed in Tables 4.31 and 4.32. Because of the potential to progress to severe sepsis, treatment with antibiotics should usually be started empirically as soon as appropriate cultures have been taken. The choice of agent is governed by the likely pathogen: if there are no clues, a broad-spectrum regimen should be used (e.g. piperacillin plus gentamicin, cefotaxime, with or without metronidazole or imipenem). In hospital units with epidemic MRSA infection, it is common practice to add vancomycin or teicoplanin to this empiric regimen. Antibiotic therapy should be reviewed daily as the illness progresses and the results of investigations

Table 4.31 Causes of septicaemia in a previously healthy adult

Site of origin	Usual pathogen(s)
Skin	*Staphylococcus aureus* and other Gram-positive cocci
Urinary tract	*Escherichia coli* and other aerobic Gram-negative rods
Respiratory tract	*Streptococcus pneumoniae*
Gall bladder or bowel	*Enterococcus faecalis*, *E. coli* and other Gram-negative rods *Bacteroides fragilis*
Pelvic organs	*Neisseria gonorrhoeae*, anaerobes

Ask the authors
www.studentconsult.com

Table 4.32 Causes of septicaemia in hospitalized patients

Clinical problem	Usual pathogen(s)
Urinary catheter	*Escherichia coli*, *Klebsiella* spp., *Proteus* spp., *Serratia* spp., *Pseudomonas* spp.
Intravenous catheter	*Staphylococcus aureus*, *Staph. epidermidis*, *Klebsiella* spp., *Pseudomonas* spp., *Candida albicans*
Peritoneal catheter	*Staph. epidermidis*
Post-surgery:	
Wound infection	*Staph. aureus*, *E. coli*, anaerobes (depending on site)
Deep infection	Depends on anatomical location
Burns	Gram-positive cocci, *Pseudomonas* spp., *Candida albicans*
Immunocompromised patients	Any of the above

become available. The general management of the sepsis syndrome is covered on page 899.

Meningococcal septicaemia

Neisseria meningitidis is found world-wide, in five major serogroups. In sub-Saharan Africa and parts of Asia where group A meningococcus is prevalent it usually causes epidemic disease. Groups Y and W can also cause epidemic infection, while groups B and C (which are the predominant strains in Europe and North America) tend to be sporadic. In England and Wales approximately 1500 cases of meningococcal disease are reported annually.

Man is the only known reservoir for the organism, which is carried asymptomatically in the nasopharynx of 5–20% of the general population. Meningococcal disease occurs when the bacteria invade the nasal mucosa and enter the bloodstream: this only happens in a small percentage of those colonized. Invasion depends on both host and bacterial factors. It is more likely to take place soon after colonization has taken place, and following viral upper respiratory infections.

Clinical features

Invasive meningococcal infection may cause meningitis, septicaemia, or both. Meningitic disease (see p. 899) usually presents with the classical triad of headache, fever and neck stiffness. Vomiting, diminished consciousness and focal neurological signs occur, although some patients, especially in the early stages, only have mild symptoms. Meningococcal septicaemia causes the typical features of septic shock such as fever, myalgia and hypotension (see p. 899), and may be accompanied by a petechial or haemorrhagic rash (Fig. 4.28). In some cases the patient can deteriorate rapidly, with shock, disseminated intravascular coagulation and multiorgan failure.

More online
www.studentconsult.com

Diagnosis

The presence of meningitis and septicaemia with a typical rash is strongly suggestive of meningococcal disease. Gram-negative diplococci may be seen on Gram stain of CSF or of aspirate from petechiae, and meningococci can also be cultured from CSF or blood, or detected by PCR.

Fig. 4.28 **Meningococcal infections, showing a purpuric rash.**

Management

N. meningitidis is sensitive to benzylpenicillin, third-generation cephalosporins and chloramphenicol: antibiotic treatment for meningococcal meningitis should be started **immediately** (see Emergency Box 21.1, p. 1154) and continued parenterally for 7 days. Meningococcal septicaemia should be managed in the same way as any other septicaemic illness (see p. 905). The mortality from meningococcal septicaemia in developed countries is currently approximately 10%, while that from meningococcal meningitis alone is less than 5% (see below). Mild neurological sequelae (especially vestibular nerve damage) are common, but serious brain damage is relatively unusual.

The meningococcal C conjugate vaccine has led to an overall reduction of meningococcal C disease in the UK. A serogroup B vaccine is not currently available but is under development. A combined A and C vaccine is also available for outbreak control (see p. 423).

For close contacts of a case of meningococcal disease, household and 'kissing' contacts should be given prophylaxis with oral rifampicin or ciprofloxacin to eradicate the bacteria from the nasopharynx. In the case of group C disease, contacts should be offered immunization.

Rheumatic fever

Rheumatic fever is an inflammatory disease that occurs in children and young adults (the first attack usually occurs at between 5 and 15 years of age) as a result of infection with group A streptococci. It affects the heart, skin, joints and central nervous system. It is common in the Middle and Far East, eastern Europe and South America. It is rare in the UK, western Europe and North America. This decline in the incidence of rheumatic fever (from 10% of children in the 1920s to 0.01% today) parallels the reduction in all streptococcal infections and is largely due to improved hygiene and the use of antibiotics.

Pharyngeal infection with group A streptococcus is followed by the clinical syndrome of rheumatic fever. This is thought to develop because of an autoimmune reaction triggered by molecular mimicry between the cell wall M proteins of the infecting *Streptococcus pyogenes* and cardiac myosin

and laminin. The condition is not due to direct infection of the heart or to the production of a toxin.

Pathology

All three layers of the heart may be affected. The characteristic lesion of rheumatic carditis is the Aschoff nodule, which is a granulomatous lesion with a central necrotic area occurring in the myocardium. Small, warty vegetations may develop on the endocardium, particularly on the heart valves. A serofibrinous effusion characterizes the acute pericarditis that occurs.

The synovial membranes are acutely inflamed during rheumatic fever, and subcutaneous nodules (which are also granulomatous lesions) are seen in the acute stage of the disease.

Clinical features

The disease presents suddenly, with fever, joint pains, malaise and loss of appetite. Diagnosis relies on the presence of two or more major clinical manifestations or one major manifestation plus two or more minor features. These are known as the modified Jones criteria (Table 4.33).

Carditis manifests as:

- new or changed heart murmurs
- development of cardiac enlargement or cardiac failure
- appearance of a pericardial effusion and ECG changes of pericarditis, myocarditis, AV block or other cardiac arrhythmias.

Non-cardiac features include the following:

- Fever.
- The arthritis associated with rheumatic fever is classically a fleeting migratory polyarthritis affecting large joints such as the knees, elbows, ankles and wrists. Once the acute inflammation disappears, the rheumatic process leaves the joints normal.
- Sydenham's chorea (or St Vitus' dance, see p. 1149) is involvement of the central nervous system that develops late after a streptococcal infection. Sufferers are noticeably 'fidgety' and display spasmodic, unintentional choreiform movements. Speech is often affected.
- Skin manifestations include erythema marginatum, a transient pink rash with slightly raised edges, which occurs in 20% of cases. The erythematous areas found mostly on the trunk and limbs coalesce into crescent- or ring-shaped patches. Subcutaneous nodules, which are painless, pea-sized, hard nodules beneath the skin, may also occur.

Table 4.33 Revised Jones criteria for the diagnosis of rheumatic fever

Major criteria
Carditis
Polyarthritis
Chorea
Erythema marginatum
Subcutaneous nodules
Minor criteria
Fever
Arthralgia
Previous rheumatic fever
Raised ESR/C-reactive protein
Leucocytosis
Prolonged PR interval on ECG
Plus evidence of antecedent streptococcal infection, e.g. positive throat cultures for group A streptococci, elevated antistreptolysin O titre (>250 U) or other streptococcal antibodies, or a history of recent scarlet fever

ESR, erythrocyte sedimentation rate.

Investigations

- **Throat swabs** are cultured for the group A streptococcus.
- Antistreptolysin O titre and antiDNAse B may be elevated.
- ESR and CRP are usually high.
- **Cardiac investigations**, e.g. ECG, echocardiogram.

Treatment

Patients with active arthritis or carditis should be rested in bed. When the clinical syndrome has subsided (e.g. no pyrexia, normal pulse rate, normal ESR, normal white cell count) the patient is mobilized.

Residual streptococcal infections should be eradicated with oral phenoxymethylpenicillin 500 mg four times daily for 1 week. This therapy should be administered even if nasal or pharyngeal swabs do not culture the streptococci.

The arthritis of rheumatic fever responds dramatically to NSAIDs. There is no good evidence that steroids are of benefit, although some experts give high-dose prednisolone if there is severe carditis. Recurrences are most common when persistent cardiac damage is present, and are prevented by the continued administration of oral phenoxymethylpenicillin 250 mg twice daily or intramuscular benzathine penicillin G 1.2 million units monthly until the age of 20 years or for 5 years after the latest attack (see p. 93). Erythromycin is used if the patient is allergic to penicillin. Any streptococcal infection that does develop should be treated promptly.

Prognosis

More than 50% of those who suffer acute rheumatic fever with carditis will later (after 10–20 years) develop chronic rheumatic valvular disease, predominantly affecting the mitral and aortic valves (see Table 13.34).

Leptospirosis

Leptospirosis is a zoonosis caused by the spirochaete *Leptospira interrogans*. There are over 200 serotypes: the main types affecting humans are *L. i. icterohaemorrhagiae* (rodents), *L. i. canicola* (dogs and pigs), *L. i. hardjo* (cattle) and *L. i. pomona* (pigs and cattle). Leptospires are excreted in the animal urine, and enter the host through a skin abrasion or through intact mucous membranes. Leptospirosis can also be caught by ingestion of contaminated water. The organism can survive for many days in warm fresh water, and for up to 24 hours in sea water.

In England and Wales only about 40 cases of leptospirosis are reported every year (although many mild infections probably go undiagnosed), and it remains largely an occupational disease of farmers, vets and others who work with animals. In some parts of the world (e.g. Hawaii, where the annual incidence is about 130/100 000) it is associated with a variety of recreational activities which bring people into

closer contact with rodents. Outbreaks of leptospirosis have also been associated with flooding.

Clinical features

Weil, in 1886, described a severe illness consisting of jaundice, haemorrhage and renal impairment caused by *L. i. icterohaemorrhagiae*, but fortunately 90–95% of infections are subclinical or cause only a mild fever. The incubation period of leptospirosis is usually 7–14 days, and the illness typically has two phases. A leptospiraemic phase, which lasts for up to a week, is followed after a couple of days' interval by an immunological phase. The first phase is characterized by severe headache, malaise, fever, anorexia and myalgia. Most patients have conjunctival suffusion. Hepatosplenomegaly, lymphadenopathy and various skin rashes are sometimes seen. The second phase is usually mild. Fifty per cent of patients have meningism, about a third of whom have a CSF lymphocytosis. The majority of patients recover uneventfully at this stage.

In severe disease there may not be a clear distinction between phases. Following the initial symptoms, patients progressively develop hepatic and renal failure, haemolytic anaemia and circulatory collapse. Cardiac failure and pulmonary haemorrhage may also occur. Even with full supportive care the mortality is around 10%, rising to 15–20% in the elderly.

Diagnosis

The diagnosis is usually a clinical one. Leptospires can be cultured from blood or CSF during the first week of illness, but culture requires special media and may take several weeks. A minority of patients may also excrete the organism in their urine from the second week onwards. More usually, the diagnosis is made serologically. Specific IgM antibodies start to appear from the end of the first week and the diagnosis is often made retrospectively with a microscopic agglutination test (MAT) showing a four-fold rise. There is typically a leucocytosis and in severe infection, thrombocytopenia and an elevated creatine phosphokinase.

Management

Early antibiotic therapy will limit the progress of the disease, but treatment should still be initiated whatever the stage of the infection. Oral doxycycline may be used in mild cases: intravenous penicillin or erythromycin is given in more severe disease. Intensive supportive care is needed for those patients who develop hepatorenal failure.

Brucellosis

Brucellosis (Malta fever, undulant fever) is a zoonosis and has a world-wide distribution, although it has been virtually eliminated from cattle in the UK where there are only about 15 human infections annually. The highest incidence is in the Mediterranean countries, the Middle East and the tropics; there are about 500 000 new cases diagnosed per year.

The organisms usually gain entry into the human body via the mouth; less frequently they may enter via the respiratory tract, genital tract or abraded skin. The bacilli travel in the lymphatics and infect lymph nodes. This is followed by haematogenous spread with ultimate localization in the reticuloendothelial system. Acquisition is usually by the ingestion of raw milk from infected cattle or goats, although occupational exposure is also common. Person-to-person transmission is rare.

Clinical features

The incubation period of acute brucellosis is 1–3 weeks. The onset is insidious, with malaise, headache, weakness, generalized myalgia and night sweats. The fever pattern is classically undulant, although continuous and intermittent patterns are also seen. Lymphadenopathy and hepatosplenomegaly are common; sacro-iliitis, arthritis, osteomyelitis, epididymo-orchitis, meningoencephalitis and endocarditis have all been described.

Untreated brucellosis can give rise to chronic infection, lasting a year or more. This is characterized by easy fatiguability, myalgia, and occasional bouts of fever and depression. Splenomegaly is usually present. Occasionally infection can lead to localized brucellosis. Bones and joints, spleen, endocardium, lungs, urinary tract and nervous system may be involved. Systemic symptoms occur in less than one-third.

Diagnosis

Blood (or bone marrow) cultures are positive during the acute phase of illness in 50% of patients (higher in *B. melitensis*), but prolonged culture is needed. In chronic disease serological tests are of greater value. The brucella agglutination test, which demonstrates a fourfold or greater rise in titre (>1 in 160) over a 4-week period, is highly suggestive of brucellosis. An elevated serum IgG level is evidence of current or recent infection; a negative test excludes chronic brucellosis. In localized brucellosis antibody titres are low, and diagnosis is usually established by culturing the organisms from the involved site. Species-specific PCR tests are also available.

Management and prevention

Brucellosis is treated with a combination of doxycycline 200 mg daily and rifampicin 600–900 mg daily for 6 weeks, but relapses occur. Alternatively, tetracycline can be combined with streptomycin, which is usually given for only the first 2 weeks of treatment. Prevention and control involve careful attention to hygiene when handling infected animals, vaccination with the eradication of infection in animals, and pasteurization of milk. No vaccine is available for use in humans.

Listeriosis

Listeria monocytogenes is an environmental organism which is widely disseminated in soil and decayed matter. It affects both animals and man: the most common route of human infection is in contaminated foodstuffs. The organism can grow at temperatures as low as 4°C, and the most commonly implicated foods are unpasteurized soft cheeses, raw vegetables and chicken pâtés. Listeriosis is a rare but serious infection affecting mainly neonates, pregnant women, the elderly and the immunocompromised. *L. monocytogenes* has also been recognized as a cause of self-limiting foodborne gastroenteritis in healthy adults, but the incidence of this is unknown.

In pregnant women listeria causes a flu-like illness, but infection of the fetus can lead to septic abortion, premature labour and stillbirth. Early treatment of listeria in pregnancy may prevent this, but the overall fetal loss rate is about 50%. In the elderly and the immunocompromised listeria can cause

www.studentconsult.com More online

meningoencephalitis. Septicaemia and a variety of other focal infections have also been described.

The diagnosis is established by culture of blood, CSF or other body fluids. The treatment of choice for adult listeriosis is ampicillin plus gentamicin. Co-trimoxazole is also effective, but the organism is resistant to cephalosporins.

Q fever

Q (query) fever is a zoonosis caused by the rickettsia-like organism *Coxiella burnetii*. Infection is widespread in domestic, farm and other animals, birds and arthropods: spread is mainly by ticks. Modes of transmission to humans are by dust, aerosol, and unpasteurized milk from infected cows. The formation of spores means that *C. burnetii* can survive in extreme environmental conditions for long periods. The infective dose is very small, so that minimal animal contact is required. One reported outbreak occurred among inhabitants of a village through which infected sheep had passed. Infection in the UK is rare but has been linked to meat processing.

Clinical features

Symptoms begin insidiously 2–4 weeks after infection. Fever is accompanied by flu-like symptoms with myalgia and headache. The acute illness usually resolves spontaneously but pneumonia or hepatitis may develop. Occasionally infection can become chronic, with endocarditis, myocarditis, uveitis, osteomyelitis or other focal infections.

C. burnetii is an obligate intracellular organism, and does not grow on standard culture media. Diagnosis is made serologically using an immunofluorescent assay. Antibody tests for two different bacterial antigens allow distinction between acute and chronic infection. A PCR assay is now available, but the sensitivity is low.

Management

Treatment with doxycycline 200 mg daily for 2 weeks reduces the duration of the acute illness, but it is not known whether this correlates with eradication of the organism. Azithromycin may also be used. For chronic Q fever, including endocarditis, doxycycline is often combined with hydroxychloroquine. Even prolonged courses of treatment may not clear the infection. A vaccine is available for those at high risk.

Lyme disease

Lyme disease is caused by a spirochaete, *Borrelia burgdorferi*, which has at least 11 different genomic species. It is a zoonosis of deer and other wild mammals. The syndrome was first recognized and named following an 'outbreak' of arthritis in the town of Lyme, Connecticut, in the mid 1970s. Since then the disease has increased in both incidence and detection: it is now known to be widespread in the USA, Europe, Russia and the Far East. Infection is transmitted from animal to man by ixodid ticks, and is most likely to occur in rural wooded areas in spring and early summer.

Clinical features

The first stage of the illness, which follows 7–10 days after infection, is characterized by the skin lesion erythema migrans at the site of the tick bite. This is often accompanied by headache, fever, malaise, myalgia, arthralgia and lymphadenopathy. Many people have no further illness after this. A second stage may follow weeks or months later, when some patients develop neurological symptoms (meningoencephalitis, cranial or polyneuropathies), radiculopathies, cardiac problems (conduction disorders, myocarditis) or arthritis. These manifestations, which are often fluctuant and changing, usually resolve spontaneously over months or years. Some patients, however, develop chronic and persistent neurological disease (e.g. paraparesis) or rheumatological disease ('late Lyme disease'), although the evidence for persistent infection is lacking (i.e. PCR negative). Acrodermatitis chronica atrophicans, usually on the backs of the hands and feet, can occur if the disease does not resolve spontaneously.

Diagnosis

The clinical features and epidemiological considerations are usually strongly suggestive. The diagnosis can only rarely be confirmed by isolation of the organisms from blood, skin lesions or CSF. IgM antibodies are detectable in the first month, and IgG antibodies are invariably present late in the disease. Commercial tests (ELISA, immunofluorescence, haemagglutination) are available but false-positive results do occur. A genuine positive IgG may be a marker of previous exposure rather than of ongoing infection.

Management

Amoxicillin or doxycycline given early in the course of the disease shortens the duration of the illness in approximately 50% of patients. Late disease should be treated with 2–4 weeks of intravenous benzylpenicillin or ceftriaxone. However, treatment is unsatisfactory, and preventative measures are necessary. In tick-infested areas, repellents and protective clothing should be worn. Prompt removal of any tick is essential as infection is unlikely to take place until the tick has been attached for more than 48 hours. Ticks should be grasped with forceps near to the point of attachment to the skin and then withdrawn by gentle traction. Antibiotic prophylaxis following a tick bite is not usually justified, even in areas where Lyme disease is common. There is currently no effective vaccine.

Tularaemia

Tularaemia is due to infection by *Francisella tularensis*, a Gram-negative organism. It is primarily a zoonosis, acquired mainly from rodents. Infection can be transmitted by arthropod vectors or by handling infected animals, when the microorganisms enter through minor abrasions or mucous membranes. Occasionally infection occurs from contaminated water or from eating uncooked meat. The disease is widely distributed in North America, Northern Europe and Asia. It is relatively rare, occurring mainly in hunters, trappers, and others in close contact with animals.

The incubation period of 2–7 days is followed by a generalized illness. The most common presentation is ulceroglandular tularaemia. A papule occurs at the site of inoculation. This ulcerates and is followed by tender, suppurative lymphadenopathy. Rarely this can be followed by bacteraemia, leading to septicaemia, pneumonia or meningitis. These forms of the disease carry a high mortality if untreated.

Diagnosis is by culture of the organism or by a rising titre seen on a bacterial agglutination test.

Tularaemia should be treated with streptomycin or gentamicin.

FURTHER READING

Feder HM, Johnson B, O'Connell S et al. A critical appraisal of chronic Lyme disease. *New England Journal of Medicine* 2007; **357**: 1422–1430.

Olive C. Progress in M-protein based subunit vaccines to prevent rheumatic fever and rheumatic heart disease. *Current Opinion in Molecular Therapeutics* 2007; **9**: 25–34.

Palaniappan R, Ramanujam S, Chang YF. Leptospirosis: pathogenesis, immunity, and diagnosis. *Current Opinion in Infectious Diseases* 2007; **20**: 284–292.

Pappas G, Akritidis N, Bosilkovski M et al. Brucellosis. *New England Journal of Medicine* 2005; **352**: 2325–2336.

Parker NR, Barralet JH, Bell AM. Q Fever. *Lancet* 2006; **367**: 679–683.

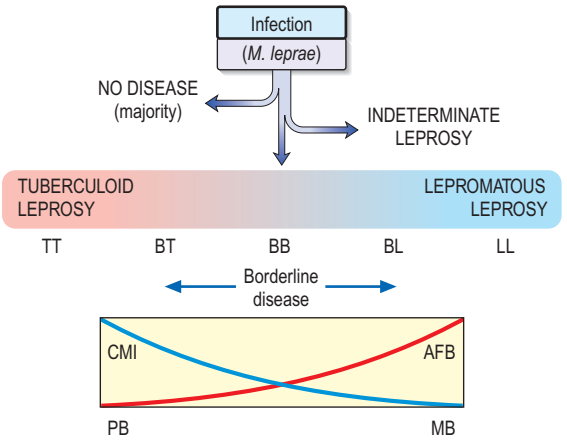


Fig. 4.29 **Clinical spectrum of leprosy with the combined Ridley–Jopling and WHO classification.** BT, borderline tuberculoid; BB, borderline; BL, borderline lepromatous; CMI, cell-mediated immunity; AFB, acid-fast bacilli; PB, paucibacillary; MB, multibacillary. Modified from Jacobson RR, Krahenbuhl JL. Leprosy. *Lancet* 1999; **353**: 655–660, with permission from Elsevier.

BACTERIAL INFECTIONS SEEN IN DEVELOPING AND TROPICAL COUNTRIES

SKIN, SOFT TISSUE AND EYE DISEASE

Leprosy

www.studentconsult.com
More online

Leprosy (Hansen's disease) is caused by the acid-fast bacillus *Mycobacterium leprae*. Unlike other mycobacteria, it does not grow in artificial media or even in tissue culture. Apart from the nine-banded armadillo, man is the only natural host of *M. leprae*, although it can be grown in the footpads of mice.

The WHO campaign to control leprosy has been hugely successful, with more than 13 million people having been cured of the disease. The number of people with active leprosy has fallen from 5.4 million in 1985 to about 286 000 at the end of 2004, largely as the result of supervised multidrug treatment regimens. The majority of the remaining cases are in India and Brazil, and despite the successes many new infections are occurring in these countries.

The precise mode of transmission of leprosy is still uncertain but it is likely that nasal secretions play a role. Infection is related to poverty and overcrowding. Once an individual has been infected, subsequent progression to clinical disease appears to be dependent on several factors. Males appear to be more susceptible than females, and there is evidence from twin studies of a genetic susceptibility. The main factor, however, is the response of the host's cell-mediated immune system.

Two *polar types* of leprosy are recognized (Fig. 4.29):

- *Tuberculoid leprosy*, a localized disease that occurs in individuals with a high degree of cell-mediated immunity (CMI). The T cell response to the antigen releases interferon which activates macrophages to destroy the bacilli (Th1 response) but with associated destruction of the tissue. The lepromin test (see below) is positive.
- *Lepromatous leprosy*, a generalized disease that occurs in individuals with impaired CMI (Fig. 4.29). Here the tissue macrophages fail to be activated and the bacilli multiply intracellularly. Th2 cytokines are produced. The lepromin test is negative.

The WHO classification of leprosy depends on the number of skin lesions and the number of bacilli detected on the skin smears: *paucibacillary leprosy* has 5 or fewer skin lesions with no bacilli; *multibacillary leprosy* has 6 or more lesions which may have bacilli.

In practice many patients will fall between these two extremes and some may move along the spectrum as the disease progresses or is treated.

Clinical features

The incubation period varies from 2 to 6 years, although it may be as short as a few months or as long as 20 years. The onset of leprosy is generally insidious. Acute onset is known to occur. Patients may present with a transient rash, features of an acute febrile illness, evidence of nerve involvement, or with any combination of these. The major signs of leprosy are:

- Skin lesions, usually anaesthetic (generally tuberculoid).
- Thickened peripheral nerves, nerves of predilection which are superficial or lie in fibro-osseous tunnels – ulnar (elbow), median (wrist), radial cutaneous (wrist), common peroneal (knee), posterior tibial and sural (ankle), facial (crossing zygomatic arch) and greater auricular (posterior triangle of the neck).

The spectrum of disease can be divided into five clinical groups:

Tuberculoid leprosy (TT). In tuberculoid leprosy the characteristic, usually single, skin lesion is a hypopigmented, anaesthetic patch with thickened, clearly demarcated edges, central healing, and atrophy. The face, gluteal region and extremities are most commonly affected. The nerve leading to the hypopigmented patch, and the regional nerve trunk, are often thickened and tender. Unlike other parts of the body, a tuberculoid patch on the face is not anaesthetic. Nerve involvement leads to marked muscle atrophy. Tuberculoid lesions are known to heal spontaneously. The prognosis is good.

Borderline tuberculoid (BT) leprosy. This resembles TT but skin lesions are usually more numerous, smaller, and may be

present as small 'satellite' lesions around larger ones. Peripheral but not cutaneous nerves are thickened, leading to deformity of hands and feet.

Borderline (BB) leprosy. Skin lesions are numerous, varying in size and form (macules, papules, plaques). The annular, rimmed lesion with punched-out, hypopigmented anaesthetic centre is characteristic (Fig. 4.30). There is widespread nerve involvement and limb deformity (Fig. 4.31a).

Borderline lepromatous (BL) leprosy. There are a large number of florid asymmetrical skin lesions of variable form, which are strongly positive for acid-fast bacilli. Skin between the lesions is normal and often negative for bacilli.

Lepromatous leprosy (LL). Although practically every organ can be involved, the changes in the skin are the earliest and most obvious manifestation. Peripheral oedema and rhinitis are the earliest symptoms. The skin lesions predominantly occur on the face, the gluteal region and the upper and lower limbs. They may be macules, papules, nodules or plaques: of these, the macule is the first to appear. Infiltration is most noticeable in the ear lobes. Thinning of the lateral margins of the eyebrows is characteristic. The mucous membranes are frequently involved, resulting in nasal stuffiness, laryngitis and hoarseness of the voice. Nasal septal perforation with collapse of the nasal cartilages produces a saddle-nose deformity. With progression of the disease the typical leonine facies due to infiltration of the skin becomes apparent. Glove and stocking anaesthesia, gynaecomastia, testicular atrophy, ichthyosis and nerve palsies (facial, ulnar, median and radial) develop late in the disease. Neurotrophic atrophy affecting the phalanges leads to the gradual disappearance of fingers (Fig. 4.31b). Nerve involvement is less pronounced than in TT.

Lepra reactions

These are immunologically mediated acute reactions that occur in patients with the borderline or lepromatous spectrum of disease, usually during treatment. Two forms are recognized.

Non-lepromatous lepra reaction (type I lepra reaction). This is seen following treatment of patients with borderline disease; it is a type IV delayed hypersensitivity reaction. Both upgrading (or reversal) reactions (i.e. a clinical change towards a more tuberculoid form) and downgrading reactions (i.e. a change towards the lepromatous form) can occur. Neurological deficits such as an ulnar nerve palsy may occur abruptly.

Erythema nodosum leprosum (ENL; type II lepra reaction). This is a humoral antibody response to an antigen–antibody complex (i.e. a type III immune complex hypersensitivity reaction). It is seen in 50% of patients with treated LL. ENL is characterized by fever, arthralgia, iridocyclitis, crops of painful, subcutaneous erythematous nodules, and other systemic manifestations. It may last from a few days to several weeks.

Diagnosis

The diagnosis of leprosy is essentially clinical with:

- hypopigmented/reddish patches with loss of sensation
- thickening of peripheral nerves
- acid-fast bacilli (AFB) seen on skin-slit smears/biopsy. Small slits are made in pinched skin and the fluid obtained is smeared on a slide and stained for AFB.

Patients should be examined for skin lesions in adequate natural light. Occasionally nerve biopsies are helpful. Detection of *M. leprae* DNA is possible in all forms of leprosy using the polymerase chain reaction, and can be used to assess the efficacy of treatment.

Management

Multidrug therapy (MDT) is essential because of developing drug resistance (up to 40% of bacilli in some areas are resistant to dapsone). Much shorter courses of treatment are now being used: the current WHO recommended drug regimens for leprosy are shown in Box 4.11 but longer therapy may be required in severe cases. Follow up including skin smears is obligatory.

Patient education is essential. Patients should be taught self-care of their anaesthetic hands and feet to prevent ulcers. If ulcers develop, no weight-bearing should be permitted. Cheap canvas shoes with cushioned insoles are protective.

Leprosy should be treated in specialist centres with adequate physiotherapy and occupational therapy support. Surgery and physiotherapy also play a role in the management of trophic ulcers and deformities of the hands, feet and face.

Treatment of lepra reactions. This is urgent, as irreversible eye and nerve damage can occur. Antileprosy therapy must be continued. Type II lepra reactions (ENL) can be treated with analgesics, chloroquine, clofazimine and antipyretics. Thalidomide was prevously used for ENL and is effective. However, because of its teratogenicity, it is no longer recommended by WHO. Prednisolone 40–60 mg daily for 3–6 months is effective in type I reactions.

Prevention

The prevention and control of leprosy depends on rapid treatment of infected patients, particularly those with LL and BL, to decrease the bacterial reservoir.

Box 4.11 Recommended treatment regimens for leprosy in adults (modified WHO guidelines)

Multibacillary leprosy (LL, BL, BB)

Rifampicin 600 mg once-monthly, supervised
Clofazimine 300 mg once-monthly, supervised
Clofazimine 50 mg daily, self-administered
Dapsone 100 mg daily, self-administered
Treatment continued for 12 months

Paucibacillary leprosy (BT, TT)

Rifampicin 600 mg once-monthly, supervised
Dapsone 100 mg daily, self-administered
Treatment continued for 6 months

Single lesion paucibacillary leprosy

Rifampicin 600 mg
Ofloxacin 400 mg
Minocycline 100 mg
} as a single dose

LL, lepromatous; BL, borderline lepromatous; BB, borderline; BT, borderline tuberculoid; TT, tuberculoid

Fig. 4.30 **Multiple asymmetrical hypopigmented anaesthetic patches.** Courtesy of Dr P Matondo, Zambia.

(a)

(b)

Fig. 4.31 **Leprosy.** (a) Claw hand due to median and ulnar nerve damage. (b) Hands showing neurotrophic atrophy.

Anthrax

Anthrax is caused by *Bacillus anthracis*. The spores of these Gram-positive bacilli are extremely hardy and withstand extremes of temperature and humidity. The organism is capable of toxin production and this property correlates most closely with its virulence. The disease occurs world-wide. Epidemics have been reported in The Gambia, in both North and South America and in southern Europe. Transmission is through direct contact with an infected animal; infection is most frequently seen in farmers, butchers, and dealers in wool and animal hides. Spores can also be ingested or inhaled. There have been cases in the USA due to the deliberate release of anthrax spores as a bioterrorist weapon (see p. 910).

Clinical features

The incubation period is 1–10 days. Cutaneous anthrax is the most common. The small, erythematous, maculopapular lesion is initially painless. It may subsequently vesiculate and ulcerate, with formation of a central black eschar. The illness is self-limiting in the majority of patients, but occasionally perivesicular oedema and regional lymphadenopathy may be marked, and toxaemia can occur.

Inhalational anthrax (woolsorter's disease) follows inhalation of spores, and bioterrorism should be suspected. A febrile illness is accompanied by non-productive cough and retrosternal discomfort; pleural effusions are common. Untreated, the mortality is about 90%, and in the cases in the USA it was 45% despite treatment.

Gastrointestinal anthrax is due to consumption of undercooked, contaminated meat. It presents as severe gastroenteritis; haematemesis and bloody diarrhoea can occur. Toxaemia, shock and death may follow.

Diagnosis

The diagnosis is established by demonstrating the organism in smears from cutaneous lesions or by culture of blood and other body fluids. Serological confirmation can be made using ELISAs detecting antibodies to both the organism and a toxin.

Management

Ciprofloxacin is considered the best treatment. In mild cutaneous infections, oral therapy for 2 weeks is adequate but therapy for 60 days was used in the recent outbreaks in the USA. In more severe infections high doses of intravenous antibiotics are needed, along with appropriate supportive care. Any suspected case should be reported to the relevant authority.

Control

Any infected animal that dies should be burned and the area in which it was housed disinfected. Where animal husbandry is poor, mass vaccination of animals may prevent widespread contamination, but needs to be repeated annually. A human vaccine is available for those at high risk, and prophylactic antibiotics may be indicated following exposure. Some countries are establishing public health policies to deal with the deliberate release of anthrax spores.

Mycobacterial ulcer (Buruli ulcer)

Buruli ulcer is seen in humid rural areas of the tropics, especially in Africa. *Mycobacterium ulcerans* is the causative organism. The mode of transmission is thought to be via water bugs, infected by mycobacterium, found in pools and muddy fields. A small subcutaneous nodule at the site of infection gradually ulcerates, involving subcutaneous tissue, muscle and fascial planes. The ulcers are usually large with undermined edges and markedly necrotic bases due to mycolactone (a toxin produced by the mycobacterium). Smears taken from necrotic tissue generally reveal numerous acid-fast bacilli. Until recently the only effective treatment was wide surgical excision with skin grafts, but this is often unavailable in areas where the disease is prevalent. Recent studies using rifampicin and streptomycin have shown significant benefit in some forms of the disease, and further trials are ongoing.

Endemic treponematoses (bejel, yaws and pinta)

These diseases are found in various parts of the tropics and subtropics, mainly in impoverished rural areas. The WHO treated over 50 million cases in the 1950s and 1960s, reducing the prevalence, but subsequently there has been a resurgence of infection. The latest estimate of global prevalence is 2.5 million cases, mainly in South America and Africa (India has recently declared the eradication of yaws). Improvements in sanitation and an increase in living standards will be required to eradicate the diseases completely as organisms are transmitted by bodily contact, usually in children, the organism entering through damaged skin.

Clinical features

Yaws

Yaws (caused by *Treponema pertenue*) is the most widespread and common of the endemic treponemal diseases. After an incubation period of weeks or months a primary inflammatory reaction occurs at the inoculation site, from which organisms can be isolated. Dissemination of the organism leads to multiple papular lesions containing treponemes; these skin lesions usually involve the palms and soles. There may also be bone involvement, particularly affecting the long bones and those of the hand.

Approximately 10% of those infected go on to develop late yaws. Bony gummatous lesions may progress to cause gross destruction and disfigurement, particularly of the skull and facial bones, the interphalangeal joints and the long bones. Plantar hyperkeratosis is characteristic. Like syphilis, there may be a latent period between the early and late phases of the disease, but visceral, neurological and cardiovascular problems do not occur.

Bejel (endemic syphilis)

Bejel is seen in Africa and the Middle East. The causative organism (*Treponema endemicum*) enters through abrasions in the skin directly or by mouth-to-mouth or skin-to-skin contact indirectly. It differs from venereal syphilis in that a primary lesion is not commonly seen. The late stages resemble syphilis, but cardiological and neurological manifestations are rare.

Pinta

Pinta, caused by *Treponema carateum*, is restricted mainly to Central and South America. It is milder than the

other trepanomatoses and is confined to the skin. The primary lesion is a pruritic red papule, usually on the hand or foot. It may become scaly but never ulcerates and is generally associated with regional lymphadenopathy. In the later stages similar lesions can continue to occur for up to 1 year, associated with generalized lymphadenopathy. Eventually the lesions heal leaving hyperpigmented or depigmented patches.

Diagnosis and management

In endemic areas the diagnosis is usually clinical. The causative organism can be identified from the exudative lesions under dark-ground microscopy. Serological tests for syphilis are positive and do not differentiate between the conditions.

The treatment is with long-acting penicillin (e.g. intramuscular benzathine penicillin, 1.2 million units) given as a single dose. Doxycycline is used when penicillin is ineffective or contraindicated.

Trachoma

Trachoma, caused by the intracellular bacterium *Chlamydia trachomatis*, is the most common cause of blindness in the world. It is estimated that there are 150 million current infections, and 6 million people who have been blinded by trachoma. It is a disease of poverty which is found mainly in the tropics and the Middle East: it is entirely preventable. Trachoma commonly occurs in children, and is spread by direct transmission or by flies. Isolated infection is probably self-limiting, and it is repeated infection which leads to chronic eye disease.

Clinical features

Infection is bilateral and begins in the conjunctiva, with marked follicular inflammation and subsequent scarring. Scarring of the upper eyelid causes entropion, leaving the cornea exposed to further damage with the eyelashes rubbing against it (trichiasis). The corneal scarring that eventually occurs leads to blindness.

Trachoma may also occur as an acute ophthalmic infection in the neonate.

Diagnosis and management

The diagnosis is generally made clinically. However this is an unreliable indicator of active infection. A newly developed near-patient immunodiagnostic dipstick test may help to better target antibiotic therapy.

Systemic therapy with a single dose of azithromycin 20 mg/kg is the treatment of choice; although tetracycline ointment applied locally each day for 6–8 weeks is effective, compliance is poor. In endemic areas repeated courses of therapy are necessary. Once infection has been controlled, surgery may be required for eyelid reconstruction and for treatment of corneal opacities.

Prevention

Community health education, improvements in water supply and sanitation (pit latrines) and earlier case reporting could have a substantial impact on disease prevalence. This is reflected in the WHO 'SAFE' approach to trachoma: surgery, antibiotics, facial cleanliness, environmental improvement. A WHO target is global eradication by 2020.

FURTHER READING

Atik B, Thanh TTK, Luong VQ. Impact of annual targeted treatment on infectious trachoma and susceptibility to reinfection. *Journal of the American Medical Association* 2006; **296**: 488–497.

Britton WJ, Lockwood DNJ. Leprosy. *Lancet* 2004; **363**: 1209–1219.

Wansborough-Jones M, Phillips R. Buruli ulcer: emerging from obscurity. *Lancet* 2006; **367**: 1849–1858.

GASTROINTESTINAL INFECTIONS

Cholera

Cholera is caused by the curved, flagellated Gram-negative bacillus, *Vibrio cholerae*. The organism is killed by temperatures of 100°C in a few seconds but can survive in ice for up to 6 weeks. One major pathogenic serogroup possesses a somatic antigen (O1) with two biotypes: classical and El Tor. The El Tor biotype replaced the classical biotype as the major cause of the seventh pandemic which began in the 1960s. Infection with the El Tor biotype generally causes milder symptoms, but can still cause severe and life-threatening disease.

The fertile, humid Gangetic plains of West Bengal have traditionally been regarded as the home of cholera. However, a series of pandemics have spread the disease across the world, usually following trade routes. The seventh pandemic currently affects large areas of Asia and sub-Saharan Africa. A new serogroup (O139 Bengal) is responsible for many cases in Bangladesh, India and South East Asia.

Transmission is by the faeco-oral route. Contaminated water plays a major role in the dissemination of cholera, although contaminated foodstuffs and contact carriers may contribute in epidemics. Achlorhydria or hypochlorhydria facilitates passage of the cholera bacilli into the small intestine. Here they proliferate, elaborating an exotoxin which produces massive secretion of isotonic fluid into the intestinal lumen (see p. 306). Cholera toxin also releases serotonin (5-HT) from enterochromaffin cells in the gut, which activates a neural secretory reflex in the enteric nervous system. This may account for at least 50% of cholera toxin's secretory activity. *V. cholerae* also produces other toxins (zona occludens toxin, ZOT, and accessory cholera toxin, ACT) which contribute to its pathogenic effect.

Clinical features

The incubation period varies from a few hours to 6 days. The majority of patients with cholera have a mild illness that cannot be distinguished clinically from diarrhoea due to other infective causes. In severe cases there is abrupt onset of profuse painless diarrhoea, followed by vomiting. As the illness progresses the typical 'rice water' stool, flecked with mucus, may be seen. There is massive fluid loss, and if this is not replaced the features of hypovolaemic shock (cold clammy skin, tachycardia, hypotension and peripheral cyanosis) and dehydration (sunken eyes, hollow cheeks and a diminished urine output) appear. The patient, though apathetic, is usually lucid. Muscle cramps may be severe. Children may present with convulsions owing to hypoglycaemia.

With adequate treatment the prognosis is good, with a gradual return to normal clinical and biochemical parameters in 1–3 days.

Diagnosis

This is largely clinical. Examination of freshly passed stools may demonstrate rapidly motile organisms (although this is not diagnostic, as *Campylobacter jejuni* may give a similar appearance). A rapid dipstick test is now also available. Stool and rectal swabs should be taken for culture to confirm the diagnosis and to establish antibiotic sensitivity. Cholera should always be reported to the appropriate public health authority.

Management

The mainstay of treatment is rehydration, and with appropriate and effective rehydration therapy mortality has decreased to less than 1%. Oral rehydration is usually adequate, but intravenous therapy is occasionally required.

Oral rehydration solutions (ORS) are based on the observation that glucose (and other carbohydrates) enhance sodium and water absorption in the small intestine, even in the presence of secretory loss due to toxins. Additions such as amylase-resistant starch to glucose-based ORS have been shown to increase the absorption of fluid. Cereal-based electrolyte solutions have been found to be as effective as sugar/salt ORS, and actually reduce stool volume as well as rehydrating. The WHO recommends the use of reduced osmolarity ORS for all types of diarrhoea, although concerns remain about the risk of hyponatraemia. Suitable solutions for rehydration are listed in Box 4.9 (see p. 132).

Mildly dehydrated individuals are given ORS 50 mL/kg in the first 4 hours, followed by a maintenance dose of 100 mL/kg daily until the diarrhoea stops. For moderate dehydration, ORS 100 mL/kg is given within the first 4 hours followed by 10–15 mL/kg/hour.

Intravenous rehydration is required only for severely dehydrated individuals with features of collapse. Several litres of intravenous fluid (Ringers lactate) are usually required to overcome the features of shock. Maintenance of hydration is effectively carried out by oral rehydration solutions.

Three-day courses of tetracycline or erythromycin have been used to help to eradicate the infection, decrease stool output, and shorten the duration of the illness. Drug resistance is becoming an increasing problem, and a single dose of ciprofloxacin or azithromycin (each 1 g) is a better choice if available.

Immunization is now recommended by the WHO in potential or actual outbreak situations. Live attenuated and killed vaccine (both oral) are available: neither protect against the O139 strain. Chemoprophylaxis with tetracycline 500 mg twice daily for 3 days for adults, or 125 mg daily for children, is effective if the organism is sensitive, but may encourage the spread of resistance. The best preventative measures, however, are good hygiene and improved sanitation.

Enteric fever

Over 17 million new cases of enteric fever occur world-wide, mainly in India and Africa, causing 600 000 deaths per year. Enteric fever is an acute systemic illness characterized by fever, headache and abdominal discomfort. *Typhoid*, the typical form of enteric fever, is caused by *Salmonella typhi*. A similar but generally less severe illness known as paratyphoid is due to infection with *S. paratyphi* A, B or C. Man is the only natural host for *S. typhi*, which is transmitted in contaminated food or water. The incubation period is 10–14 days.

Clinical features

After ingestion, the bacteria invade the small bowel wall via Peyer's patches, from where they spread to the regional lymph nodes and then to the blood. The onset of illness is insidious and non-specific, with intermittent fever, headache and abdominal pain. Physical findings in the early stages include abdominal tenderness, hepatosplenomegaly, lymphadenopathy and a scanty maculopapular rash ('rose spots'). Without treatment (and occasionally even after treatment) serious complications can arise, usually in the third week of illness. These include meningitis, lobar pneumonia, osteomyelitis, intestinal perforation and intestinal haemorrhage. The fourth week of the illness is characterized by gradual improvement, but in developing countries up to 30% of those infected will die, and 10% of untreated survivors will relapse. This compares with a mortality rate of 1–2% in the USA.

After clinical recovery 5–10% of patients will continue to excrete *S. typhi* for several months: these are termed convalescent carriers. Between 1% and 4% will continue to carry the organism for more than a year: this is chronic carriage. The usual site of carriage is the gall bladder, and chronic carriage is associated with the presence of gallstones. However, in parts of the Middle East and Africa where urinary schistosomiasis is prevalent, chronic carriage of *S. typhi* in the urinary bladder is also common.

Diagnosis

The definitive diagnosis of enteric fever requires the culture of *S. typhi* or *S. paratyphi* from the patient. Blood culture is positive in most cases in the first 2 weeks. Culture of intestinal secretions, faeces and urine is also used, although care must be taken to distinguish acute infection from chronic carriage. Bone marrow culture is more sensitive than blood culture, but is rarely required except in patients who have already received antibiotics. Leucopenia is common but non-specific. Serological tests such as the Widal antigen test are of little practical value, are easily misinterpreted, and should not be used.

Management

Increasing antibiotic resistance is seen in isolates of *S. typhi*, especially in the Indian subcontinent. Chloramphenicol, co-trimoxazole and amoxicillin may all still be effective in some cases, but quinolones (e.g. ciprofloxacin 500 mg twice daily) are now the treatment of choice, although increased resistance to these agents is being seen: in such cases azithromycin may be effective. (Infection from the Indian subcontinent, Middle East and South East Asia is often resistant to multiple antibacterial agents.) The patient's temperature may remain elevated for several days after starting antibiotics, and this alone is not a sign of treatment failure. Prolonged antibiotic therapy may eliminate the carrier state, but in the presence of gall bladder disease it is rarely effective. Cholecystectomy is not usually justified on clinical or public health grounds.

Prevention

This is mainly through improved sanitation and clean water. Travellers should avoid drinking untreated water, **ice in drinks** and eating ice creams. Vaccination with injectable

inactivated or oral live attenuated vaccines gives partial protection.

SYSTEMIC INFECTIONS

Tuberculosis

Tuberculosis is caused by *Mycobacterium tuberculosis*, and occasionally *M. bovis* or *M. africanum*. These are slow-growing bacteria, and unlike other mycobacteria, are facultative intracellular organisms. The prevalence of tuberculosis increases with poor social conditions, inadequate nutrition and overcrowding. In developing countries it is most commonly acquired in childhood.

The impact of tuberculosis in the developing world has been magnified in the past 20 years by the emergence of the HIV pandemic (see p. 184) (Fig. 4.32).

Widespread misuse of antibiotics, combined with the breakdown of healthcare systems in parts of Africa, Russia and East Europe, has led to the emergence of drug-resistant tuberculosis. Multidrug-resistant tuberculosis (MDRTB) is caused by bacteria that are resistant to both rifampicin and isoniazid, two drugs which form the mainstay of treatment. It is now widespread in many parts of the world, including Asia, Eastern Europe and Africa. Extensively drug resistant TB (XDRTB) is additionally resistant to quinolones and injectable second line agents. MDRTB, and especially XDRTB, are very difficult to treat, and carry significant mortality even with the best medical care (see p. 866).

In most people, the initial primary tuberculosis is asymptomatic or causes only a mild illness. The focus of the disease heals.

Occasionally the primary infection progresses locally to a more widespread lesion. Haematogenous spread at this stage may give rise to miliary tuberculosis.

Tuberculosis in the adult may be the result of reactivation of old disease (post-primary tuberculosis), primary infection, or more rarely reinfection.

Pulmonary tuberculosis is the most common form; this is described on page 863, along with the chemotherapeutic regimens. Tuberculosis also affects other parts of the body:

- The gastrointestinal tract, mainly the ileocaecal area, but occasionally the peritoneum, producing ascites (see p. 318).
- The genitourinary system. The kidneys are most commonly involved, but tuberculosis can also cause painless, craggy swellings in the epididymis, and salpingitis, tubal abscesses and infertility in females.
- The central nervous system, causing tuberculous meningitis and tuberculomas (see p. 1155).
- The skeletal system, causing septic arthritis and osteomyelitis.
- The skin, giving rise to lupus vulgaris.
- The eyes, where it can cause choroiditis or iridocyclitis.
- The pericardium, producing constrictive pericarditis (see p. 795).
- The adrenal glands, causing destruction and producing Addison's disease.
- Lymph nodes. This is a common mode of presentation, especially in young adults and children. Any group of lymph nodes may be involved, but hilar and paratracheal lymph nodes are the most common. Initially the nodes are firm and discrete but later they become matted and can suppurate with sinus formation. Scrofula is the term used to describe massive cervical lymph node enlargement with discharging sinuses. Mycobacterial lymph node disease may also be caused by non-tuberculous mycobacteria.

FURTHER READING

Bhan MK, Bahl R, Bhatnagar S. Typhoid and paratyphoid fever. *Lancet* 2005; **366**: 749–762.

Sack DA, Sack RB, Chaignant C-L. Getting serious about cholera. *New England Journal of Medicine* 2006; **355**: 649–652.

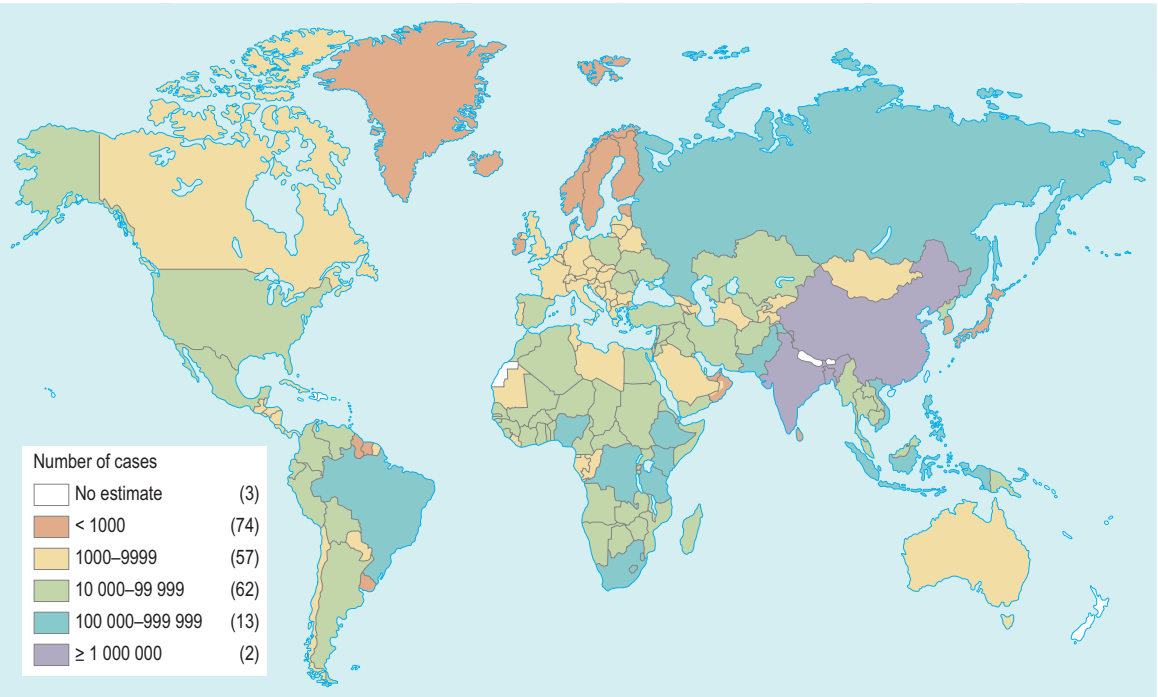


Fig. 4.32 **Tuberculosis – geographical distribution.** From Frieden TR, Sterling TR, Munsiff SS et al. Tuberculosis. *Lancet* 2003; **362**: 888, with permission from Elsevier.

Non-tuberculous mycobacterial infections

The majority of mycobacterial species are environmental organisms, and are rarely pathogenic. Some have been found to cause disease in man, particularly in immunocompromised patients or those with pre-existing chronic lung disease (Table 4.34).

Plague

Plague is caused by *Yersinia pestis*, a Gram-negative bacillus. Sporadic cases of plague (as well as occasional epidemics) occur world-wide: about 2000 cases per year are reported to the WHO, with a 10% mortality. The majority of cases are seen in sub-Saharan Africa, although the disease is occasionally seen in developed countries in people undertaking outdoor pursuits. The main reservoirs are woodland rodents, which transmit infection to domestic rats (*Rattus rattus*). The usual vector is the rat flea, *Xenopsylla cheopis*. These fleas bite humans when there is a sudden decline in the rat population. Occasionally, spread of the organisms may be through infected faeces being rubbed into skin wounds, or through inhalation of droplets.

Clinical features

Four clinical forms are recognized: bubonic, pneumonic, septicaemic and cutaneous.

Bubonic plague

This is the most common form and occurs in about 90% of infected individuals. The incubation period is 2–7 days. The onset of illness is acute, with high fever, chills, headache, myalgia, nausea, vomiting and, when severe, prostration. This is rapidly followed by the development of lymphadenopathy (buboes), most commonly involving the inguinal region. Characteristically these are matted and tender, and suppurate in 1–2 weeks.

Pneumonic plague

This is characterized by the abrupt onset of features of a fulminant pneumonia with bloody sputum, marked respiratory distress, cyanosis and death in almost all affected patients.

Septicaemic plague

This presents as an acute fulminant infection with evidence of shock and disseminated intravascular coagulation (DIC). If left untreated, death usually occurs in 2–5 days. Lymphadenopathy is unusual.

Cutaneous plague

This presents either as a pustule, eschar or papule or an extensive purpura, which can become necrotic and gangrenous.

Diagnosis

This is based on clinical and epidemiological and laboratory findings. Microscopy (on blood or lymph node aspirate), or a rapid antigen detection test can provide a presumptive diagnosis in an appropriate clinical setting. Blood or lymph node culture, or paired serological tests, are required for confirmation.

Management

Treatment is urgent and should be instituted before the results of culture studies are available. The treatment of choice is intramuscular streptomycin 1 g twice daily or gentamicin 1 mg/kg i.v. three times daily for 10 days. Oral tetracycline 500 mg four times daily and chloramphenicol are also effective.

Prevention

Prevention of plague is largely dependent on the control of the flea population. Outhouses, or huts, should be sprayed with insecticides that are effective against the local flea. During epidemics rodents should not be killed until the fleas are under control, as the fleas will leave dead rodents to bite humans. Tetracycline 500 mg four times daily for 7 days is an effective chemoprophylactic agent. A partially effective formalin-killed vaccine is available for use by travellers to plague-endemic areas.

Relapsing fevers

These conditions are so named because, after apparent recovery from the initial infection, one or more recurrences may occur after a week or more without fever. They are caused by spirochaetes of the genus *Borrelia*.

Clinical features

Louse-borne relapsing fever (caused by *B. recurrentis*) is spread by body lice, and only humans are affected. Classically it is an epidemic disease of armies and refugees,

Table 4.34 Non-tuberculous mycobacteria causing disease in man

Clinical	Common cause	Rare cause
Chronic lung disease	*Mycobacterium avium-intracellulare* *M. kansasii*	*M. malmoense* *M. xenopi*
Local lymphadenitis	*M. avium-intracellulare* *M. scrofulaceum*	*M. malmoense* *M. fortuitum*
Skin and soft tissue infection		
Fish tank granuloma	*M. marinum*	
Abscesses, ulcers, sinuses	*M. fortuitum* *M. chelonae*	*M. haemophilum*
Bone and joint infection	*M. kansasii* *M. avium-intracellulare*	*M. scrofulaceum*
Disseminated infection (in HIV)	*M. avium-intracellulare*	

More online
www.studentconsult.com

although it is also endemic in the highlands of Ethiopia, Yemen and Bolivia. Lice are spread from person to person when humans live in close contact in impoverished conditions. Infected lice are crushed by scratching, allowing the spirochaete to penetrate through the skin. Symptoms begin 3–10 days after infection and consist of a high fever of abrupt onset with rigors, generalized myalgia and headache. A petechial or ecchymotic rash may be seen. The general condition then deteriorates, with delirium, hepatosplenomegaly, jaundice, haemorrhagic problems and circulatory collapse. Although complete recovery may occur at this time, the majority experience one or more relapses of diminishing intensity over the weeks following the initial illness. The severity of the illness varies enormously, and some cases have only mild symptoms. However, in some epidemics mortality has exceeded 50%.

Tick-borne relapsing fever is caused by *B. duttoni* and other *Borrelia* species, spread by soft (argasid) ticks. Rodents are also infected, and humans are incidental hosts, acquiring the spirochaete from the saliva of the infected tick. This disease is mainly found in countries where traditional mud huts are the form of shelter, and is a common cause of febrile illness in parts of Africa. The illness is generally similar to the louse-borne disease, although neurological involvement is more common.

Diagnosis and management

Spirochaetes can be demonstrated microscopically in the blood during febrile episodes: organisms are more numerous in louse-borne relapsing fever. Treatment is usually with tetracycline or doxycycline (see p. 96). A severe Jarisch–Herxheimer reaction (see p. 179) occurs in many patients, often requiring intensive nursing care and intravenous fluids.

Prevention

Control of infection relies on elimination of the vector. Ticks live for years and remain infected, passing the infection to their progeny. These reservoirs of infection should be controlled by spraying houses with insecticides and by reducing the number of rodents. Patients infested with lice should be deloused by washing with a suitable insecticide. All clothes must be thoroughly disinfected.

Typhus

Typhus is the collective name given to a group of diseases caused by *Rickettsia* species (Table 4.35). *Rickettsiae* (and the closely related *Orientiae*) are small intracellular bacteria that are spread to humans by arthropod vectors, including body lice, fleas, hard ticks and larval mites. Rickettsiae inhabit the alimentary tract of these arthropods and the disease is spread to the human host by inoculation of their faeces through broken human skin, generally produced by scratching. Rickettsiae multiply intracellularly and can enter most mammalian cells, although the main lesion produced is a vasculitis due to invasion of endothelial cells of small blood vessels. Multisystem involvement is usual.

Clinical features

Typhus fever group

Epidemic typhus

The vector of epidemic typhus is the human body louse, and like louse-borne relapsing fever, epidemics are associated with war and refugees. Outbreaks have occurred in Africa, Central and South America, and Asia.

The incubation period of 1–3 weeks is followed by an abrupt febrile illness associated with profound malaise and generalized myalgia. Headache is severe and there may be conjunctivitis with orbital pain. A measles-like eruption appears around the fifth day. At the end of the first week, signs of meningoencephalitis appear and CNS involvement may progress to coma. At the height of the illness, splenomegaly, pneumonia, myocarditis and gangrene at the peripheries may be evident. Oliguric renal failure occurs in fulminating disease, which is usually fatal. Recovery begins in the third week but is generally slow. The disease may recur many years after the initial attack owing to rickettsiae that lie dormant in lymph nodes. The recrudescence is known as Brill–Zinsser disease. The factors that precipitate recurrence are not clearly defined, although other infections may play a role.

Endemic (murine) typhus

This is an infection of rodents that is inadvertently spread to humans by rat fleas. The disease closely resembles epidemic typhus but is much milder and rarely fatal.

Scrub typhus

Found throughout Asia and the Western Pacific, this disease is spread by larval trombiculid mites (chiggers). An eschar (a black, crusted, necrotic papule) can often be found at the site of the bite. The clinical illness is very variable, ranging from a mild illness to fulminant and potentially fatal disease.

Table 4.35 Infections caused by rickettsiae

Disease	Organism	Reservoir	Vector
Typhus fever group			
Epidemic typhus	*Rickettsia prowazekii*	Man	Human body louse
Endemic (murine) typhus	*R. typhi*	Rodents	Rat flea
Scrub typhus	*Orientia tsutsugamushi*	Trombiculid mite	Trombiculid mite
Spotted fever group			
African tick typhus	*R. africae*	Various mammals	Hard tick
Mediterranean spotted fever	*R. conorii*	Rodents, dog	Hard tick
Rocky Mountain spotted fever	*R. rickettsii*	Rodents, dog	Hard tick
Rickettsial pox	*R. akari*	Rodents	Mite
Flea-borne spotted fever	*R. felis*	Various mammals	Flea

The more severe cases resemble epidemic typhus. Unlike other types of typhus the organism is passed on to subsequent generations of mites, which consequently act as both reservoir and vector.

Spotted fever group

A variety of *Rickettsia* species, collectively known as the spotted fever group rickettsiae, cause the illnesses known as spotted fevers. In all except for rickettsial pox (which is transmitted by a rodent mite) the vector is a hard tick. Although the causative organism and the name of the illness vary from place to place the clinical course is common to all. After an incubation period of 4–10 days an eschar may develop at the site of the bite in association with regional lymphadenopathy. There is abrupt onset of fever, myalgia and headache, accompanied by a maculopapular rash which may become petechial. Neurological, haematological and cardiovascular complications occur as in epidemic typhus, although these are uncommon.

Diagnosis

The diagnosis is generally made on the basis of the history and clinical course of the illness. It can be confirmed serologically or by PCR.

Treatment and prevention

Doxycycline or tetracycline given for 5–7 days is the treatment of choice. Ciprofloxacin is also effective. Doxycycline 200 mg weekly protects against scrub typhus; it is reserved for highly endemic areas. Rifampicin is also used when resistance to tetracycline has occurred. Seriously ill patients need intensive care. Control of typhus is achieved by eradication of the arthropod vectors. Lice and fleas can be eradicated from clothing by insecticides (0.5% malathion or DDT). Control of rodents is necessary in endemic typhus and some of the spotted fevers. Areas of vegetation infested with trombiculid mites can be cleared by chemical spraying from the air. Bites from ticks and mites should be avoided by wearing protective clothing on exposed areas of the body. The likelihood of infection from ticks is related to the duration of feeding, and in high-risk areas the body should be inspected twice a day as the bites are painless, and any ticks should be removed (see p. 139).

Bartonellosis and ehrlichiosis

Bartonella spp. and *Ehrlichia* spp. are intracellular bacteria closely related to the rickettsiae. A number of human diseases can be caused by these organisms; like rickettsial disease, infection is usually spread from animals via an arthropod vector (Table 4.36).

Carrión's disease

This disease is restricted mainly to the habitat of its main vector, the sandfly, in the river valleys of the Andes mountains at an altitude of 500–3000 m. Two clinical presentations are seen, which may occur alone or consecutively. *Oroya fever* is an acute febrile illness causing myalgia, arthralgia, severe headache and confusion, followed by a haemolytic anaemia. *Verruga peruana* consists of eruptions of reddish-purple haemangiomatous nodules, resembling bacillary angiomatosis. It may follow 4–6 weeks after Oroya fever, or be the presenting feature of infection. Spontaneous resolution may occur over a period of months or years. Carrión's disease is frequently complicated by superinfection, especially with *Salmonella* spp.

The diagnosis is made by culturing bacilli from blood or peripheral lesions. Serological tests have been developed but are not widely available.

Treatment with chloramphenicol or tetracycline is very effective in acute disease, but less so in verruga peruana.

Cat-scratch disease and bacillary angiomatosis

These are described on page 125.

Trench fever

Trench fever is caused by *Bartonella quintana*, and transmitted by human body lice. It is mainly seen in refugees and the homeless. It is characterized by cyclical fever (typically every 5 days), chills and headaches, accompanied by myalgia and pretibial pain. The disease is usually self-limiting but it can be treated with erythromycin or doxycycline if symptoms are severe.

Melioidosis

The term melioidosis refers to infections caused by the Gram-negative bacteria *Burkholderia pseudomallei*. This environmental organism, which is found in soil and surface water, is distributed widely in the tropics and subtropics. The majority of clinical cases of melioidosis occur in South East Asia. Infection follows inhalation or direct inoculation. More than half of all patients with melioidosis have predisposing underlying disease: it is particularly common in diabetics.

Table 4.36 Human infections caused by *Bartonella* spp. and *Ehrlichia* spp.

Disease	Organism	Reservoir	Vector
Bartonella			
Carrión's disease	*Bartonella bacilliformis*	Unknown	Sandfly
Cat scratch disease	*B. henselae*	Cat	Cat flea
Bacillary angiomatosis	*B. henselae*	Cat	Cat flea
Trench fever	*B. quintana*	Human	Body louse
Ehrlichia			
Human ehrlichiosis	*Ehrlichia chafeensis*	Deer	Hard ticks
'Human granulocytic ehrlichia' (HE)	*Anaplasma phagocytophilum**	Small mammals and deer	Hard ticks

*Formerly known as *Ehrlichia phagocytophilia*.

B. pseudomallei causes a wide spectrum of disease, and the majority of infections are probably subclinical. Illness may be acute or chronic, localized or disseminated, but one form of the disease may progress to another and individual patients may be difficult to categorize. The most serious form is septicaemic melioidosis, which is often complicated by multiple metastatic abscesses: this is frequently fatal. Serological tests are available, but definitive diagnosis depends on isolating the organism from blood or appropriate tissue. *B. pseudomallei* has extensive intrinsic antibiotic resistance. The most effective agent is ceftazidime, which is given intravenously for 2–4 weeks; this should be followed by several months of co-amoxiclav to prevent relapses.

Actinomycosis

Actinomyces spp. are Gram-positive, branching higher bacteria which are normal mouth and intestine commensals; they are particularly associated with poor mouth hygiene. *Actinomyces* have a world-wide distribution but are a rare cause of disease in the West.

Clinical features

- *Cervicofacial actinomycosis*, the most common form, usually occurs following dental infection or extraction. It is often indolent and slowly progressive, associated with little pain, and results in induration and localized swelling of the lower part of the mandible. Sinuses and tracts develop with discharge of 'sulphur' granules.
- *Thoracic actinomycosis* follows inhalation of organisms, usually into a previously damaged lung. The clinical picture is not distinctive and is often mistaken for malignancy or tuberculosis. Symptoms such as fever, malaise, chest pain and haemoptysis are present. Empyema occurs in 25% of patients and local extension produces chest-wall sinuses with discharge of 'sulphur' granules.
- *Abdominal actinomycosis* most frequently affects the caecum. Characteristically, a hard indurated mass is felt in the right iliac fossa. Later, sinuses develop. The differential diagnosis includes malignancy, tuberculosis, Crohn's disease and amoeboma. The incidence of pelvic actinomycosis appears to be increasing with wider use of intrauterine contraceptive devices.

Occasionally actinomycosis becomes disseminated to involve any site.

Diagnosis and management

Diagnosis is by microscopy and culture of the organism. Treatment often involves surgery as well as antibiotics: penicillin is the drug of choice. Intravenous penicillin 2.4 g 4-hourly is given for 4–6 weeks, followed by oral penicillin for some weeks after clinical resolution. Tetracyclines are also effective.

Nocardia infections

Nocardia spp. are Gram-positive branching bacteria, which are found in soil and decomposing organic matter. *N. asteroides*, and less often *N. brasiliensis*, are the main human pathogens.

Clinical features

Mycetoma is the most common illness. This is a result of local invasion by *Nocardia* spp. and presents as a painless swelling, usually on the sole of the foot (Madura foot). The swelling of the affected part of the body continues inexorably. Nodules gradually appear which eventually rupture and discharge characteristic 'grains', which are colonies of organisms. Systemic symptoms and regional lymphadenopathy are rare. Sinuses may occur several years after the onset of the first symptom. A similar syndrome may be produced by other branching bacteria, and also by species of eumycete fungi such as *Madurella mycetomi* (see p. 153).

Pulmonary disease, which follows inhalation of the organism, presents with cough, fever and haemoptysis: it is usually seen in the immunocompromised. Pleural involvement and empyema occur. In severely immunosuppressed patients initial pulmonary infection may be followed by disseminated disease.

Diagnosis and management

The diagnosis is often difficult to establish, as *Nocardia* is not easily detected in sputum cultures or on histological section. Severe pulmonary or disseminated infection may require parenteral treatment. Co-trimoxazole, linezolid, ceftriaxone and amikacin have all been used successfully, but in vitro sensitivities are variable and there is no consensus on the best treatment.

FURTHER READING

Dukes Hamilton C, Sterling T, Blumberg H et al. Extensively drug resistant tuberculosis. *Clinical Infectious Diseases* 2007; **45**: 338–342.

Peacock SJ. Melioidosis. *Current Opinion in Infectious Diseases* 2006; **19**: 421–428.

Prentice MB, Rahalison L. Plague. *Lancet* 2007; **369**: 1196–1207.

FUNGAL INFECTIONS

Morphologically, fungi can be grouped into three major categories:

- yeasts and yeast-like fungi, which reproduce by budding
- moulds, which grow by branching and longitudinal extension of hyphae
- dimorphic fungi, which behave as yeasts in the host but as moulds in vitro (e.g. *Histoplasma capsulatum* and *Sporothrix schenckii*).

Despite the fact that fungi are ubiquitous, systemic fungal infections are uncommon which contrasts to superficial fungal infections of the skin, nails and orogenital mucosae (see p. 1232). Systemic mycoses are increasing in selected populations of patients such as the immunocompromised (e.g. on corticosteroids) and transplant patients, those managed in high-dependency units and in those with severe immunodeficiency, including HIV. Fungal infections are transmitted by inhalation of spores or by contact with the skin. Fungi do not produce endotoxin, but exotoxin (e.g. aflatoxin) production has been documented in vitro. Fungi may also produce allergic pulmonary disease. Some fungi such as *Candida albicans* are human commensals. Diseases are usually divided into systemic, subcutaneous or superficial (Table 4.37).

SYSTEMIC FUNGAL INFECTIONS

Candidiasis

Candidiasis is the most common fungal infection in humans and is predominantly caused by *Candida albicans* although

other species of *Candida* are increasingly recognized. Candida are small asexual fungi. Most species that are pathogenic to humans are normal oropharyngeal and gastrointestinal commensals. Candidiasis is found world-wide.

Clinical features

Any organ in the body can be invaded by candida, but vaginal infection and oral thrush are the most common forms. This latter is seen in the very young, in the elderly, following antibiotic therapy and in those who are immunosuppressed. Candidal oesophagitis presents with painful dysphagia. Cutaneous candidiasis typically occurs in intertriginous areas. It is also a cause of paronychia. Balanitis and vaginal infection are also common (see p. 182).

Chronic mucocutaneous candidiasis is a rare manifestation, usually occurring in children, and is associated with a T cell defect. It presents with hyperkeratotic plaque-like lesions on the skin, especially the face, and on the fingernails. It is associated with several endocrinopathies, including hypothyroidism and hypoparathyroidism. Dissemination of candidiasis may lead to haematogenous spread, with meningitis, pulmonary involvement, endocarditis or osteomyelitis.

Diagnosis and treatment

The fungi can be demonstrated in scrapings from infected lesions, tissue secretions or in invasive disease, from blood cultures.

Treatment varies depending on the site and severity of infection. Oral lesions respond to nystatin, oral amphotericin B or fluconazole. For systemic infections, parenteral therapy with amphotericin B, fluconazole, voriconazole or caspofungin is necessary. Polysymptomatic patients often complaining of widespread candidiasis have a psychiatric disorder (see p. 1194).

Table 4.37 Common fungal infections

Systemic	**Subcutaneous**
Histoplasmosis	Sporotrichosis
Cryptococcosis	Subcutaneous zygomycosis
Coccidioidomycosis	Chromoblastomycosis
Blastomycosis	Mycetoma
Zygomycosis (mucormycosis)	**Superficial**
Candidiasis	Dermatophytosis
Aspergillosis	Superficial candidiasis
Pneumocystis jiroveci (formerly *P. carinii*)	*Malassezia* infections

Histoplasmosis

Histoplasmosis is caused by *Histoplasma capsulatum*, a non-encapsulated, dimorphic fungus. Spores can survive in moist soil for several years, particularly when it is enriched by bird and bat droppings. Histoplasmosis occurs world-wide but is only commonly seen in Ohio and the Mississippi river valleys where over 80% of the population have been subclinically exposed. Transmission is mainly by inhalation of the spores, particularly when clearing out attics, barns and bird roosts or exploring caves.

Clinical features

Figure 4.33 summarizes the pathogenesis, main clinical forms and sequelae of *Histoplasma* infection. *Primary pulmonary histoplasmosis* is usually asymptomatic. The only

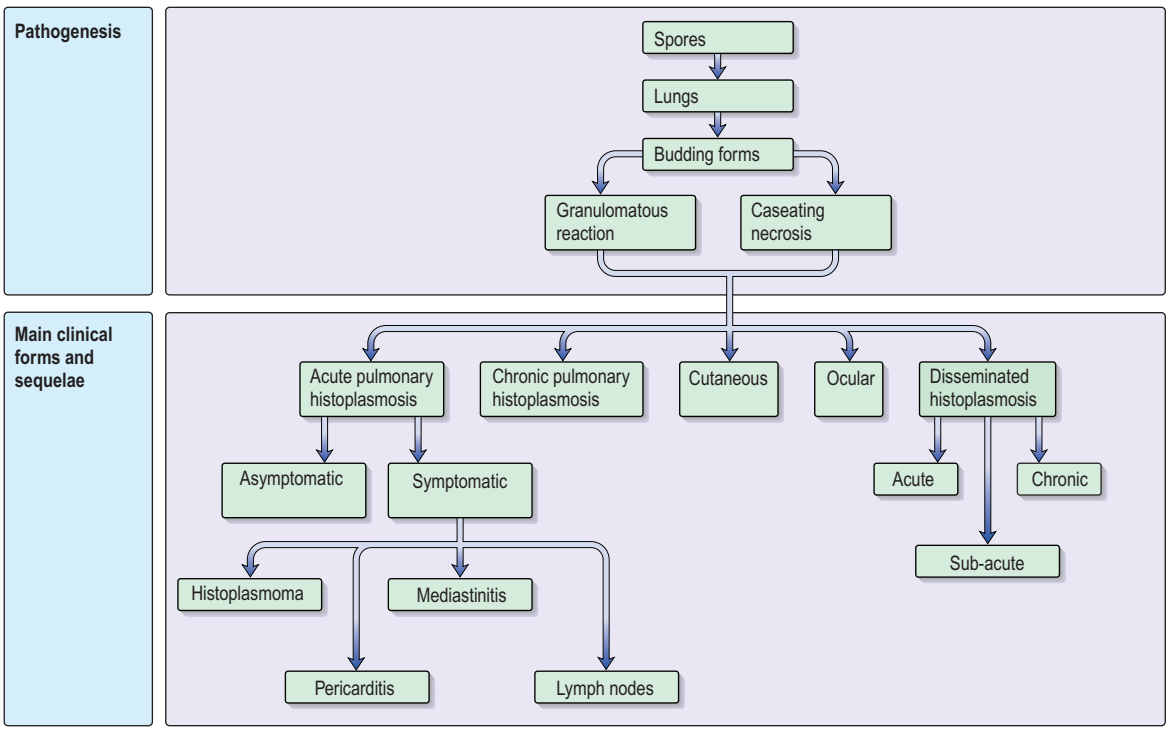


Fig. 4.33 **Histoplasma infection.** Summary of pathogenesis, main clinical forms and sequelae.

Ask the authors
www.studentconsult.com

evidence of infection is conversion of a histoplasmin skin test from negative to positive, and radiological features similar to those seen with the Ghon primary complex of tuberculosis. Calcification in the lungs, spleen and liver occurs in patients from areas of high endemicity. When symptomatic, primary pulmonary histoplasmosis generally presents as a mild influenza-like illness, with fever, chills, myalgia and cough. The systemic symptoms are pronounced in severe disease.

Complications such as atelectasis, secondary bacterial pneumonia, pleural effusions, erythema nodosum and erythema multiforme also occur.

Chronic pulmonary histoplasmosis is clinically indistinguishable from pulmonary tuberculosis (see p. 864). It is usually seen in American white males over the age of 50 years. Radiologically, pulmonary cavities, infiltrates and characteristic fibrous streaking from the periphery towards the hilum are seen.

The presentation of disseminated histoplasmosis resembles disseminated tuberculosis clinically. Fever, lymphadenopathy, hepatosplenomegaly, weight loss, leucopenia and thrombocytopenia are common. Rarely, features of meningitis, hepatitis, hypoadrenalism, endocarditis and peritonitis may dominate the clinical picture.

Diagnosis

Definitive diagnosis is possible only by culturing the fungi or by demonstrating them on histological sections. The histoplasmin skin test is of limited diagnostic value and can be negative in acute disseminated disease. *H. capsulatum* glycoprotein can be detected in the urine and serum in those with acute pulmonary and disseminated infection. Antibodies usually develop within 3 weeks of the onset of illness and are best detected by complement-fixation or immunodiffusion.

Management

Only symptomatic acute pulmonary histoplasmosis, chronic histoplasmosis and acute disseminated histoplasmosis require therapy. Itraconazole or ketoconazole are indicated for moderate disease. Severe infection is treated with intravenous amphotericin B for 1–2 weeks followed by itraconazole for a total of 12 weeks. Methylprednisolone is recommended for those who develop respiratory complications. Patients with AIDS usually require treatment with parenteral amphotericin B followed by maintenance therapy with itraconazole 200 mg twice daily where HAART is unavailable. Surgical excision of histoplasmomas (pulmonary granuloma due to *H. capsulatum*) or chronic cavitatory lung lesions and release of adhesions following mediastinitis are often required.

African histoplasmosis

This is caused by *Histoplasma duboisii*, the spores of which are larger than those of *H. capsulatum*. Skin lesions (e.g. abscesses, nodules, lymph node involvement and lytic bone lesions) are prominent. Pulmonary lesions do not occur. Treatment is similar to that for *H. capsulatum* infection.

Aspergillosis

Aspergillosis is caused by one of several species of dimorphic fungi of the genus *Aspergillus*. Of these, *A. fumigatus* is the most common, although *A. flavus* and *A. niger* are also recognized. These fungi are ubiquitous in the environment and are commonly found on decaying leaves and trees. Humans are infected by inhalation of the spores. Disease manifestation depends on the dose of the spores inhaled as well as the immune response of the host. Three major forms of the disease are recognized:

- Bronchopulmonary allergic aspergillosis (see p. 875) shows symptoms suggestive of bronchial asthma.
- Aspergilloma (see p. 876) is sometimes referred to as a pulmonary mycetoma.
- Invasive aspergillosis, which occurs in immunosuppressed patients and presents as acute pneumonia, meningitis or an intracerebral abscess, lytic bone lesions, and granulomatous lesions in the liver; less commonly endocarditis, paranasal *Aspergillus* granuloma or keratitis may occur. Urgent treatment with intravenous voriconazole or amphotericin B is required. Caspofungin provides another alternative for refractory disease.

The diagnosis and treatment are described in more detail on page 875.

Cryptococcosis

Cryptococcosis is caused by the yeast-like fungus, *Cryptococcus neoformans*. It has a world-wide distribution and appears to be spread by birds, especially pigeons, in their droppings. The spores gain entry into the body through the respiratory tract, where they elicit a granulomatous reaction. Pulmonary symptoms are, however, uncommon; meningitis which usually occurs in those with HIV or lymphoma is the usual mode of presentation and often develops subacutely. Less commonly, lung cavitation, hilar lymphadenopathy, pleural effusions and occasionally pulmonary fibrosis occur. Skin and bone involvement is rare.

Diagnosis and treatment

This is established by demonstrating the organisms in appropriately stained tissue sections. A positive latex cryptococcal agglutinin test performed on the CSF is diagnostic of cryptococcosis.

Amphotericin B (0.7–1.0 mg/kg daily i.v.) alone or in combination with flucytosine (100–200 mg/kg daily) for 2 weeks followed by fluconazole 400 mg daily. Therapy should be continued for 8 weeks if meningitis is present. Fluconazole has greater CSF penetration and is used when toxicity is encountered with amphotericin B and flucytosine and as maintenance therapy in immunocompromised patients, especially those with HIV (see p. 200).

Coccidioidomycosis

Coccidioidomycosis is caused by the non-budding spherical form (spherule) of *Coccidioides immitis*. This is a soil saprophyte and is found in the southern USA, Central America and parts of South America. Humans are infected by inhalation of the thick-walled barrel-shaped spores called arthrospores. Occasionally epidemics of coccidioidomycosis have been documented following dust storms.

Clinical features

The majority of patients are asymptomatic and the infection is only detected by the conversion of a skin test using

coccidioidin (extract from a culture of mycelial growth of *C. immitis*) from negative to positive. Acute pulmonary coccidioidomycosis presents, after an incubation period of about 10 days, with fever, malaise, cough and expectoration. Erythema nodosum, erythema multiforme, phlyctenular conjunctivitis and, less commonly, pleural effusions may occur. Complete recovery is usual.

Pulmonary cavitation with haemoptysis, pulmonary fibrosis, meningitis, lytic bone lesions, hepatosplenomegaly, skin ulcers and abscesses may occur in severe disease.

Diagnosis

Because of the high infectivity of this fungus, and consequent risk to laboratory personnel, serological tests (rather than culture of the organism) are widely used for diagnosis. These include the highly specific latex agglutination and precipitin tests (IgM) which are positive within 2 weeks of infection and decline thereafter. Other tests include complement fixation, ELISA and radioimmunoassay.

A complement-fixation test (IgG) performed on the CSF is diagnostic of coccidioidomycosis meningitis and becomes positive within 4–6 weeks and remains so for many years.

Treatment

Mild pulmonary infections are self-limiting and require no treatment, but progressive and disseminated disease requires urgent therapy. Ketoconazole, itraconazole or fluconazole for 6 months is the treatment of choice for primary pulmonary disease with more prolonged courses for cavitating or fibronodular disease. Fluconazole in high-dose (600–1000 mg daily) is given for meningitis. Itraconazole provides an alternative. Surgical excision of cavitatory pulmonary lesions or localized bone lesions may be necessary.

Blastomycosis

Blastomycosis is a systemic infection caused by the biphasic fungus *Blastomyces dermatitidis*. Although initially believed to be confined to certain parts of North America, it has been reported from South America, India and the Middle East.

Clinical features

Blastomycosis primarily involves the skin, where it presents as non-itchy papular lesions that later develop into ulcers with red verrucous margins. The ulcers are initially confined to the exposed parts of the body but later involve the unexposed parts as well. Atrophy and scarring may occur. Pulmonary involvement presents as a solitary lesion resembling a malignancy or gives rise to radiological features similar to the primary complex of tuberculosis. Systemic symptoms such as fever, malaise, cough and weight loss are usually present. Bone lesions are common and present as painful swellings.

Diagnosis and treatment

The diagnosis is confirmed by demonstrating the organism in histological sections or by culture, although results can be negative in 30–50% of cases. Enzyme immunoassay may be helpful although there is some cross-reactivity of antibodies to blastomyces with histoplasma.

Itraconazole is preferred for treating mild to moderate disease in the immunocompetent for periods up to 6 months. Ketoconazole or fluconazole are also used. In severe or unresponsive disease and in the immunocompromised, amphotericin B is indicated.

Mucormycosis

Invasive zygomycosis (mucormycosis) is rare and is caused by several fungi, including *Mucor* spp., *Rhizopus* spp. and *Absidia* spp. It occurs in severely ill patients. The hallmark of the disease is vascular invasion with marked haemorrhagic necrosis.

Rhinocerebral mucormycosis is the most common form. Nasal stuffiness, facial pain and oedema, and necrotic, black nasal turbinates are characteristic. It is rare and is mainly seen in diabetics with ketoacidosis.

Subcutaneous zygomycosis presents as a brawny, woody infiltration involving the limbs, neck and trunk and rarely the pharynx and orbital regions in immunosuppressed patients.

Other forms include pulmonary and disseminated infection (immunosuppressed) and gastrointestinal infection (in malnutrition).

Treatment is with amphotericin B and sometimes judicious debridement. Oral saturated potassium iodide has been used in the subcutaneous variety.

FURTHER READING

Galgiani JN, Ampel NM, Blair JE et al. Coccidioidomycosis. *Clinical Infectious Diseases* 2005; **41**: 1217–1223.

Pappas PG, Rex JH, Sobel JD et al. Guidelines for treatment of candidiasis. *Clinical Infectious Diseases* 2004; **38**: 161–189.

Patterson TF. Advances and challenges in management of invasive mycoses. *Lancet* 2005; **366**: 1013–1025.

Segal BH, Walsh TJ. Current approaches to the diagnosis and treatment of invasive aspergillosis. *American Journal of Respiratory and Critical Care Medicine* 2006; **173**: 707–714.

SUBCUTANEOUS INFECTIONS

Sporotrichosis

Sporotrichosis is due to the saprophytic fungus *Sporothrix schenckii*, which is found world-wide. Infection usually follows cutaneous inoculation, at the site of which a reddish, non-tender, maculopapular lesion develops – referred to as 'plaque sporotrichosis'. Pulmonary involvement and disseminated disease rarely occur.

Treatment with itraconazole 100–200 mg/day for 3–6 months is usually curative.

Subcutaneous zygomycosis

Subcutaneous zygomycosis, a disease seen in the tropics, is caused by several filamentous fungi of the *Basidiobolus* genus. The disease usually remains confined to the subcutaneous tissues and muscle fascia. It presents as a brawny, woody infiltration involving the limbs, neck and trunk. Less commonly, the pharyngeal and orbital regions may be affected in immunocompromised patients and especially those with poorly controlled diabetes mellitus. It is locally erosive and can prove fatal. Amphotericin B is the drug of choice.

Treatment is with saturated potassium iodide solution given orally.

Chromoblastomycosis

Chromoblastomycosis (chromomycosis) is caused by fungi of various genera including *Phialophora*, *Wangella* and *Fonsecaea*. These are found mainly in tropical and subtropical countries. It presents initially as a small papule, usually at the site of a previous injury. This persists for several months before ulcerating. The lesion later becomes warty and encrusted and gradually spreads. Satellite lesions may be present. Itching is frequent. The drug of choice is amphotericin B in combination with itraconazole or voriconazole. Cryosurgery is used to remove local lesions.

Mycetoma (Madura foot)

Mycetoma may be due to subcutaneous infection with fungi (*Eumycetes* spp.) or bacteria (see p. 149). It is largely confined to the tropics. Infection results in local swelling which may discharge through sinuses. Bone involvement may follow.

Treatment includes surgical debridement with ketoconazole or antibacterials, such as co-trimoxazole, according to the aetiological agent.

Pneumocystis jiroveci (formerly *P. carinii*) infection

Genetic analysis has shown this organism to be homologous with fungi. It exists as a trophozoite which is probably motile and reproduces by binary fission. After invasion the trophozoite wall thickens and forms a cyst. On maturation further division takes place to yield eight merozoites which after cell wall rupture develop into trophozoites. Infection probably occurs in infancy but in otherwise healthy infants it remains undetected. It is usually cleared from the lungs. *Pneumocystis jiroveci* disease in adults is associated with immunodeficiency states, particularly AIDS, and is discussed on page 200.

FURTHER READING

Richardson MD, Warnock DW (eds). *Fungal Infection: Diagnosis and Management,* 3rd edn. Oxford: Blackwell Science, 2003.

SUPERFICIAL INFECTIONS

Dermatophytosis

Dermatophytoses are chronic fungal infections of keratinous structures such as the skin, hair or nails. *Trichophyton* spp., *Microsporum* spp., *Epidermophyton* spp. and *Candida* spp. can also infect keratinous structures.

Malassezia infection

Malassezia spp. are found on the scalp and greasy skin and are responsible for seborrhoeic dermatitis, pityriasis versicolor (hypo- or hyperpigmented rash on trunk) and *Malassezia* folliculitis (itchy rash on back).

Treatment is with topical antifungals or oral ketoconazole if infection is refractory or more extensive.

PROTOZOAL INFECTIONS

Protozoa are unicellular eukaryotic organisms. They are more complex than bacteria, and belong to the animal kingdom. Although many protozoa are free-living in the environment some have become parasites of vertebrates, including man, often developing complex life cycles involving more than one host species. In order to be transmitted to a new host, some protozoa transform into hardy cyst forms which can survive harsh external conditions. Others are transmitted by an arthropod vector, in which a further replication cycle takes place before infection of a new vertebrate host.

BLOOD AND TISSUE PROTOZOA

Malaria

Human malaria can be caused by four species of the genus *Plasmodium*: *P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae*. Occasionally other species of malaria usually found in primates can affect man. Malaria probably originated from animal malarias in central Africa, but was spread around the globe by human migration. Public health measures and changes in land use have eradicated malaria in most developed countries, although the potential for malaria transmission still exists in many areas. Five hundred million people are infected every year, and over one million die yearly. Twenty five thousand international travellers per year are infected.

More online
www.studentconsult.com

Epidemiology

Malaria is transmitted by the bite of female anopheline mosquitoes. The parasite undergoes a temperature-dependent cycle of development in the gut of the insect, and its geographical range therefore depends on the presence of the appropriate mosquito species and on adequate temperature. The disease occurs in endemic or epidemic form throughout the tropics and subtropics except for some areas above 2000 m (Fig. 4.34). Australia, the USA and most of the Mediterranean littoral are also malaria-free. In *hyperendemic areas* (51–75% rate of parasitaemia, or palpable spleen in children 2–9 years of age) and *holoendemic areas* (>75% rate) where transmission of infection occurs year round, the bulk of the mortality is seen in infants. Those who survive to adulthood acquire significant immunity; low-grade parasitaemia is still present, but causes few symptoms. In *mesoendemic areas* (11–50%) there is regular seasonal transmission of malaria. Mortality is still mainly seen in infants, but older children and adults may develop chronic ill health due to repeated infections. In *hypoendemic areas* (0–10%), where infection occurs in occasional epidemics, little immunity is acquired and the whole population is susceptible to severe and fatal disease.

Malaria can also be transmitted in contaminated blood transfusions. It has occasionally been seen in injecting drug users sharing needles and as a hospital-acquired infection related to contaminated equipment. Rare cases are acquired outside the tropics when mosquitoes are transported from endemic areas ('airport malaria'), or when the local mosquito population becomes infected by a returning traveller.

Parasitology

The female mosquito becomes infected after taking a blood meal containing gametocytes, the sexual form of the malarial

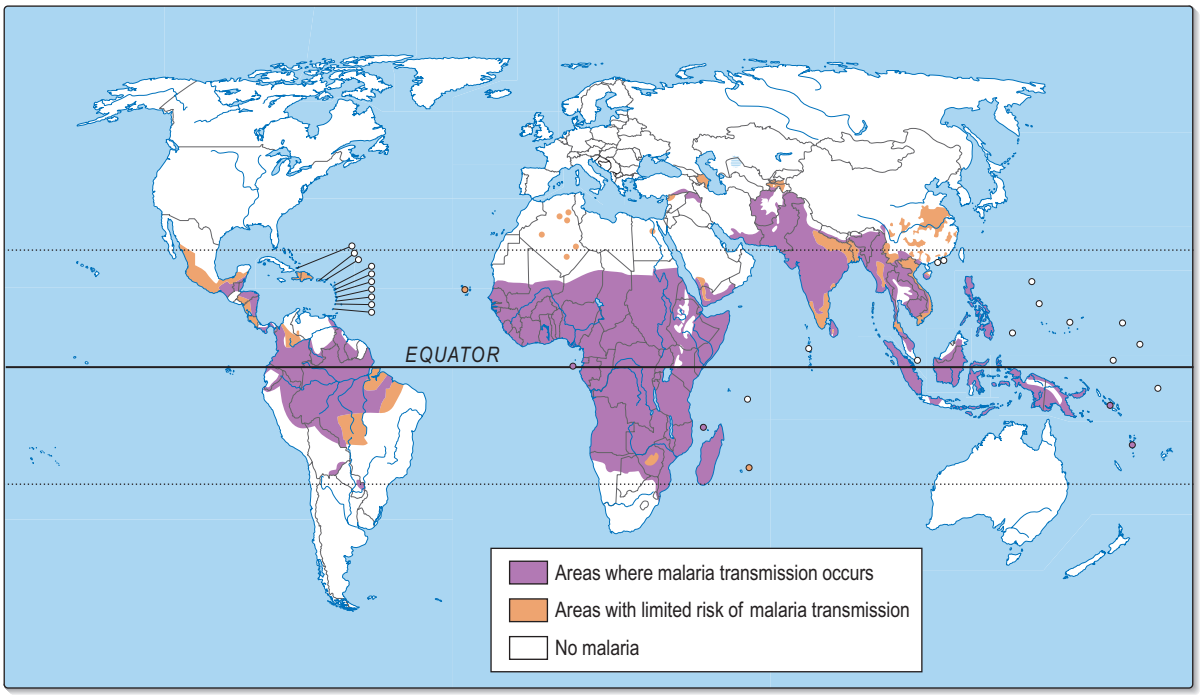


Fig. 4.34 **Malaria – geographical distribution.** Reproduced from http://www.who.int/ith/maps/malaria2007.jpg with permission of the World Health Organization.

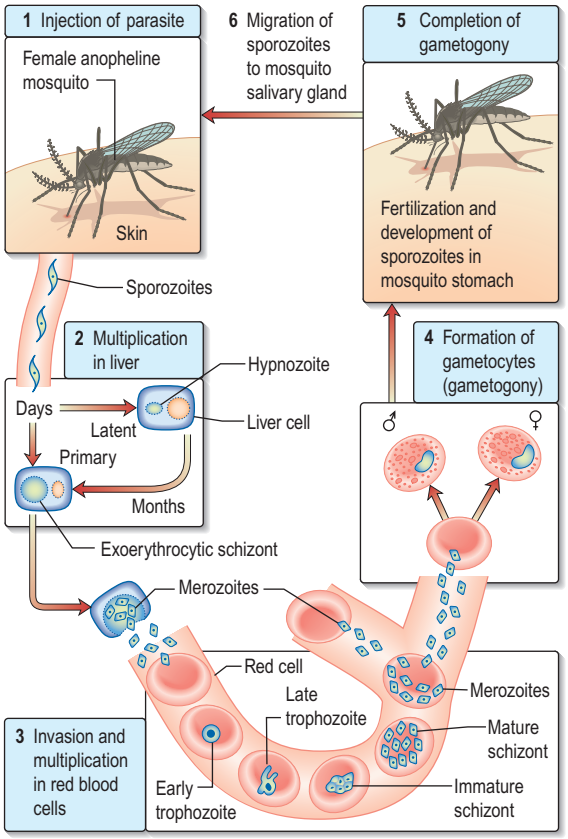


Fig. 4.35 **A schematic life cycle of *Plasmodium vivax*.**

parasite (Fig. 4.35). The developmental cycle in the mosquito usually takes 7–20 days (depending on temperature), culminating in infective sporozoites migrating to the insect's salivary glands. The sporozoites are inoculated into a new human host, and those which are not destroyed by the immune response are rapidly taken up by the liver. Here they multiply inside hepatocytes as merozoites: this is pre-erythrocytic (or hepatic) sporogeny. After a few days the infected hepatocytes rupture, releasing merozoites into the blood from where they are rapidly taken up by erythrocytes. In the case of *P. vivax* and *P. ovale*, a few parasites remain dormant in the liver as hypnozoites. These may reactivate at any time subsequently, causing relapsing infection.

Inside the red cells the parasites again multiply, changing from merozoite, to trophozoite, to schizont, and finally appearing as 8–24 new merozoites. The erythrocyte ruptures, releasing the merozoites to infect further cells. Each cycle of this process, which is called erythrocytic schizogony, takes about 48 hours in *P. falciparum*, *P. vivax* and *P. ovale*, and about 72 hours in *P. malariae*. *P. vivax* and *P. ovale* mainly attack reticulocytes and young erythrocytes, while *P. malariae* tends to attack older cells; *P. falciparum* will parasitize any stage of erythrocyte.

A few merozoites develop not into trophozoites but into gametocytes. These are not released from the red cells until taken up by a feeding mosquito to complete the life cycle.

Table 4.38 Causes of anaemia in malaria infection

Haemolysis of infected red cells
Haemolysis of non-infected red cells (blackwater fever)
Dyserythropoiesis
Splenomegaly and sequestration
Folate depletion

Pathogenesis

The pathology of malaria is related to anaemia, cytokine release, and in the case of *P. falciparum*, widespread organ damage due to impaired microcirculation. The anaemia seen in malaria is multifactorial (Table 4.38). In *P. falciparum*

malaria, red cells containing schizonts adhere to the lining of capillaries in the brain, kidneys, gut, liver and other organs. As well as causing mechanical obstruction these schizonts rupture, releasing toxins and stimulating further cytokine release.

After repeated infections partial immunity develops, allowing the host to tolerate parasitaemia with minimal ill effects. This immunity is largely lost if there is no further infection for a couple of years. Certain genetic traits also confer some immunity to malaria. People who lack the Duffy antigen on the red cell membrane (a common finding in West Africa) are not susceptible to infection with *P. vivax*. Certain haemoglobinopathies (including sickle cell trait) also give some protection against the severe effects of malaria: this may account for the persistence of these otherwise harmful mutations in tropical countries. Iron deficiency may also have some protective effect. The spleen appears to play a role in controlling infection, and splenectomized people are at risk of overwhelming malaria. Some individuals appear to have a genetic predisposition for developing cerebral malaria following infection with *P. falciparum*. Pregnant women are especially susceptible to severe disease.

Clinical features

Typical malaria is seen in non-immune individuals. This includes children in any area, adults in hypoendemic areas, and any visitors from a non-malarious region.

The normal incubation period is 10–21 days, but can be longer. The most common symptom is fever, although malaria may present initially with general malaise, headache, vomiting or diarrhoea. At first the fever may be continual or erratic: the classical tertian or quartan fever only appears after some days. The temperature often reaches 41°C, and is accompanied by rigors and drenching sweats.

P. vivax or P. ovale infection

The illness is relatively mild. Anaemia develops slowly, and there may be tender hepatosplenomegaly. Spontaneous recovery usually occurs within 2–6 weeks, but hypnozoites in the liver can cause relapses for many years after infection. Repeated infections often cause chronic ill health due to anaemia and hyperreactive splenomegaly.

P. malariae infection

This also causes a relatively mild illness, but tends to run a more chronic course. Parasitaemia may persist for years, with or without symptoms. In children, *P. malariae* infection is associated with glomerulonephritis and nephrotic syndrome.

P. falciparum infection

This causes, in many cases, a self-limiting illness similar to the other types of malaria, although the paroxysms of fever are usually less marked. However it may also cause serious complications (Box 4.12), and the vast majority of malaria deaths are due to *P. falciparum*. Patients can deteriorate rapidly, and children in particular progress from reasonable health to coma and death within hours. A high parasitaemia (>1% of red cells infected) is an indicator of severe disease, although patients with apparently low parasite levels may also develop complications. *Cerebral malaria* is marked by diminished consciousness, confusion and convulsions, often progressing to coma and death. Untreated it is universally fatal. *Blackwater fever* is due to widespread intravascular haemolysis, affecting both parasitized and unparasitized red cells, giving rise to dark urine.

Box 4.12 Some features of severe falciparum malaria

CNS

Prostration
Cerebral malaria (coma convulsion ≈ 3 seizures in 24 hours)

Renal

Haemoglobinuria (blackwater fever)
Oliguria
Uraemia (serum creatinine > 250 μmol/L) (acute tubular necrosis)

Blood

Severe anaemia (<5 g/dL) (haemolysis and dyserythropoiesis)
Disseminated intravascular coagulation (DIC)
Bleeding, e.g. retinal haemorrhages

Respiratory

Tachypnoea
Acute respiratory distress syndrome

Metabolic

Hypoglycaemia (<2mmol/L) (particularly in children)
Metabolic acidosis (blood pH < 7.25)

Gastrointestinal/liver

Diarrhoea
Jaundice (bilirubin > 50 μmol/L)
Splenic rupture

Other

Shock – hypotensive (<80 systolic pressure) and Gram-negative septicaemia
Hyperpyrexia

Hyperreactive malarial splenomegaly (tropical splenomegaly syndrome, TSS)

This is seen in older children and adults in areas where malaria is hyperendemic. It is associated with an exaggerated immune response to repeated malaria infections, and is characterized by anaemia, massive splenomegaly and elevated IgM levels. Malaria parasites are scanty or absent. TSS usually responds to prolonged treatment with prophylactic antimalarial drugs.

Diagnosis

Malaria should be considered in the differential diagnosis of anyone who presents with a febrile illness in, or having recently left, a malarious area. Falciparum malaria is unlikely to present more than 3 months after exposure, even if the patient has been taking prophylaxis, but vivax malaria may cause symptoms for the first time up to a year after leaving a malarious area.

Diagnosis is usually made by identifying parasites on a Giemsa-stained thick or thin blood film (thick films are more difficult to interpret, and it may be difficult to speciate the parasite, but they have a higher yield). At least three films should be examined before malaria is declared unlikely. Quantitative buffy coat analysis (QBC) is quicker but less accurate. Rapid antigen detection tests are available for near-patient use. In many endemic areas malaria is overdiagnosed on clinical grounds and it is necessary to make a

definite diagnosis wherever possible. Serological tests are of no diagnostic value.

Parasitaemia is common in endemic areas, and the presence of parasites does not necessarily mean that malaria is the cause of the patient's symptoms. Further investigation, including a lumbar puncture, may be needed to exclude bacterial infection.

Management

The drug of choice for **susceptible** parasites is chloroquine (Box 4.13). *P. vivax, P. ovale* and *P. malariae* are almost always sensitive to this drug, the only exception being some strains of *P. vivax* from Oceania. Following successful treatment of *P. vivax* or *P. ovale* malaria, it is necessary to give a 2- to 3-week course of primaquine (15 mg daily) to eradicate the hepatic hypnozoites and prevent relapse. This drug can precipitate haemolysis in patients with G6PD deficiency (see p. 413).

Box 4.13 Drug treatment of uncomplicated malaria in adults

Type of malaria	Drug treatment
Plasmodium vivax, P. ovale, P. malariae	Cloroquine: 600 mg 300 mg 6 hours later 300 mg 24 hours later 300 mg 24 hours later
P. falciparum (almost all are resistant to chloroquine)	Quinine: 600 mg 3 times daily for 7 days *Plus* Doxycycline: 200 mg × 1 daily for 7 days *Or* Fansidar (SP): 3 tablets as single dose *Alternative therapies* Artesunate-mefloquine 12/25 mg/kg daily for 3 days *or* Dihydroartemisinin/piperaquine 6.3 mg/kg daily for 3 days *or* Malarone: 4 tablets daily for 3 days *or* Coartemether/Riamet: 4 tablets 12-hourly for 3 days *or* Lapdap (chlorproguanil/dapsone)

SP, Fansidar = pyrimethamine/sulfadoxine; Malarone = atovaquone/proguanil; Coartemether/Riamet = artemether lamufantrine.
Chloroquine doses quoted are for base drug.
Quinine dose applies to sulphate, hydrochloride or dihydrochloride.

P. falciparum. Chloroquine resistance is now found almost world-wide and chloroquine is not recommended. Despite this, cost and availability mean that chloroquine is still the most commonly used antimalarial. In sub-Saharan Africa, older malarial drugs are now ineffective and artemisinin-based combination therapies (ACT) may be used for *P. falciparum* infection (see below).

Treatment should ideally be based on a knowledge of local sensitivity, but this is often not known. In developed countries *P. falciparum* is commonly treated with quinine, usually as quinine sulphate. In mild cases this can be given orally, but in more severe illness it is given by intravenous infusion. Tinnitus and nausea are predictable side-effects of quinine, and do not require dose reduction unless severe. Some resistance to quinine is emerging, and another antimalarial (either Fansidar (pyrimethamine/sulfadoxine) or tetracycline) should be given at the end of the course of quinine.

Other antimalarial drugs include mefloquine, artemisinin derivatives, and a number of fixed-dose combination preparations including Malarone (atovaquone/proguanil), coartemether (artemether/lamefantrine), and Lapdap (chlorproguanil/dapsone). Other (non-fixed dose) combinations are also in use: amodiaquine/sulfadoxine-pyramethamine has been shown to be highly effective in some settings. Despite a number of recent trials it is unclear which combinations are the most effective or sustainable, especially in the poorest developing countries. Combinations containing artemisinin derivatives are widely perceived to be the best treatment and are written into the national policies of many countries, but cost often prevents their use in practice.

Treatment of severe falciparum malaria. Severe malaria, indicated by the presence of any of the complications discussed above, or a parasite count above 1% in a non-immune patient, is a medical emergency (Emergency Box 4.1).

- Intravenous artesunate is at least as effective as intravenous quinine, and is a better choice where available. When quinine is used the loading dose should be omitted if the patient has already received quinine or mefloquine.
- Intensive care facilities may be needed, including mechanical ventilation and dialysis.

Emergency Box 4.1

Drug treatment of severe *Plasmodium falciparum* malaria in adults

	Full hospital facilities	No infusion available	No injection available
CQ-sensitive *P. falciparum*	Chloroquine: 10 mg/kg infused over 8 hours, followed by 15 mg/kg over 24 hours	Chloroquine: 2.5 mg/kg every 4 hours by intramuscular injection (to total of 25 mg/kg)	Chloroquine: by nasogastric tube (as in oral regimen) *or*
CQ-resistant *P. falciparum*	Quinine salt: 20 mg/kg infused over 4 hours, followed by 10 mg/kg over 4 hours every 8 hours *or* Artesunate 2.4 mg/kg by intravenous injection, then 1.2 mg/kg	Quinine dihydrochloride given by divided intramuscular injection (regimen as for i.v.)	Artemisinin rectal suppositories (limited availability)

Chloroquine (CQ) doses quoted are for base drug. Quinine doses apply to sulphate, hydrochloride and dihydrochloride.

Box 4.14 Malaria prophylaxis for adult travellers

Area visited	Prophylactic regimen	Alternative
No chloroquine resistance	Chloroquine 300 mg weekly	Proguanil 200 mg daily
Limited chloroquine resistance	Chloroquine 300 mg weekly *plus* Proguanil 200 mg daily	Doxycycline 100 mg daily *or* Malarone 1 tablet daily *or* Mefloquine 250 mg weekly
Significant chloroquine resistance	Mefloquine 250 mg weekly	Doxycycline 100 mg daily *or* Malarone 1 tablet daily

- Severe anaemia may require transfusion.
- Careful monitoring of fluid balance is essential: both pulmonary oedema and prerenal failure are common.
- Hypoglycaemia can be induced both by the infection itself and by quinine treatment.
- Superadded bacterial infection is common. In very heavy infections (parasitaemia > 10%), there may be a role for exchange transfusion, if the facilities are available.

Prevention and control

As with many vector-borne diseases, control of malaria relies on a combination of case treatment, vector eradication, and personal protection from vector bites, e.g. insecticide (permethrin) treated nets. Mosquito eradication is usually achieved either by the use of insecticides, house spraying with DDT, or by manipulation of the habitat (e.g. marsh drainage). After some initial successes, a WHO campaign to eliminate malaria foundered in the mid-1960s. Since then the emergence of both parasite resistance to drugs and mosquito resistance to insecticides has rendered the task more difficult. However, malaria is once again a priority for the WHO, which announced a new 'Roll Back Malaria' campaign in 1998. This has had good success in some countries who have co-ordinated programmes, including the use of insecticide-treated bednets and indoor residual spraying with DDT.

Non-immune travellers to malarious areas should take measures to avoid insect bites, such as using insect repellent (diethyltoluamide (DEET) 20–50% in lotions and sprays) and sleeping under mosquito nets. Antimalarial prophylaxis should also be taken in most cases, although this is never 100% effective (Box 4.14). The precise choice of prophylactic regimen depends both on the individual traveller and on the specific itinerary; further details can be found in National Formularies or from travel advice centres. Despite considerable efforts, there is still no effective vaccine available for malaria.

Trypanosomiasis

African trypanosomiasis (sleeping sickness)

Sleeping sickness is caused by trypanosomes transmitted to humans by the bite of the tsetse fly (genus *Glossina*). It is endemic in a belt across sub-Saharan Africa, extending to about 14°N and 20°S: this marks the natural range of the tsetse fly. Two subspecies of trypanosome cause human sleeping sickness: *Trypanosoma brucei gambiense* ('Gambian sleeping sickness'), and *T. b. rhodesiense* ('Rhodesian sleeping sickness').

Epidemiology

Sleeping sickness due to T. b. gambiense is found from Uganda in Central Africa, west to Senegal and south as far as Angola. Man is the major reservoir, and infection is transmitted by riverine *Glossina* species (e.g. *G. palpalis*).

Sleeping sickness due to T. b. rhodesiense occurs in East and Central Africa from Ethiopia to Botswana. It is a zoonosis of both wild and domestic animals. In endemic situations it is maintained in game animals and transmitted by savanna flies such as *G. morsitans*. Epidemics are usually related to cattle, and the vectors are riverine flies.

Political upheavals during the 1990s disrupted established treatment and control programmes, resulting in major epidemics in Angola, the Democratic Republic of Congo and Uganda. By 1997 as many as 500 000 people were affected by sleeping sickness. A concerted control programme has now brought this number down to about 70 000.

Parasitology

Tsetse flies bite during the day, and unlike most arthropod vectors both males and females take blood meals. An infected insect may deposit metacyclic trypomastigotes (the infective form of the parasite) into the subcutaneous tissue. These cause local inflammation ('trypanosomal chancre') and regional lymphadenopathy. Within 2–3 weeks the organisms invade the bloodstream, subsequently spreading to all parts of the body including the brain.

Clinical features

T. b. gambiense causes a chronic, slowly progressive illness. Episodes of fever and lymphadenopathy occur over months or years, and hepatosplenomegaly may develop. Eventually infection reaches the central nervous system, causing headache, behavioural changes, confusion and daytime somnolence. As the disease progresses patients may develop tremors, ataxia, convulsions and hemiplegias; eventually coma and death supervene. Histologically there is a lymphocytic meningoencephalitis, with scattered trypanosomes visible in the brain substance.

T. b. rhodesiense sleeping sickness is a much more acute disease. Early systemic features may include myocarditis, hepatitis and serous effusions, and patients can die before the onset of CNS disease. If they survive, cerebral involvement occurs within weeks of infection, and is rapidly progressive.

Diagnosis

Trypanosomes may be seen on Giemsa-stained smears of thick or thin blood films, or of lymph node aspirate. Blood films are usually positive in *T. b. rhodesiense*, but may be negative in *T. b. gambiense*: concentration techniques may increase the yield. Serological tests are useful for screening for infection: the card agglutination test for trypanosomiasis (CATT) is a robust and easy-to-use field assay. Examination of cerebrospinal fluid is essential in patients with evidence of trypanosomal infection. CNS involvement causes lymphocytosis and elevated protein in the CSF, and parasites may be seen in concentrated specimens.

Management

The treatment of sleeping sickness had remained largely unchanged for more than 40 years, but there have been recent improvements in the management of *T. b. gambiense* infection. In both forms, treatment is usually effective if given before the onset of CNS involvement. The drug of choice in early trypanosomiasis is suramin, given intravenously at a dose of 20 mg/kg at 5- to 7-day intervals up to a total dose of 5 g. Severe reactions are relatively common, and a test dose of 100 mg is usually given prior to this regimen. A single dose of suramin should be given to patients with parasitaemia prior to lumbar puncture, to avoid inoculation into the CSF.

The only effective drugs which penetrate the CSF in trypanocidal concentrations were the arsenicals, of which the most widely used is melarsoprol. These are given intravenously; a variety of dosing schedules are in use. Melarsoprol is extremely toxic: 2–10% of patients develop an acute encephalopathy, with a 50–75% mortality, and peripheral neuropathy and hepatorenal toxicity are also common. Prednisolone 1 mg/kg/day decreases the treatment-related mortality by 50% in *T. b. gambiense* infection; it has not been fully assessed in *T. b. rhodesiense* disease. Between 3% and 6% of patients relapse following melarsoprol treatment. In *T. b. gambiense* sleeping sickness, eflornithine (difluoromethylornithine) is effective, clearing parasites from blood and CSF: it has a variable effect on *T. b. rhodesiense* infection. Eflornithine is much less toxic than the arsenical drugs, but was prohibitively expensive. It is now being donated free of charge by the manufacturer, allowing widespread use.

Control

The morbidity and mortality of sleeping sickness could be considerably reduced by early detection and treatment of cases. Control programmes have been effective in some areas, but many have been discontinued because of lack of funding or political upheaval. As in many vector-borne diseases, prevention depends largely on elimination, control or avoidance of the vector. Again this requires considerable coordination and money, and has only been implemented in a few places.

South American trypanosomiasis (Chagas' disease)

Chagas' disease is widely distributed in rural areas of South and Central America. It is caused by *Trypanosoma cruzi*, which is transmitted to humans in the faeces of blood-sucking reduviid bugs (also called cone-nose, or assassin bugs). Faeces infected with *T. cruzi* trypomastigotes are rubbed in through skin abrasions, mucosa or conjunctiva. The bugs, which live in mud or thatch buildings, feed on a variety of vertebrate hosts (e.g. rats, opossums) at night, defecating as they do so.

The parasites spread in the bloodstream, before entering host cells and multiplying. Cell rupture releases them back into the circulation, where they can be taken up by a feeding bug. Further multiplication takes place in the insect gut, completing the trypanosome life cycle. Human infection can also occur via contaminated blood transfusion, or occasionally by transplacental spread.

Clinical features

Acute infection, which usually occurs in children, often passes unnoticed. A firm reddish papule is sometimes seen at the site of entry, associated with regional lymphadenopathy. In the case of conjunctival infection there is swelling of the eyelid, which may close the eye (Romaña's sign). There may be fever, lymphadenopathy, hepatosplenomegaly and rarely meningoencephalitis. Acute Chagas' disease is occasionally fatal in infants, but normally there is full recovery within a few weeks or months.

After a latent period of many years, 10–30% of people go on to develop *chronic Chagas' disease*. The pathogenesis of this is unclear: it is possibly due to an autoimmune response triggered by the initial infection, although recent evidence has thrown doubt on this mechanism. The heart is commonly affected, with conduction abnormalities, arrhythmias, aneurysm formation and cardiac dilatation. Gastrointestinal involvement leads to progressive dilatation of parts of the gastrointestinal tract: this commonly results in megaoesophagus (causing dysphagia and aspiration pneumonia) and megacolon (causing severe constipation).

Diagnosis

Trypanosomes may be seen on a stained blood film during the acute illness. In chronic disease, parasites may be detected by xenodiagnosis: infection-free reduviid bugs are allowed to feed on the patient, and the insect gut subsequently examined for parasites. Serological tests can detect both acute and chronic Chagas' disease.

Management and control

Nifurtimox is the drug of choice for treating Chagas' disease, but it is variably available throughout the world. Benznidazole (5–10 mg/kg/day for 60 days) is also used. The cure rate is about 80% in acute infection. Both are toxic with adverse reactions in up to 50% of patients. The drug treatment of chronic infection is controversial, as most of the tissue damage is thought to be immune-mediated, but there is some recent evidence suggesting that parasite elimination improves the outcome. Antiarrhythmic drugs and pacemakers may be needed in cardiac disease, and surgical treatment is sometimes needed for gastrointestinal complications.

In the long term, prevention of Chagas' disease relies on improved housing and living conditions. In the interim, local vector control programmes may be effective and the countries of the 'Southern Cone' of South America run a successful joint programme to control the disease by spraying houses with insecticide. Impregnated bed nets with pyrethroid and insect repellents should be used.

Leishmaniasis

This group of diseases is caused by protozoa of the genus *Leishmania*, which are transmitted by the bite of the female phlebotomine sandfly (Table 4.39). Leishmaniasis is seen in localized areas of Africa, Asia (particularly India and Bangladesh), Europe, the Middle East and South and Central America. Certain parasite species are specific to each geographical area. The clinical picture is dependent on the species of parasite, and on the host's cell-mediated immune response. Asymptomatic infection, in which the parasite is suppressed or eradicated by a strong immune response, is common in endemic areas, as demonstrated by a high incidence of positive leishmanin skin tests. Symptomatic infection may be confined to the skin (sometimes with spread to the mucous membranes), or widely disseminated throughout the body (visceral leishmaniasis). Relapse of previously

Table 4.39 ***Leishmania*** **species causing visceral and cutaneous disease in man**

	Species complex	Species
Visceral leishmania	*L. donovani*	*L. donovani*
		L. infantum
		L. chagasi
Cutaneous leishmania	*L. tropica*	*L. tropica*
	L. major	*L. major*
	L. aethiopica	*L. aethiopica*
	L. mexicana	*L. mexicana*
		L. amazonensis
		L. garnhami
		L. pifanoi
		L. venezuelensis
	L. braziliensis	*L. braziliensis*
		L. guyanensis
		L. panamanensis
		L. peruviana
Mucocutaneous leishmania		*L. braziliensis*

asymptomatic infection is seen in patients who become immunocompromised, especially those with HIV infection.

In some areas leishmania is primarily zoonotic, whereas in others, man is the main reservoir of infection. In the vertebrate host the parasites are found as oval amastigotes (Leishman–Donovan bodies). These multiply inside the macrophages and cells of the reticuloendothelial system, and are then released into the circulation as the cells rupture. Parasites are taken into the gut of a feeding sandfly (genus *Phlebotomus* in the Old World, genus *Lutzomyia* in the New World), where they develop into the flagellate promastigote form. These migrate to the salivary glands of the insect, where they can be inoculated into a new host.

Visceral leishmaniasis

Clinical features

Visceral leishmaniasis (kala azar) is caused by *L. donovani*, *L. infantum* or *L. chagasi*, and is prevalent in localized areas of Asia, Africa, the Mediterranean littoral and South America. In parts of India, where man is the main host, the disease occurs in epidemics. In most other areas it is endemic, and it is mainly children and visitors to the area who are at risk. The main animal reservoirs in Europe and Asia are dogs and foxes, while in Africa it is carried by various rodents.

The incubation period is usually 1–2 months, but may be several years. The onset of symptoms is insidious, and the patient may feel quite well despite markedly abnormal physical findings. Fever is common, and although usually low-grade, it may be high and intermittent. The liver, and especially the spleen, become enlarged; lymphadenopathy is common in African kala azar. The skin becomes rough and pigmented. If the disease is not treated, profound pancytopenia develops, and the patient becomes wasted and immunosuppressed. Death usually occurs within a year, and is normally due to bacterial infection or uncontrolled bleeding.

Diagnosis

Specific diagnosis is made by demonstrating the parasite in stained smears of aspirates of bone marrow, lymph node, spleen or liver. The organism can also be cultured from these specimens. Specific serological tests are positive in 95% of cases. Pancytopenia, hypoalbuminaemia and hypergammaglobulinaemia are common. The leishmanin skin test is negative, indicating a poor cell-mediated immune response.

Management

The most widely used drugs for visceral leishmaniasis are the pentavalent antimony salts (e.g. sodium stibogluconate, which contains 100 mg of antimony per mL), given intravenously or intramuscularly at a dose of 20 mg of antimony per kg for 21 days. In India, meglumine antimonate is used. Resistance to antimony salts is increasing, and relapses may occur following treatment. The drug of choice where resources permit is intravenous amphotericin B (preferably given in the liposomal form). However, this drug is expensive and not widely available in many areas where the disease is prevalent. An oral drug, miltefosine, has been shown in India to be highly effective, and may replace older therapies. Intercurrent bacterial infections are common and should be treated with antibiotics. Blood transfusion may occasionally be required.

Successful treatment may be followed in a small proportion of patients by a skin eruption called *post-kala azar dermal leishmaniasis* (PKDL). It starts as a macular maculopapular nodular rash which spreads over the body. It is most often seen in the Sudan and India. There have been reports of PKDL responding to miltefosine.

Cutaneous leishmaniasis

Cutaneous leishmaniasis is caused by a number of geographically localized species, which may be zoonotic or anthroponotic. Following a sandfly bite, leishmania amastigotes multiply in dermal macrophages. The local response depends on the species of leishmania, the size of the inoculum, and the host immune response. Single or multiple painless nodules occur on exposed areas within 1 week to 3 months following the bite. These enlarge and ulcerate with a characteristic erythematous raised border. An overlying crust may develop. The lesions heal slowly over months or years, sometimes leaving a disfiguring scar.

L. major and *L. tropica* are found in Russia and Eastern Europe, the Middle East, Central Asia, the Mediterranean littoral and sub-Saharan Africa. The reservoir for *L. major* is desert rodents, while *L. tropica* has a mainly urban distribution with dogs and humans as reservoirs. *L. aethiopica* is found in the highlands of Ethiopia and Kenya, where the animal reservoir is the hyrax. The skin lesions usually heal spontaneously with scarring: this may take a year or more in the case of *L. tropica*. Leishmaniasis recidivans is a rare chronic relapsing form caused by *L. tropica*.

L. mexicana is found predominantly in Mexico, Guatemala, Brazil, Venezuela and Panama: infection usually runs a benign course with spontaneous healing within 6 months. *L. braziliensis* infections (which are seen throughout tropical South America) also usually heal spontaneously, but may take longer.

L. mexicana amazonensis and *L. aethiopica* may occasionally cause diffuse cutaneous leishmaniasis. This is rare, and is characterized by diffuse infiltration of the skin by Leishman–Donovan bodies. Visceral lesions are absent.

Diagnosis and treatment

The diagnosis can often be made clinically in a patient who has been in an endemic area. Giemsa stain on a split-skin

smear will demonstrate leishmania parasites in 80% of cases. Biopsy tissue from the edge of the lesion can be examined histologically, and parasites identified by PCR; culture is less often successful. The leishmanin skin test is positive in over 90% of cases, but does not distinguish between active and resolved infection. Serology is unhelpful.

Small lesions usually require no treatment. Large lesions or those in cosmetically sensitive sites can sometimes be treated locally, by curettage, cryotherapy or topical antiparasitic agents. In other cases, systemic treatment (as for visceral leishmaniasis) is required, although treatment is less successful as antimonials are poorly concentrated in the skin; *L. aethiopica* is not sensitive to antimonials.

Mucocutaneous leishmaniasis

Mucocutaneous leishmaniasis occurs in 3–10% of infections with *L. b. braziliensis*, and is commonest in Bolivia and Peru. The cutaneous sores are followed months or years later by indurated or ulcerating lesions affecting mucosa or cartilage, typically on the lips or nose ('espundia'). The condition can remain static, or there may be progression over months or years affecting the nasopharynx, uvula, palate and upper airways.

Diagnosis and treatment

Biopsies usually show only very scanty organisms, although parasites can be detected by PCR; serological tests are frequently positive.

Amphotericin B is the treatment of choice if available, although systemic antimonial compounds are widely used; miltefosine may also be effective. Relapses are common following treatment. Patients may die because of secondary bacterial infection, or occasionally laryngeal obstruction.

Prevention

Prevention of leishmaniasis relies on control of vectors and/or reservoirs of infection. Insecticide spraying, control of host animals, and treating infected humans may all be helpful. Personal protection against sandfly bites is also necessary, especially in travellers visiting endemic areas. Sandflies are poor fliers and sleeping off the ground helps prevent bites.

Toxoplasmosis

Toxoplasmosis is caused by the intracellular protozoan parasite *Toxoplasma gondii*. The sexual form of the parasite lives in the gut of the definitive host, the cat, where it produces oocysts. After a period of maturing in the environment, these oocysts become the source of infection for secondary hosts which may ingest them. In the secondary hosts (which include man, cattle, sheep, pigs, rodents and birds) there is disseminated infection. Following a successful immune response the infection is controlled, but dormant parasites remain encysted in host tissue for many years. The life cycle is completed when carnivorous felines eat infected animal tissue. Humans are infected either from contaminated cat faeces, or by eating undercooked infected meat; transplacental infection may also occur.

Clinical features

Toxoplasmosis is common: seroprevalence in adults in the UK is about 25%, rising to 90% in some parts of Europe. Most infections are asymptomatic or trivial. Symptomatic patients usually present with lymphadenopathy, mainly in the head and neck. There may be fever, myalgia and general malaise; occasionally there are more severe manifestations including hepatitis, pneumonia, myocarditis and choroidoretinitis. Lymphadenopathy and fatigue can sometimes persist for months after the initial infection.

Congenital toxoplasmosis may also be asymptomatic, but can produce serious disease. Clinical manifestations include microcephaly, hydrocephalus, encephalitis, convulsions and mental retardation. Choroidoretinitis is common; occasionally this may be the only feature.

Immunocompromised patients, especially those with HIV infection, are at risk of serious infections with *T. gondii*. In acquired immunodeficiency states this is usually due to reactivation of latent disease (see p. 201).

Diagnosis

Diagnosis is usually made serologically. IgG antibodies detectable by the Sabin–Feldman dye test remain positive for years; acute infection can be confirmed by demonstrating a rising titre of specific IgM.

Management

Acquired toxoplasmosis in an immunocompetent host rarely requires treatment. In those with severe disease (especially eye involvement) sulfadiazine 2–4 g daily and pyrimethamine 25 mg daily are given for 4 weeks, along with folinic acid. The management of pregnant women with toxoplasmosis aims to decrease the risk of fetal complications. However there is little good evidence that giving spiramycin either alone or in combination with sulfadiazine (which is the recommended treatment) has any significant effect on the frequency or severity of fetal damage. Infected infants should be treated from birth. The treatment of toxoplasmosis in HIV-positive patients is covered on page 201.

Babesiosis

Babesiosis is a tick-borne parasitic disease, diagnosed most commonly in North America and Europe. It is a zoonosis of rodents and cattle, and is occasionally transmitted to humans: infection is more common and more severe in those who are immunocompromised following splenectomy. The causative organisms are the plasmodium-like *Babesia microti* (rodents) and *B. divergens* (cattle).

The incubation period averages 10 days. In patients with normal splenic function, the illness is usually mild. In splenectomized individuals, systemic symptoms are more pronounced and haemolysis is associated with haemoglobinuria, jaundice and renal failure. Examination of a peripheral blood smear may reveal the characteristic plasmodium-like organisms.

The standard treatment was a combination of quinine 650 mg and clindamycin 600 mg orally three times daily for 7 days but a regimen of atovaquone and azithromycin is as effective, with fewer adverse reactions.

GASTROINTESTINAL PROTOZOA

The major gastrointestinal parasites of man are shown in Table 4.40.

Table 4.40 Pathogenic human intestinal protozoa

Amoebae
Entamoeba histolytica

Flagellates
Giardia intestinalis

Ciliates
Balantidium coli

Coccidia
Cryptosporidium parvum
Isospora belli
Sarcocystis spp.
Cyclospora cayetanensis

Microspora
Enterocytozoon bieneusi
Encephalitozoon spp.

Amoebiasis

Amoebiasis is caused by *Entamoeba histolytica*. The organism formerly known as *E. histolytica* is known to consist of two distinct species: *E. histolytica*, which is pathogenic, and *E. dispar*, which is non-pathogenic. Cysts of the two species are identical, but can be distinguished by molecular techniques after culture of the trophozoite. *E. histolytica* can be distinguished from all amoebae except *E. dispar*, and from other intestinal protozoa, by microscopic appearance. Amoebiasis occurs world-wide, although much higher incidence rates are found in the tropics and subtropics.

The organism exists both as a motile trophozoite and as a cyst that can survive outside the body. Cysts are transmitted by ingestion of contaminated food or water, or spread directly by person-to-person contact. Trophozoites emerge from the cysts in the small intestine and then pass on to the colon, where they multiply.

Clinical features

It is believed that many individuals can carry the pathogen without obvious evidence of clinical disease (asymptomatic cyst passers). However, this may be due in some cases to the misidentification of non-pathogenic *E. dispar* as *E. histolytica*, and it is not clear how often true *E. histolytica* infection is symptomless. In affected people *E. histolytica* trophozoites invade the colonic epithelium, probably with the aid of their own cytotoxins and proteolytic enzymes. The parasites continue to multiply and finally frank ulceration of the mucosa occurs. If penetration continues, trophozoites may enter the portal vein, via which they reach the liver and cause intrahepatic abscesses. This invasive form of the disease is serious and may even be fatal.

The incubation period of intestinal amoebiasis is highly variable and may be as short as a few days or as long as several months. The usual course is chronic, with mild intermittent diarrhoea and abdominal discomfort. This may progress to bloody diarrhoea with mucus, and is sometimes accompanied by systemic symptoms such as headache, nausea and anorexia. Less commonly, infection may present as acute amoebic dysentery, resembling bacillary dysentery or acute ulcerative colitis.

Complications are unusual, but include toxic dilatation of the colon, chronic infection with stricture formation, severe haemorrhage, amoeboma and amoebic liver abscess. Amoebomas, which develop most commonly in the caecum or rectosigmoid region, are sometimes mistaken for carcinoma. They may bleed, cause obstruction or intussuscept. Amoebic liver abscesses often develop in the absence of a recent episode of colitis. Tender hepatomegaly, a high swinging fever and profound malaise are characteristic, although early in the course of the disease both symptoms and signs may be minimal. The clinical features are described in more detail on page 361.

Diagnosis

Microscopic examination of fresh stool or colonic exudate obtained at sigmoidoscopy is the simplest way of diagnosing colonic amoebic infection. To confirm the diagnosis motile trophozoites containing red blood cells must be identified: the presence of amoebic cysts alone does not imply disease. Sigmoidoscopy and barium enema examination may show colonic ulceration but are rarely diagnostic.

The amoebic fluorescent antibody test is positive in at least 90% of patients with liver abscess and in 60–70% with active colitis. Seropositivity is low in asymptomatic cyst passers.

Management

Metronidazole 800 mg three times daily for 5 days is given in amoebic colitis; a lower dose (400 mg three times daily for 5 days) is usually adequate in liver abscess. Tinidazole is also effective: dehydroemetine and chloroquine are rarely used. After treatment of the invasive disease, the bowel should be cleared of parasites with a luminal amoebicide such as diloxanide furoate.

Prevention

Amoebiasis is difficult to eradicate because of the substantial human reservoir of infection. The only progress will be through improved standards of hygiene, sanitation and better access to clean water. Cysts are destroyed by boiling, but chlorine and iodine sterilizing tablets are not always effective.

Giardiasis

Giardia intestinalis is a flagellate (Fig. 4.36) that is found world-wide. It causes small intestinal disease, with diarrhoea and malabsorption. Prevalence is high in many developing countries, and it is the most common parasitic infection in travellers returning to the UK. In certain parts of Europe, and in some rural areas of North America, large water-borne epidemics have been reported. Person-to-person spread may occur in day nurseries and residential institutions. Like *E. histolytica*, the organism exists both as a trophozoite and a cyst, the latter being the form in which it is transmitted.

The organism sometimes colonizes the small intestine and may remain there without causing detriment to the host. In other cases, severe malabsorption may occur which is thought to be related to morphological damage to the small intestine. The changes in villous architecture are usually mild partial villous atrophy; subtotal villous atrophy is rare. The mechanism by which the parasite causes alteration in mucosal architecture and produces diarrhoea and intestinal

Fig. 4.36 ***Giardia intestinalis*** **on small intestinal mucosa.** Courtesy of Dr A Phillips, Department of Electron Microscopy, Royal Free Hospital, London.

malabsorption is unknown: there is evidence that the morphological damage may be immune-mediated. Bacterial overgrowth has also been found in association with giardiasis and may contribute to fat malabsorption.

Clinical features

Many individuals excreting giardia cysts have no symptoms. Others become ill within 1–3 weeks after ingesting cysts: symptoms include diarrhoea, often watery in the early stage of the illness, nausea, anorexia, and abdominal discomfort and bloating. In most people affected these resolve after a few days, but in some the symptoms persist. Stools may then become paler, with the characteristic features of steatorrhoea. If the illness is prolonged, weight loss ensues, which can be marked. Chronic giardiasis frequently seen in developing countries can result in growth retardation in children.

Diagnosis

Both cysts and trophozoites can be found in the stool, but negative stool examination does not exclude the diagnosis since the parasite may be excreted at irregular intervals. The parasite can also be seen in duodenal aspirates (obtained either at endoscopy or with an Enterotest capsule), and in histological sections of jejunal mucosa.

Management

Metronidazole 2 g as a single dose on 3 successive days will cure the majority of infections, although sometimes a second or third course is necessary. Alternative drugs include tinidazole, mepacrine and albendazole. Preventative measures are similar to those outlined above for *E. histolytica*.

Cryptosporidiosis

This organism is found world-wide, cattle being the major natural reservoir. It has also been demonstrated in supplies of drinking water in the UK. The parasite is able to reproduce both sexually and asexually; it is transmitted by oocysts excreted in the faeces.

In healthy individuals cryptosporidiosis is a self-limiting illness. Acute watery diarrhoea is associated with fever and general malaise lasting for 7–10 days. In immunocompromised patients, especially those with HIV, diarrhoea is severe and intractable (see p. 201).

Diagnosis is usually made by faecal microscopy, although the parasite can also be detected in intestinal biopsies. There is no reliable treatment, although nitazoxanide may be of benefit.

Balantidiasis

Balantidium coli is the only ciliate that produces clinically significant infection in humans. It is found throughout the tropics, particularly in Central and South America, Iran, Papua New Guinea and the Philippines. It is usually carried by pigs, and infection is most common in those communities that live in close association with swine. Its life cycle is identical to that of *E. histolytica*. *B. coli* causes diarrhoea, and sometimes a dysenteric illness with invasion of the distal ileal and colonic mucosa. Trophozoites rather than cysts are found in the stool. Treatment is with tetracycline or metronidazole.

Blastocystis hominis infection

B. hominis is a strictly anaerobic protozoan pathogen that inhabits the colon. For decades its pathogenicity for humans was questioned, but there is increasing evidence that it may cause diarrhoea. It is sensitive to metronidazole.

Cyclospora cayetanensis infection

Cyclospora cayetanensis, a coccidian protozoal parasite, was originally recognized as a cause of diarrhoea in travellers to Nepal. It has been detected in stool specimens from immunocompetent and immunodeficient people world-wide. Infection is usually self-limiting, but can be treated with co-trimoxazole.

Microsporidiosis

Protozoa of the phylum *Microsporea* can cause diarrhoea in patients with HIV/AIDS (see p. 201).

FURTHER READING

Barrett MP. The rise and fall of sleeping sickness. *Lancet* 2006; **367**: 1377–1379.

Bhattacharya SK, Sinha PK, Sundar S et al. Phase 4 trial of miltefosine for the treatment of Indian visceral leishmaniasis. *Journal of Infectious Diseases* 2007; **196**: 591–598.

Duffy P, Mutabingwa T. Artemisinin combination therapies. *Lancet* 2006; **367**: 2037–2038.

Farthing MJ. Treatment options for the eradication of intestinal protozoa. *National Clinical Practice in Gastroenterology and Hepatology* 2006; **3**: 436–445.

Greenwood BM, Bojang K, Whitty C et al. Malaria. *Lancet* 2005; **365**: 1487–1498.

Maguire JH. Chagas' disease – can we stop the deaths? *New England Journal of Medicine* 2006; **355**: 760–761.

HELMINTHIC INFECTIONS

Worm infections are very common in developing countries, causing much disease in both humans and domestic animals. They are frequently imported into industrialized countries. The most common human helminth infections are listed in Table 4.41. Neglected tropical diseases include a group of 13 parasitic and bacterial infections causing major disability amongst the poorest people in the world. They include three soil-transmitted helminth infections – ascariasis, hookworm and trichuriasis.

Helminths are the largest internal human parasite. They reproduce sexually, generating millions of eggs or larvae. *Nematodes* and *trematodes* have a mouth and intestinal tract, while *cestodes* absorb nutrients directly through the outer tegument. All worms are motile, although once the adults are established in their definitive site they rarely migrate further. Adult helminths may be very long-lived: up to 30 years in the case of the schistosomes.

Many helminths have developed complex life cycles, involving more than one host. Both primary and intermediate hosts are often highly specific to a particular species of worm. In some cases of human infection man is the primary host, while in others humans are a non-specific intermediary or are coincidentally infected. Multiple infections with different helminths are common in endemic areas. Mass treatment programmes, in which one or more anthelminthic drugs are given on a regular (usually annual) basis, are used to keep the total worm load down (Table 4.42)

NEMATODES

Human infections can be divided into:

- Tissue-dwelling worms including the filarial worms, and the Guinea worm *Dracunculus medinensis*.

Table 4.42 Drugs used in mass treatment programmes for helminth infections (alone or in combination)

Drug	Infection
Diethylcarbamazine (DEC)	Loiaisis Filariasis
Ivermectin	Loiaisis Filariasis Onchocerciasis Strongyloidiasis
Albendazole	Filariasis (with DEC) Intestinal helminths
Praziquantel	Schistosomiasis

Table 4.41 Helminths commonly infecting man

	Helminth	Common name/disease caused
Nematodes (roundworms)		
Tissue-dwelling worms	*Wuchereria bancrofti*	Filariasis
	Brugia malayi/timori	Filariasis
	Loa loa	Loiasis
	Onchocerca volvulus	River blindness
	Dracunculus medinensis	Dracunculiasis
	Mansonella perstans	Mansonellosis
Intestinal human nematodes	*Enterobius vermicularis*	Threadworm
	Ascaris lumbricoides	Roundworm
	Trichuris trichiura	Whipworm
	Necator americanus	Hookworm
	Ancylostoma duodenale	Hookworm
	Strongyloides stercoralis	Strongyloidosis
Zoonotic nematodes	*Toxocara canis*	Toxocariasis
	Trichinella spiralis	Trichinellosis
Trematodes (flukes)		
Blood flukes	*Schistosoma* species	Schistosomiasis
Lung flukes	*Paragonimus* species	Paragonimiasis
Intestinal/hepatic flukes	*Fasciolopsis buski*	
	Fasciola hepatica	
	Clonorchis sinensis	
	Opisthorchis felineus	
Cestodes (tapeworms)		
Intestinal adult worms	*Taenia saginata*	Beef tapeworm
	Taenia solium	Pork tapeworm
	Diphyllobothrium latum	Fish tapeworm
	Hymenolepis nana	Dwarf tapeworm
Larval tissue cysts	*Taenia solium*	Cysticercosis
	Echinococcus granulosus	Hydatid disease
	Echinococcus multilocularis	Hydatid disease
	Spirometra mansoni	Sparganosis

- Human intestinal worms, including the human hookworms, the common roundworm (*Ascaris lumbricoides*) and *Strongyloides stercoralis*, which are the most common helminthic parasites of man. The adult worms live in the human gut, and do not usually invade tissues, but many species have a complex life cycle involving a migratory larval stage.
- Zoonotic nematodes which accidentally infect man and are not able to complete their normal life cycle. They often become 'trapped' in the tissues, causing a potentially severe local inflammatory response.

Tissue-dwelling worms

Filariasis

Several nematodes belonging to the superfamily Filarioidea can infect humans. The adult worms are long and threadlike, ranging from 2 cm to 50 cm in length; females are generally much larger than males. Larval stages are inoculated into humans by various species of biting flies (each specific to a particular parasite). The adult worms which develop from these larvae mate, producing millions of offspring (microfilariae), which migrate in the blood or skin. These are ingested by feeding flies, in which the remainder of the life cycle takes place. Disease, which may be caused by either the adult worms or by microfilariae, is caused by host immune response to the parasite and is characterized by massive eosinophilia. Adult worms are long-lived (10–15 years), and reinfection is common, so that disease tends to be chronic and progressive.

Lymphatic filariasis

Lymphatic filariasis, which may be caused by different species of filarial worm, has a scattered distribution in the tropics and subtropics (Table 4.43). Nearly 1 billion people in developing countries are at risk. *Wuchereria bancrofti* is transmitted to man by a number of mosquito species, mainly *Culex fatigans*. Adult female worms (which are 5–10 cm long) live in the lymphatics, releasing large numbers of microfilariae into the blood. Generally this occurs at night, coinciding with the nocturnal feeding pattern of *C. fatigans*. Non-periodic forms of *W. bancrofti*, transmitted by day-biting species of mosquito, are found in the South Pacific. *Brugia malayi* (and the closely related *B. timori*) are very similar to *W. bancrofti*, exhibiting the same nocturnal periodicity. The usual vectors are mosquitoes of the genus *Mansonia*, although other mosquitoes have been implicated.

Clinical features

Following the bite of an infected mosquito, the larvae enter the lymphatics and are carried to regional lymph nodes. Here they grow and mature for 6–18 months.

Adult worms produce allergic lymphangitis. The clinical picture depends on the individual immune response, which in turn may depend on factors such as age at first exposure. In endemic areas many people have asymptomatic infection. Sometimes early infection is marked by bouts of fever accompanied by pain, tenderness and erythema along the course of affected lymphatics. Involvement of the spermatic cord and epididymis are common in Bancroftian filariasis. These acute attacks subside spontaneously in a few days, but usually recur. Recurrent episodes cause intermittent lymphatic obstruction, which in time can become fibrotic and irreversible. Obstructed lymphatics may rupture, causing cellulitis and further fibrosis; there may also be chylous pleural effusions and ascites. Over time there is progressive enlargement, coarsening and fissuring of the skin, leading to the classical appearances of elephantiasis. The limbs or scrotum may become hugely swollen. Eventually the adult worms will die, but the lymphatic obstruction remains and tissue damage continues. Elephantiasis takes many years to develop, and is only seen in association with recurrent infection in endemic areas.

Occasionally the predominant features of filarial infection are pulmonary. Microfilariae become trapped in the pulmonary capillaries, generating intense local allergic response. The resulting pneumonitis causes cough, fever, weight loss and shifting radiological changes, associated with a high peripheral eosinophil count. This is known as *tropical pulmonary eosinophilia* (see p. 874).

Diagnosis

The diagnosis of filariasis is usually made on clinical grounds, supported by a high eosinophil count. Serological tests are sensitive, especially in the earlier stages, but can cross-react with other nematodes: they become negative 1–2 years after effective treatment. Microfilariae start to appear in the peripheral blood about a year after infection, and may be detected on a stained nocturnal blood film. Parasitological diagnosis is difficult in established elephantiasis. The clinical picture is usually diagnostic, but similar lymphatic damage may occasionally be caused by silicates absorbed through the feet from volcanic soil (podoconiosis).

Table 4.43 Diseases caused by the filarial worms

Organism	Adult worm	Microfilariae	Major vector	Clinical signs	Distribution
Wuchereria bancrofti	Lymphatics	Blood	*Culex* species	Fever Lymphangitis Elephantiasis	Tropics
Brugia timori/malayi	Lymphatics	Blood	*Mansonia* species	Fever Lymphangitis Elephantiasis	East and South East Asia, South India, Sri Lanka
Loa loa	Subcutaneous	Blood	*Chrysops* species	'Calabar' swellings Urticaria	West and Central Africa
Onchocerca	Subcutaneous	Skin, eye	*Simulium* species	Subcutaneous nodules Eye disease	Africa, South America
Mansonella perstans	Retroperitoneal	Blood	*Culicoides* species	Allergic eosinophilia	Sub-Saharan Africa, South America

Treatment

Diethylcarbamazine (DEC) kills both adult worms and microfilariae. Serious allergic responses may occur as the parasites are killed, and particular care is needed when using DEC in areas endemic for loiasis. Old multi-dose regimens are now being replaced by single-dose treatment (6 mg/kg, usually repeated annually) often in combination with albendazole. Ivermectin is also used. Associated bacterial infections should be treated promptly, and reconstructive surgery may be needed to remove excess tissue. Mass chemotherapy can decrease the prevalence and severity of infection in endemic areas and 80 million people have already been treated under the WHO eradication programme. Early diagnosis and treatment prevent the development of elephantiasis. These approaches must be combined with vector control to achieve permanent results, while individual protection depends on avoidance of mosquito bites.

Loiasis

Loiasis is found in the humid forests of West and Central Africa. The causative parasite, *Loa loa*, is a small (3–7 cm) filarial worm which is found in the subcutaneous tissues. The microfilariae circulate in the blood during the day, but cause no direct symptoms. The vectors are day-biting flies of the genus *Chrysops*.

Adult worms migrate around the body in subcutaneous tissue planes. Worms may be present for years, frequently without causing symptoms. From time to time localized, tender, hot, soft tissue swellings occur due to hypersensitivity (Calabar swellings) often near to a joint. These are produced in response to the passage of a worm and usually subside over a few days or weeks. There may also be more generalized urticaria and pruritus. Occasionally a worm may be seen crossing the eye under the conjunctiva; they may also enter retro-orbital tissue, causing severe pain. Short-term residents of endemic areas often have more severe manifestations of the disease.

Microfilariae may be seen on stained blood films, although these are often negative. Serological tests are relatively insensitive, and cross-react with other microfilariae. There is usually massive eosinophilia. DEC may cause severe allergic reactions (including fever, urticaria, myalgia, and occasionally encephalitis) associated with parasite killing and is being replaced by newer agents. Ivermectin in single doses of 200–400 μg/kg is effective: it may occasionally cause severe reactions. Albendazole, which causes a more gradual reduction in microfilarial load, may be preferable in heavily-infected patients. Mass treatment with either DEC or ivermectin can decrease the transmission of infection, but the mainstay of prevention is vector avoidance and control.

Onchocerciasis

Onchocerciasis (river blindness) affects 18 million people world-wide, of whom 250 000 are blind and 500 000 visually impaired; most of these are in West and Central Africa, with small foci in the Yemen, and Central and South America. It is the result of infection with *Onchocerca volvulus*. Infection is transmitted by day-biting flies of the genus *Simulium*.

Pathogenesis

Infection occurs when larvae are inoculated into humans by the bite of an infected fly. The worms mature in 2–4 months, and can live for more than 15 years. Adult worms, which can reach lengths of 50 cm (although less than 0.5 mm in diameter), live in the subcutaneous tissues. They may form fibrotic nodules, especially over bony prominences and sites of trauma. Huge numbers of microfilariae are distributed in the skin, and may invade the eyes. Live microfilariae cause relatively little harm, but dead parasites may cause severe allergic reactions, with hyaline necrosis and loss of tissue collagen and elastin. In the eye a similar process causes conjunctivitis, sclerosing keratitis, uveitis and secondary glaucoma. Choroidoretinitis is also occasionally seen.

Clinical features

Symptoms usually start about a year after infection. Initially there is generalized pruritus, with urticaria and fleeting oedema. Subcutaneous nodules (which can be detected by ultrasound) start to appear, and in dark-skinned individuals, hypo- and hyperpigmentation from excoriation and inflammatory changes. Over time more chronic inflammatory changes appear, with roughened, inelastic skin. Superficial lymph nodes become enlarged, and in the groin may hang down in loose folds of skin ('hanging groin'). Eye disease, which is associated with chronic heavy infection, usually first manifests as itching and conjunctival irritation. This gradually progresses to more extensive eye disease and eventually to blindness.

Diagnosis

In endemic areas the diagnosis can often be made clinically, especially if supported by finding eosinophilia on a blood film. In order to identify parasites, skin snips taken from the iliac crest or shoulder are placed in saline under a cover slip. After 4 hours, microscopy will show microfilariae wriggling free on the slide. If this is negative, DEC can be applied topically under an occlusive dressing: this will provoke an allergic rash in the majority of infected people (modified Mazzotti reaction) but this is not routinely performed as it is unpleasant. Slit-lamp examination of the eyes may reveal the microfilariae. Serological tests are frequently positive in endemic areas, but this does not imply current infection; they may be negative in expatriates with a low worm load.

Management and prevention

Ivermectin, in a single dose of 150 μg/kg, kills microfilariae and prevents their return for 6–12 months. There is little effect on adult worms, so annual (or more frequent) retreatment is needed. In patients co-infected with *Loa loa*, ivermectin may occasionally induce severe allergic reactions, including a toxic encephalopathy.

Since 1974 the WHO Onchocerciasis Control Programme has had a considerable impact on onchocerciasis in West Africa. A combination of vector control measures and, more recently, mass treatment with ivermectin, has led to a decrease in both infection rates and progression to serious disease. Humans are the only host but measures are required over a long period because of the longevity of the worm (10–15 years).

Mansonellosis

Mansonella perstans is a filarial worm transmitted by biting midges of the genus *Culicoides*. Small numbers of microfilariae are found in the blood, and although they do not cause serious disease there may be minor allergic reactions and an eosinophilia.

Dracunculiasis

Infection with the Guinea worm, *Dracunculus medinensis*, occurs when water fleas (copepods) containing the parasite larvae are swallowed in contaminated drinking water. Ingested larvae mature and the female worm, which can reach over a metre in length, migrates through connective and subcutaneous tissue for 9–18 months before surfacing on the skin. The uterus of the worm ruptures, releasing larvae which are ingested by the small crustacean water fleas, and the cycle is completed.

The diagnosis is clinical.

The traditional treatment, extracting the worm over several days by winding it round a stick, is probably still the most effective. The worm should not be damaged. Antibiotics may be needed to control secondary infection.

Water fleas (and thus infective larvae) can be removed from drinking water by chemical treatment or by simple filtration. Large-scale eradication programmes have been in place for several years and the incidence of infection has fallen from 900 000 cases in 1989 to 10 600 in 2005. The disease is now confined to small areas of Africa, mainly in Sudan and Ghana. Man is the only host of *D. medinensis*, and it should therefore be possible to completely eradicate this parasite.

Human intestinal nematodes

Adult intestinal nematodes (also sometimes referred to as soil-transmitted helminths, or geohelminths) live in the human gut. There are two main types of life cycle, both including a soil-based stage. In some cases infection is spread by ingestion of eggs (which often require a period of maturation in the environment), while in others the eggs hatch in the soil and larvae penetrate directly through the skin of a new host. *Ascaris lumbricoides* larvae invade the duodenum and enter the venous system, via which they reach the lungs. They are eventually expectorated and swallowed, entering the intestine where they complete their maturation. *Strongyloides* is also unusual, in that it is the only nematode that is able to complete its life cycle in humans. Larvae may hatch before leaving the colon, and so are able to reinfect the host by penetrating the intestinal wall and entering the venous system.

www.studentconsult.com
More online

Ascariasis (roundworm infection)

Ascaris lumbricoides is a pale yellow worm, 20–35 cm in length (Fig. 4.37). It is found world-wide but is particularly common in poor rural communities, where there is heavy faecal contamination of the immediate environment. Larvae migrate through the tissues to the lungs before being expectorated and swallowed; adult worms are found in the small intestine. Ova are deposited in faeces, and require a 2- to 4-month maturation in the soil before they are infective.

Infection is usually asymptomatic, although heavy infections are associated with nausea, vomiting, abdominal discomfort and anorexia. Worms can sometimes obstruct the small intestine, the most common site being at the ileocaecal valve. They may also occasionally invade the appendix, causing acute appendicitis, or the bile duct, resulting in biliary obstruction and suppurative cholangitis. Larvae in the lung may produce pulmonary eosinophilia. Heavy infection in children, especially those who are already malnourished, may have significant effects on nutrition and development. Serious morbidity and mortality are rare in ascariasis, but the huge number of people infected means that on a global basis roundworm infection causes a significant burden of disease, especially in children.

Ascaris eggs can be identified in the stool, and occasionally adult worms emerge from the mouth or the anus. They may also be seen on barium enema studies. Appropriate drug treatments are shown in Box 4.15. Very rarely, surgical or endoscopic intervention may be required for intestinal or biliary obstruction.

Fig. 4.37 ***Ascaris lumbricoides*, approximately 20 cm long.**

Box 4.15 Drugs used for treating human intestinal nematodes (single dose unless otherwise stated)

		Ascaris	Hookworm	*Enterobius*	*Trichuris*	*Strongyloides*
Piperazine	75 mg/kg	++	+	++	–	–
Pyrantal pamoate	10 mg/kg	++	+	++	–	–
Oxantel pamoate	10 mg/kg	++	+	n/a	++	–
Albendazole	400 mg*	++	++	++	+	+
Mebendazole	500 mg†	++	++	++	+	+
Tiabendazole	25 mg/kg*	n/a	n/a	n/a	n/a	++
Levamisole	5 mg/kg	++	+	n/a	n/a	–
Ivermectin	200 µg/kg‡	n/a	n/a	n/a	n/a	++

++, Highly effective; +, moderately effective; –, ineffective; n/a, drug not used for this indication/no data available.
*Twice daily for 3 days in strongyloidiasis.
†WHO recommended dose for developing countries; in UK commonly given as 100 mg single dose for threadworm, or 100 mg twice daily for 3 days for whipworm.
‡Once daily for 2 days.

www.studentconsult.com
Ask the authors

Threadworm (Enterobius vermicularis)

E. vermicularis is a small (2–12 mm) worm, which is common throughout the world. Larval development takes place mainly in the small intestine, and adult worms are normally found in the colon. The gravid female deposits eggs around the anus causing intense itching, especially at night. Unlike *A. lumbricoides*, the eggs do not require a maturation period in soil, and infection is often directly transmitted from anus to mouth via the hands. Eggs may also be deposited on clothing and bed linen, and are subsequently either ingested or inhaled. Apart from discomfort and local excoriation, infection is usually harmless.

Ova can be collected either using a moistened perianal swab, or by applying adhesive cellophane tape to the perianal skin. They can then be identified by microscopy.

The most commonly used drugs are mebendazole and piperazine (Box 4.15). However, isolated treatment of an affected person is often ineffective. Other family members (especially small children) may also need to be treated, and the whole family should be given advice about personal hygiene. Two courses of treatment 2 weeks apart may break the cycle of autoinfection.

Whipworm (Trichuris trichiura)

Infections with whipworm are common world-wide, especially in poor communities with inadequate sanitation. Adult worms, which are 3–5 cm long, inhabit the terminal ileum and caecum, although in heavy infection they are found throughout the large bowel. The head of the worm is embedded in the intestinal mucosa. Ova are deposited in the faeces, and require a maturation period of 3–4 weeks in the soil before becoming infective.

Infection is usually asymptomatic, but mucosal damage can occasionally be so severe that there is colonic ulceration, dysentery or rectal prolapse.

Diagnosis is made by finding ova on stool microscopy, or occasionally by seeing adult worms on sigmoidoscopy. Drug treatment is shown in Box 4.15.

Hookworm infection

Hookworm infections, caused by the human hookworms *Ancylostoma duodenale* and *Necator americanus*, are found world-wide. They are relatively rare in developed countries, but very common in areas with poor sanitation and hygiene: overall about 25% of the world's population is affected. Hookworm infection is a major contributing factor to anaemia in the tropics. *A. duodenale* is found mainly in East Asia, North Africa and the Mediterranean, while *N. americanus* is the predominant species in South and Central America, South East Asia and sub-Saharan Africa.

Adult worms (which are about 1 cm long) live in the duodenum and upper jejunum, where they are often found in large numbers. They attach firmly to the mucosa using the buccal plate, feeding on blood. Eggs passed in the faeces develop in warm moist soil, producing infective filariform larvae. These penetrate directly through the skin of a new host, and are carried in the bloodstream to the lungs. Having crossed into the alveoli, the parasites are expectorated and then swallowed, thus arriving at their definitive home.

Clinical features

Local irritation as the larvae penetrate the skin ('ground itch') may be followed by transient pulmonary signs and symptoms, often accompanied by eosinophilia. Light infections, especially in a well-nourished person, are often asymptomatic. Heavier worm loads may be associated with epigastric pain and nausea, resembling peptic ulcer disease. Chronic heavy infection, particularly on a background of malnourishment, may cause iron deficiency anaemia, and hypoproteinaemia. Heavy infection in children is associated with delays in physical and mental development.

Diagnosis and treatment

The diagnosis is made by finding eggs on faecal microscopy. In infections heavy enough to cause anaemia these will be present in large numbers. The aim of treatment in endemic areas is reduction of worm burden rather than complete eradication: albendazole or mebendazole, which can both be given as a single dose, are the best drugs (Box 4.15). The WHO is promoting mass treatment programmes for schoolchildren in many parts of the world, together with treatment for schistosomiasis where appropriate.

Strongyloidiasis

Strongyloides stercoralis is a small (2-mm long) worm which lives in the small intestine. It is found in many parts of the tropics and subtropics, and is especially common in Asia. Eggs hatch in the bowel, and larvae are found in the stool. Usually these are non-infective rhabditiform larvae, which require a further period of maturation in the soil before they can infect a new host, but sometimes this maturation can occur in the large bowel. Infective filariform larvae can therefore penetrate directly through the perianal skin, reinfecting the host. In this way autoinfection may continue for years or even decades. Some war veterans who were imprisoned in the Far East during the Second World War have been found to have active strongyloidiasis over 50 years later. After skin penetration the life cycle is similar to that of the hookworm, except that the adult worms may burrow into the intestinal mucosa, causing a local inflammatory response.

Clinical features

Skin penetration by *S. stercoralis* causes a similar local dermatitis to hookworm. In autoinfection this manifests as a migratory linear weal around the buttocks and lower abdomen (cutaneous larva currens). In heavy infections damage to the small intestinal mucosa can cause malabsorption, diarrhoea and even perforation. There is usually a persistent eosinophilia. In patients who are immunosuppressed (e.g. by corticosteroid therapy or intercurrent illness) filariform larvae may penetrate directly through the bowel wall in huge numbers, causing an overwhelming and usually fatal generalized infection (the strongyloidiasis hyperinfestation syndrome). This condition is often complicated by septicaemia due to bowel organisms, and therefore a broad-spectrum antibiotic should always be given.

Diagnosis and treatment

Motile larvae may be seen on stool microscopy, especially after a period of incubation. Serological tests are also useful. The best drug for treating strongyloidiasis is ivermectin (200 μg/kg daily for 2 days); albendazole (15 mg/kg 12-hourly for 3 days) is also effective.

Zoonotic nematodes

A number of nematodes which are principally parasites of animals may also affect man. The most common are described below.

Trichinosis

The normal hosts of *Trichinella spiralis*, the cause of trichinosis, include pigs, bears and warthogs. Man is infected by eating undercooked meat from these animals. Ingested larvae mature in the small intestine, where adults release new larvae which penetrate the bowel wall and migrate through the tissues. Eventually these larvae encyst in striated muscle.

Light infections are usually asymptomatic. Heavier loads of worms produce gastrointestinal symptoms as the adults establish themselves in the small intestine, followed by systemic symptoms as the larvae invade. The latter include fever, oedema and myalgia. Massive infection may occasionally be fatal, but usually the symptoms subside once the larvae encyst.

The diagnosis can usually be made from the clinical picture, associated eosinophilia, and serological tests. If necessary it can be confirmed by muscle biopsy a few weeks after infection. Albendazole (20 mg/kg for 7 days) given early in the course of the illness will kill the adult worms and decrease the load of larvae reaching the tissues. Analgesia and steroids may be needed for symptomatic relief.

Toxocariasis (visceral larva migrans)

Eggs of the dog roundworm, *Toxocara canis*, are occasionally ingested by humans, especially children. The eggs hatch and the larvae penetrate the small intestinal wall and enter the mesenteric circulation, but are then unable to complete their life cycle in a 'foreign' host. Many are held up in the capillaries of the liver, where they generate a granulomatous response, but some may migrate into other tissues including lungs, striated muscle, heart, brain and eye. In most cases infection is asymptomatic, and the larvae die without causing serious problems. In heavy infections there may be generalized symptoms (fever and urticaria) and eosinophilia, as well as focal signs related to the migration of the parasites. Pulmonary involvement may cause bronchospasm and chest X-ray changes. Ocular infection may produce a granulomatous swelling mimicking a retinoblastoma, while cardiac or neurological involvement may occasionally be fatal. Rarely, larvae survive in the tissues for many years, causing symptoms long after infection.

Isolation of the larvae is difficult, and the diagnosis is usually made serologically. Albendazole 400 mg daily for a week is the most effective treatment.

Cutaneous larva migrans (CLM)

CLM is caused by the larvae of the non-human hookworms *Ancylostoma braziliense* and *A. caninum*. Like human hookworms, these hatch in warm moist soil, and then penetrate the skin. In man they are unable to complete a normal life cycle, and instead migrate under the skin for days or weeks until they eventually die. The wandering of the larva is accompanied by a clearly defined, serpiginous, itchy rash, which progresses at the rate of about 1 cm per day. There are usually no systemic symptoms. The diagnosis is purely clinical. Single larvae may be treated with a 15% solution of topical tiabendazole; multiple lesions require systemic therapy with a single dose of albendazole 400 mg or ivermectin 150–200 μg/kg.

FURTHER READING

Bethony J, Brooker S, Albonico M et al. Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. *Lancet* 2006; **367**: 1521–1536.

Hoerauf A, Buttner D, Adjei O et al. Onchocerciasis. *British Medical Journal* 2003; **326**: 207–210.

Hotez PJ, Brooker S, Bethony JM et al. Hookworm infections. *New England Journal of Medicine* 2004; **353**: 799–807.

Hotez PJ, Molyneux DH, Fenwick A et al. Control of neglected tropical diseases. *New England Journal of Medicine* 2007; **357**: 1018–1027.

TREMATODES

Trematodes (flukes) are flat leaf-shaped worms. They have complex life cycles, often involving fresh water snails and intermediate mammalian hosts. Disease is caused by the inflammatory response to eggs or to the adult worms.

Water-borne flukes

Schistosomiasis

Schistosomiasis affects over 200 million people in the tropics and subtropics, mostly in sub-Saharan Africa. Chronic infection causes significant morbidity, and after malaria it has the most socioeconomic impact of any parasitic disease. Schistosomiasis is largely a disease of the rural poor, but has also been associated with major development projects such as dams and irrigation schemes.

Parasitology and pathogenesis

There are three species of schistosome which commonly cause disease in man: *Schistosoma mansoni*, *S. haematobium* and *S. japonicum*. *S. mekongi* and *S. intercalatum* also affect man but have very restricted distribution (Fig. 4.38). Eggs are passed in the urine or faeces of an infected person, and hatch in fresh water to release the miracidia. These ciliated organisms penetrate the tissue of the intermediate host, a species of water snail specific to each species of schistosome. After multiplying in the snail, large numbers of fork-tailed cercariae are released back into the water, where they can survive for 2–3 days. During this time the cercariae can penetrate the skin or mucous membranes of the definitive host, man. Transforming into schistosomulae, they pass through the lungs before reaching the portal vein, where they mature into adult worms (the male is about 20 mm long, and the female a little larger). Worms pair in the portal vein before migrating to their final destination: mesenteric veins in the case of *S. mansoni* and *S. japonicum*, and the vesicular plexus for *S. haematobium*. Here they may remain for many years, producing vast numbers of eggs. The majority of these are released in urine or faeces, but a small number become embedded in the bladder or bowel wall, and a few are carried in the circulation to the liver or other distant sites.

The pathology of schistosome infection varies with species and stage of infection. In the early stages there may be local and systemic allergic reactions to the migrating parasites. As eggs start to be deposited there may be a local

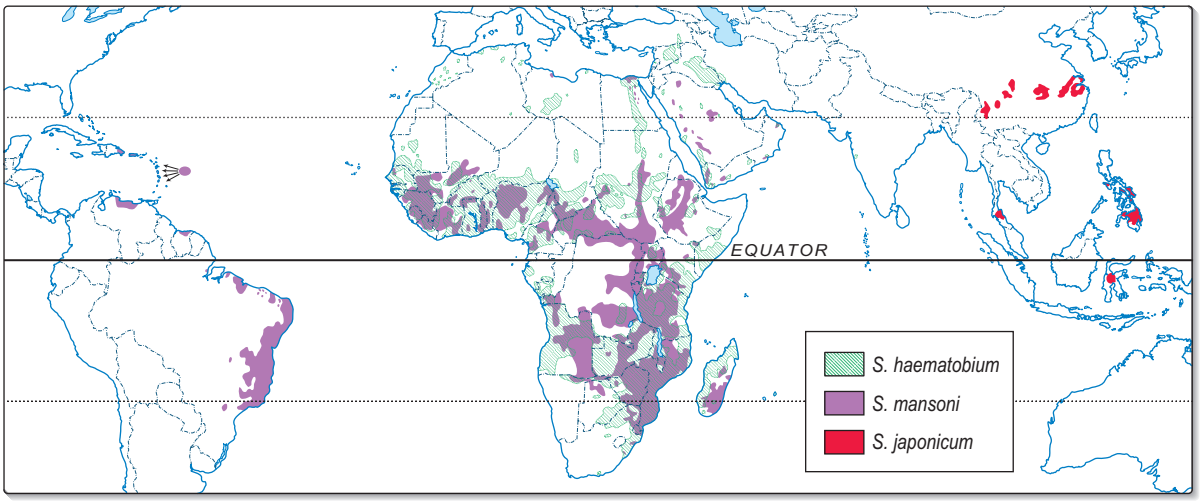


Fig. 4.38 **Schistosomiasis** – geographical distribution.

inflammatory response in the bowel or bladder, while ectopic eggs may produce granulomatous lesions anywhere in the body. Chronic heavy infection, in which large numbers of eggs accumulate in the tissues, leads to fibrosis, calcification, and in some cases, dysplasia and malignant change. Morbidity and mortality are related to duration of infection and worm load, as well as to the species of parasite. Children in endemic areas tend to have the heaviest worm load, because of both increased exposure to infection, and differences in the immune response between adults and children.

Clinical features

Cercarial penetration of the skin may cause local dermatitis ('swimmer's itch'). After a symptom-free period of 3–4 weeks, systemic allergic features may develop, including fever, rash, myalgia and pneumonitis (Katayama fever). These allergic phenomena are common in non-immune travellers, but are rarely seen in local populations, who are usually exposed to infection from early childhood onwards. If infection is sufficiently heavy, symptoms from egg deposition may start to appear 2–3 months after infection.

S. haematobium infection (bilharzia). The earliest symptom is usually painless terminal haematuria. As bladder inflammation progresses there is increased urinary frequency and groin pain. Obstructive uropathy develops, leading to hydronephrosis, renal failure and recurrent urinary infection. There is a strong association between chronic urinary schistosomiasis and squamous cell bladder carcinoma. The genitalia may also be affected, and ectopic eggs may cause pulmonary or neurological disease.

S. mansoni usually affects the large bowel. Early disease produces superficial mucosal changes, accompanied by blood-stained diarrhoea. Later the mucosal damage becomes more marked, with the formation of rectal polyps, deeper ulceration, and eventually fibrosis and stricture formation. Ectopic eggs are carried to the liver, where they cause an intense granulomatous response. Hepatitis is followed by progressive periportal fibrosis, leading to portal hypertension, oesophageal varices and splenomegaly (see p. 348). Hepatocellular function is usually well preserved.

S. japonicum, unlike the other species, infects numerous other mammals apart from man. It is similar to *S. mansoni*, but infects both large and small bowel, and produces a greater number of eggs. Disease therefore tends to be more severe, and rapidly progressive. Hepatic involvement is more common, and neurological involvement is seen in about 5% of cases.

Diagnosis

Schistosomiasis is suggested by relevant symptoms following fresh water exposure in an endemic area. In the early allergic stages the diagnosis can only be made clinically. When egg deposition has started, the characteristic eggs (with a terminal spine in the case of *S. haematobium*, and a lateral spine in the other species) can be detected on microscopy. In *S. haematobium* infection, the best specimen for examination is a filtered mid-day urine sample. Parasites may also be found in semen, and in rectal snip preparations. *S. mansoni* and *S. japonicum* eggs can usually be found in faeces or in a rectal snip. Serological tests are available, and may be useful in the diagnosis of travellers returning from endemic areas, although the test may not become positive for 12 weeks after infection: a parasitological diagnosis should always be made if possible. In chronic disease X-rays, ultrasound examinations and endoscopy may show abnormalities of the bowel or urinary tract, although these are non-specific. Liver biopsy may show the characteristic periportal fibrosis.

Management

The aim of treatment in endemic areas is to decrease the worm load and therefore minimize the chronic effects of egg deposition. It may not always be possible (or even desirable) to eradicate adult worms completely, and reinfection is common. However, a 90% reduction in egg output has been achieved in mass treatment programmes, and in light infections where there is no risk of re-exposure the drugs are usually curative. The most widely used is praziquantel (Table 4.44), which is effective against all species of schistosome, well-tolerated and reasonably cheap.

Prevention

Prevention of schistosomiasis is difficult, and relies on a combination of approaches. Mass treatment of the

FURTHER READING

Fenwick A, Rollinson D, Southgate V. Implementation of human schistosomiasis control: challenges and prospects. *Advances in Parasitology* 2006; **61**: 567–622.

Gryseels B, Polman K, Clerinx J et al. Human schistosomiasis. *Lancet* 2006; **368**:1106–1118.

King CH. Towards the elimination of schistosomiasis. *New England Journal of Medicine* 2009; **360**:106–109.

WHO Partners for Parasite Control. http://www.who.int/wormcontrol.

Table 4.44 Treatment of trematode infections

Parasite	Drug and dose
Schistosoma mansoni	Praziquantel 40 mg/kg single dose*
S. haematobium	Praziquantel 40 mg/kg single dose*
S. japonicum	Praziquantel 60 mg/kg single dose*
Paragonimus spp.	Praziquantel 25 mg/kg 8-hourly for 3 days
Clonorchis sinensis	Praziquantel 25 mg/kg 8-hourly for 2 days
Opisthorcis spp.	Praziquantel 25 mg/kg 8-hourly for 1 day
Fasciolopsis buski	Praziquantel 25 mg/kg 8-hourly for 1 day
Fasciola hepatica	Triclabendazole 10 mg/kg single dose†

*May be split to minimize nausea.
†Repeated if necessary.

population (especially children) will decrease the egg load in the community. Health education programmes, the provision of latrines and access to a safe water supply should decrease the contact with infected water. Attempts to eradicate the snail host have generally been unsuccessful, although man-made bodies of water can often be made less 'snail-friendly'. Travellers should be advised to avoid potentially infected water.

Food-borne flukes

Many flukes infect man via ingestion of an intermediate host, often fresh water fish.

Paragonimiasis

Over 20 million people are infected with lung flukes of the genus *Paragonimus*. The adult worms (of which the major species is *P. westermani*) live in the lungs, producing eggs which are expectorated or swallowed and passed in the faeces. Miracidia emerging from the eggs penetrate the first intermediate host, a fresh water snail. Larvae released from the snail seek out the second intermediate host, fresh water crustacea, in which they encyst as metacercariae. Humans and other mammalian hosts become infected after consuming uncooked shellfish. Cercariae penetrate the small intestinal wall, and migrate directly from the peritoneum to the lungs across the diaphragm. Having established themselves in the lung, the adult worms may survive for 20 years.

The common clinical features are fever, cough and mild haemoptysis. In heavy infections the disease may progress, sometimes mimicking pneumonia or pulmonary tuberculosis. Ectopic worms may cause signs in the abdomen or the brain.

The diagnosis is made by detection of ova on sputum or stool microscopy. Radiological appearances are variable and non-specific. Treatment is with praziquantel, and prevention involves avoidance of inadequately cooked shellfish.

Liver flukes

The human liver flukes, *Clonorchis sinensis*, *Opisthorchis felineus* and *O. viverrini*, are almost entirely confined to East and South East Asia, where they infect more than 20 million people. Adults live in the bile ducts, releasing eggs into the faeces. The parasite requires two intermediate hosts, a fresh water snail and a fish, and humans are infected by consumption of raw fish. The cycle is completed when excysted worms migrate from the small intestine into the bile ducts.

Infection is often asymptomatic, but may be associated with cholangitis and biliary carcinoma. The diagnosis is made by identifying eggs on stool microscopy. Treatment is with praziquantel, and infection can be avoided by cooking fish adequately.

Other fluke infections

Man can also be infected with a variety of animal flukes, notably the liver fluke *Fasciola hepatica*, and the intestinal fluke *Fasciolopsis buski*. Both require a water snail as an intermediate host; cercariae encyst on aquatic vegetation, and then are consumed by animals or man. After ingestion, *F. hepatica* penetrates the intestinal wall before migrating to the liver: during this stage it causes systemic allergic symptoms. After reaching the bile ducts, it causes similar problems to those of the other liver flukes. *F. buski* does not migrate after it excysts, and causes mainly bowel symptoms.

Treatment

See Table 4.44.

CESTODES

Cestodes (tapeworms) are ribbon-shaped worms which vary from a few millimetres to several metres in length. Adult worms live in the human intestine, where they attach to the epithelium using suckers on the anterior portion (scolex). From the scolex arises a series of progressively developing segments, called proglottids. The mature distal segments contain eggs, which may either be released directly into the faeces, or are carried out with an intact detached proglottid. The eggs are consumed by intermediate hosts, after which they hatch into larvae (oncospheres). These penetrate the intestinal wall of the host (pig or cattle) and encyst in the tissues. Man ingests the cysts in undercooked meat, and the cycle is completed when the parasites excyst in the stomach and develop into adult worms in the small intestine. Infections are usually solitary, but several adult tapeworms may coexist. The exceptions to this life cycle are the dwarf tapeworm, *Hymenolepis nana*, which has no intermediate host and is transmitted from person to person by the faeco-oral route, and *Taenia solium*, which produces cysticercosis (see below).

Taenia saginata

T. saginata, the beef tapeworm, may reach a length of several metres. It is common in all countries where undercooked beef is eaten. The adult worm causes few if any symptoms. Infection is usually discovered when proglottids are found in faeces or on underclothing, often causing considerable anxiety. Ova may also be seen on stool microscopy. Infection can be cleared with a single dose of praziquantel (10 mg/kg). It can be prevented by careful meat inspection, or by thorough cooking of beef.

Taenia solium and cysticercosis

T. solium, the pork tapeworm, is generally smaller than *T. saginata*, although it can still reach 6 metres in length. It is

particularly common in South America, South Africa, China and parts of South East Asia. As with *T. saginata* infection is usually asymptomatic. The ova of the two species are identical, but the proglottids can be distinguished on inspection.

Pork tapeworm infection is acquired by eating uncooked pork. Cysticercosis, caused by ingestion of cysts rather than the adult worm, follows the ingestion of eggs from contaminated food and water. Faeco-oral autoinfection can occur but is rare. Patients with tapeworms do not usually develop cysticercosis and patients with cysticercosis do not usually harbour tapeworms. Following the ingestion of eggs, the larvae are liberated, penetrate the intestinal wall, and are carried to various parts of the body where they develop into cysticerci. These are cysts, 0.5–1 cm in diameter, containing the scolex of a new adult worm. Common sites for cysticerci include subcutaneous tissue, skeletal muscle and brain.

Superficial cysts may be felt under the skin, but usually cause no significant symptoms. Cysts in the brain can cause a variety of problems including epilepsy, personality change, hydrocephalus and focal neurological signs (see p. 1157). These may only appear many years after infection.

Muscle cysts tend to calcify, and are often visible on X-rays. Cutaneous cysts can be excised and examined. Brain cysts are less prone to calcification, and are often only seen on CT or MRI scan. Serological tests may support the diagnosis.

Treatment is again with praziquantel or niclosamide. There is no evidence that drug treatment should be accompanied by a purgative, as was previously believed.

Treatment of cerebral cysticercosis (see p. 1157) The role of anthelminthics in cysticercosis remains controversial. Even in neurocysticercosis there is little evidence of benefit, although symptomatic patients with viable neurocysts should probably be treated. Albendazole 15 mg/kg daily for 8–20 days is the drug of choice; the alternative is praziquantel 50 mg/kg daily (in divided doses) for 15 days.

Successful treatment is accompanied by increased local inflammation, and corticosteroids should be given during and after the course of anthelminthic. Anticonvulsants should be given for epilepsy, and surgery may be indicated if there is hydrocephalus. Prevention of cysticercosis depends on good hygiene, as well as on the eradication of human *T. solium* infection.

Diphyllobothrium latum

Infection with the fish tapeworm, *D. latum*, is common in northern Europe and Japan, owing to the consumption of raw fish. The adult worm reaches a length of several metres, but like the other tapeworms usually causes no symptoms. A megaloblastic anaemia (due to competitive utilization of B_{12} by the parasite) may occur. Diagnosis and treatment are the same as for *Taenia* species.

Hydatid disease

Hydatid disease occurs when humans become an intermediate host of the dog tapeworm, *Echinococcus granulosus*. The adult worm lives in the gut of domestic and wild canines, and the larval stages are usually found in sheep, cattle and camels. Man may become infected either from direct contact with dogs, or from food or water contaminated with dog faeces. After ingestion the parasites excyst, penetrate the small intestine wall, and are carried to the liver and other organs in the bloodstream. A slow-growing, thick-walled cyst is formed, inside which further larval stages of the parasite develop. The life cycle cannot be completed unless the cyst is eaten by a dog. Hydatid disease is prevalent in areas where dogs are used in the control of livestock, especially sheep. It is common in Australia, Argentina, the Middle East and parts of East Africa; small foci of infection are still found in North Wales and rural Scotland.

Symptoms depend mainly on the site of the cyst. The liver is the most common organ affected (60%), followed by the lung (20%), kidneys (3%), brain (1%) and bone (1%). The symptoms are those of a slowly growing benign tumour. Pressure on the bile ducts may cause jaundice. Rupture into the abdominal cavity, pleural cavity or biliary tree may occur. In the latter situation, intermittent jaundice, abdominal pain and fever associated with eosinophilia result. A cyst rupturing into a bronchus may result in its expectoration and spontaneous cure, but if secondary infection supervenes a chronic pulmonary abscess will form. Focal seizures can occur if cysts are present in the brain. Renal involvement produces lumbar pain and haematuria. Calcification of the cyst occurs in about 40% of cases.

A related parasite of foxes, *E. multilocularis*, causes a similar but more severe infection, alveolar hydatid disease. These cysts are invasive and metastases may occur.

The diagnosis and treatment of hydatid liver disease are described on page 361.

FURTHER READING

Garcia HH. *Taenia solium* cysticercosis. *Lancet* 2003; **362**: 547–556.

McManus DP, Zhang W, Li J et al. Echinococcosis. *Lancet* 2003; **362**: 1295–1304.

ARTHROPOD ECTOPARASITES

Arthropods, which include the arachnid ticks and mites as well as insects, may be responsible for human disease in several ways.

Local hypersensitivity reactions

Local lesions may be caused by hypersensitivity to allergens in arthropod saliva. This common reaction, known as papular urticaria, is non-specific and is seen in the majority of people in response to the bite of a variety of blood-sucking arthropods including mosquitoes, bugs, ticks, lice and mites. Occasionally tick bites may cause a more severe systemic allergic response, especially in previously sensitized individuals.

Most of these parasites alight on man only to feed, but some species of lice live in very close proximity to the skin: body lice in clothing, and head and pubic lice on human hairs (see p. 1235).

Resident ectoparasite infections

Other ectoparasites are actually resident within the skin, causing more specific local lesions.

Scabies

See page 1234.

Jiggers

Jiggers is due to infection with the jigger flea, *Tunga penetrans*, and is common throughout South America and Africa. The pregnant female flea burrows into the sole of the foot, often between the toes. The egg sac grows to about 0.5 cm in size, before the eggs are discharged onto the ground. The

main danger is bacterial infection or tetanus. The flea should be removed with a needle or scalpel, and the area kept clean until it heals.

Myiasis

Myiasis is caused by invasion of human tissue by the larva of certain flies, principally the Tumbu fly, *Cordylobia anthropophaga* (found in sub-Saharan Africa), and the human botfly, *Dermatobia hominis* (Central and South America). The larvae, which hatch from eggs laid on laundry and linen, burrow into the skin to form boil-like lesions: a central breathing orifice may be visible. Again, the main risk is secondary infection. It is not always easy to extract the larva: covering it with petroleum jelly may bring it up in search of air.

Systemic envenoming

Many arthropods can cause local or systemic illness through envenoming, i.e. injection of venom.

The main role of arthropods in causing human disease is as vectors of parasitic and viral infections. Some of these infections are shown in Table 4.4, and discussed in detail elsewhere.

FURTHER READING

Moss PJ, Beeching NJ. Ectoparasites. In: Cohen W, Powderly V (eds) *Infectious Diseases,* 2nd edn. Edinburgh: Mosby, 2003, 157–164.

Townson H, Nathan MB, Zaim M et al. Exploiting the potential of vector control for disease prevention. *Bulletin of the World Health Organization* 2005; **83**: 942–947.

SEXUALLY TRANSMITTED INFECTIONS

Sexually transmitted infections (STIs) are among the most common causes of illness in the world and remain epidemic in all societies. The public health, social and economic consequences are extensive, both for acute infections and their longer-term sequelae. From 1997 to 2006 the number of new HIV diagnoses in the UK tripled, and new diagnoses of sexually transmitted infections (other than HIV) rose by 63%. STIs have become an increasing world-wide problem.

Those most likely to acquire STIs are young people, homo- and bisexual men and black and ethnic minority populations. Changes in incidence reflect high-risk sexual behaviour and inconsistent use of condoms. Increased travel both within and between countries, recreational drug use, alcohol and more frequent partner change are also implicated. Multiple infections frequently coexist, some of which may be asymptomatic and facilitate spread. Many people attend GUM clinics to seek information, advice and checks of their sexual health, but have no active STI.

Approach to the patient

Patients presenting with possible STIs are frequently anxious, embarrassed and concerned about confidentiality. Staff must be alert to these issues and respond sensitively. The clinical setting must ensure privacy and reinforce confidentiality.

History

The history of the presenting complaint frequently focuses on genital symptoms, the three most common being vaginal discharge (Table 4.45), urethral discharge (Table 4.46) and genital ulceration (Table 4.47). Details should be obtained of any associated fever, pain, itch, malodour, genital swelling, skin rash, joint pains and eye symptoms. All patients should be asked about dysuria, haematuria and loin pain. A full general medical, family and drug history, particularly of any recent antibacterial or antiviral treatment, allergies and use of oral contraceptives, must be obtained. In women, menstrual, contraception and obstetric history should be obtained. Any past or current history of drug misuse should be explored.

A detailed sexual history should be taken and include the number and types of sexual contacts (genital/genital, oral/genital, anal/genital, oral/anal) with dates, partner's sex, whether regular or casual partner, use of condoms and other forms of contraception, previous history of STIs including dates and treatment received, HIV testing and results and hepatitis B vaccination status.

Enquiries should be made concerning travel abroad to areas where antibiotic resistance is known or where particular pathogens are endemic.

Table 4.45 Causes of vaginal discharge

Infective	Non-infective
Candida albicans	Cervical polyps
Trichomonas vaginalis	Neoplasms
Bacterial vaginosis	Retained products (e.g. tampons)
Neisseria gonorrhoeae	Chemical irritation
Chlamydia trachomatis	
Herpes simplex virus	

Table 4.46 Causes of urethral discharge

Infective	Non-infective
Neisseria gonorrhoeae	Physical or chemical trauma
Chlamydia trachomatis	Urethral stricture
Mycoplasma genitilium	Non-specific (aetiology unknown)
Ureaplasma urealyticum	
Trichomonas vaginalis	
Herpes simplex virus	
Human papillomavirus (meatal warts)	
Urinary tract infection (rare)	
Treponema pallidum (meatal chancre)	

Table 4.47 Causes of genital ulceration

Infective	Non-infective
Syphilis:	Behçet's syndrome
Primary chancre	Toxic epidermal necrolysis
Secondary mucous patches	Stevens–Johnson syndrome
Tertiary gumma	Carcinoma
Chancroid	Trauma
Lymphogranuloma venereum	
Donovanosis	
Herpes simplex:	
Primary	
Recurrent	
Herpes zoster	

Examination

General examination must include the mouth, throat, skin and lymph nodes in all patients. Signs of HIV infection are covered on page 187. The inguinal, genital and perianal areas should be examined with a good light source. The groins should be palpated for lymphadenopathy and hernias. The pubic hair must be examined for nits and lice. The external genitalia must be examined for signs of erythema, fissures, ulcers, chancres, pigmented or hypopigmented areas and warts. Signs of trauma may be seen.

In men, the penile skin should be examined and the foreskin retracted to look for balanitis, ulceration, warts or tumours. The urethral meatus is located and the presence of discharge noted. Scrotal contents are palpated and the consistency of the testes and epididymis noted. A rectal examination/proctoscopy should be performed in patients with rectal symptoms, those who practise anoreceptive intercourse and patients with prostatic symptoms. A search for rectal warts is indicated in patients with perianal lesions.

In women, Bartholin's glands must be identified and examined. The cervix should be inspected for ulceration, discharge, bleeding and ectopy and the walls of the vagina for warts. A bimanual pelvic examination is performed to elicit adnexal tenderness or masses, cervical tenderness, and to assess the position, size and mobility of the uterus. Rectal examination and proctoscopy are performed if the patient has symptoms or practises anoreceptive intercourse.

Investigations

Although the history and examination will guide investigation, it must be remembered that multiple infections may coexist, some being asymptomatic. Full screening is indicated in any patient who may have been in contact with an STI.

Asymptomatic STI screening

- A guide for the investigation of asymptomatic patients is given in Table 4.48a.
- All patients attending GUM clinical services should be routinely offered an HIV antibody test as part of the initial screening for sexually transmitted infections.
- Asymptomatic screening for hepatitis viruses identifies patients that fall into groups at increased risk and who, if susceptible, should be offered vaccination.

Screening tests for hepatitis B. Tests should include hepatitis B surface antigen which enables identification of chronic carriers (Hep BSag) and IgG anti-hepatitis B core, a marker of past infection and hence natural immunity. Screening is recommended for: men who have sex with men (MSM) and their sexual partners, sex workers, intravenous drug users and their contacts, recipients of blood/blood products, needle stick recipients, people who have been sexually assaulted, HIV-positive people, sexual partners of those who are HBsAg-positive and individuals from areas where hepatitis B is endemic.

Hepatitis A screening and vaccination should be offered to men who have sex with men in regions where an outbreak of hepatitis A has been reported, injecting drug users, and patients with chronic hepatitis B or C, or other causes of chronic liver disease.

Hepatitis C. Although no vaccine is available, hepatitis C screening is recommended in intravenous drug users, needle stick recipients, patients with HIV and the sexual partners of HCV-positive people.

Investigations for symptomatic patients

The investigations will depend on the clinical presentation. Common presentations include urethral discharge in men, vaginal discharge in women and genital ulceration. Recommended investigations are shown in Table 4.48b.

Treatment, prevention and control

The treatment of specific conditions is considered in the appropriate section. Many GUM clinics keep basic stocks of medication and dispense directly to the patient. Treatment, particularly for those conditions that may be asymptomatic, is a major way to prevent onward transmission.

Tracing the sexual partners of patients is crucial in controlling spread of STIs. The aims are to prevent the spread of infection within the community and to ensure that people with asymptomatic infection are properly treated. Appropriate antibiotic therapy may be offered to those who have had recent intercourse with someone known to have an active infection (epidemiological treatment). Interviewing people about their sexual partners requires considerable tact and sensitivity, and specialist health advisors are available in GUM clinics.

Over half of all STI diagnoses in the UK are in young adults (16–24 years). Prevention starts with education and information. People begin sexual activity at ever-younger ages and education programmes need to include school pupils as well as young adults. Education of health professionals is also crucial. Public health campaigns aimed at informing groups at particular risk are being implemented at a national level. The National Chlamydia Screening Programme in England (NCSP) aims to provide earlier detection and treatment for *Chlamydia* by providing easy access for under 25s to *Chlamydia* testing in community settings.

Avoiding multiple partners, correct and consistent use of condoms and avoiding sex with people who have symptoms of infection may reduce the risks of acquiring an STI. For those who change their sexual partners frequently, regular check-ups (approximately 3-monthly) are advisable. Once people develop symptoms they should be encouraged to seek medical advice as soon as possible to reduce complications and spread to others.

Sexual assault

The medical and psychological management of people who have been sexually assaulted requires particular sensitivity and should be undertaken by an experienced clinician in ways that reduce the risks of further trauma. Specialist sexual assault centres exist, providing multidisciplinary medical, forensic and psychological care in secure and sensitive settings. Post-traumatic stress disorder is common. Although most frequently reported by women, both women and men may suffer sexual assault. Investigations for and treatment of sexually transmitted infections in people who have been raped can be carried out in GUM departments. Collection of material for use as evidence, however, should be carried out within 7 days of the assault by a physician trained in forensic medicine and must take place before any other medical examinations are performed.

History

In addition to the general medical, gynaecological and contraceptive history, full details of the assault, including the

Table 4.48 Recommendations for screening and testing for sexually transmitted infections

(a) Asymptomatic patients

	Gonorrhoea (GC)	*Chlamydia*	Serological investigations
Heterosexual men	Urethral culture for GC *or* Urine NAAT (re-confirm if positive) for GC if no urethral swab	Urethral NAAT for *Chlamydia* *and/or* Urine NAAT	Syphilis: EIA or TPPA or cardiolipin test and TPHA HIV: EIA for antibody (hepatitis viruses if indicated)
Men who have sex with men (MSM)	Urethral, rectal and oro-pharyngeal cultures for GC Urine NAAT if a urethral specimen is not obtained	Urethra and urine: NAAT	Syphilis: EIA or TPPA or cardiolipin test and TPHA Hepatitis B: EIA for HBsAg, HBcAb and HBsAb HIV: EIA for antibody (HAV, HCV if indicated)
Women	Cervix: culture for GC Urine NAAT*	Cervix: NAAT Vagina: self-taken tampons or swabs, vulval–introital–posterior fornix NAAT *Urine: NAAT if urethral specimen not available	Syphilis: EIA or TPPA or cardiolipin test and TPHA HIV: EIA for Antibody (Hepatitis viruses if indicated)

Adapted from UK national guidance, BASHH 2006.
*'Dual' NAAT testing of urine for GC, *Chlamydia* may be used as a screening investigation.

(b) Symptomatic patients

Men Genital discharge	***Urethral swabs:*** microscopy, GC culture, *Chlamydia* NAAT. *Trichomonas* should be sought in men in whom other pathology has been excluded	***Urine:*** for NAATs for *Chlamydia* and GC if no urethral swab is obtainable	***Rectal and oropharyngeal swabs*** for GC and *Chlamydia* culture should be taken if indicated by the history or symptom pattern	***Serology*** for syphilis. HIV: EIA (Hepatitis viruses, if indicated)
Women Genital discharge	***Cervical swabs:*** microscopy and culture for GC and *Chlamydia* NAATs	***Urine*** for *Chlamydia*. NAAT if cervical/vaginal specimen has not been obtained	***Vaginal swabs:*** microscopy and culture for *Candida*, *Trichomonas* and bacterial vaginosis. NAATS for *Chlamydia* and GC if no other material available ***Rectal and oropharyngeal swabs*** for GC and *Chlamydia* culture should be taken if indicated by the history or symptom pattern	***Serology*** for: syphilis HIV: EIA (Hepatitis viruses, if indicated)
Additional investigations in men and women with genital ulceration	***Material from the ulcer:*** microscopy for evidence of early syphilis (dark ground microscopy), Donovanosis, chancroid or LGV PCR techniques where available, for herpes, syphilis and LGV Culture for herpes simplex, *Haemophilus ducreyi*		***Blood tests*** Syphilis: EIA (IgM & IgG) and TPPA and cardiolipin tests for syphilis Herpes simplex virus: IgG by type-specific EIA, Immunoblot or Western blot Complement fixation tests for LGV	

Other investigations depending on clinical circumstances may include:
Cervical cytology
Pregnancy testing
Stools for *Giardia*, *Shigella* or *Salmonella* from those practising oral/anal sex
Smears and swabs from the subpreputial area in men with balanoposthitis (inflammation of glans penis and prepuce) for candidiasis
Midstream urine for microscopy, culture and sensitivity

EIA, enzyme immunoassay; LGV lymphogranuloma venereum; NAAT, nucleic acid amplification test; TPHA, *Treponema pallidum* haemagglutination; TPPA, *Treponema pallidum* particle agglutination assay.
Adapted from UK national guidance, BASHH 2006.

exact sites of penetration, ejaculation by the assailant and condom use should be obtained, together with details of the sexual history both before and after the event.

Examination

Any injuries requiring immediate attention must be dealt with prior to any other examination or investigations. Accurate documentation of any trauma is necessary. Forced oral penetration may result in small palatal haemorrhages. In cases of forced anal penetration, anal examination including proctoscopy should be carried out, noting any trauma.

Investigations

The risk of STI acquisition from rape is variable and depends on the incidence of STI in the population. Ideally a full STI screen at presentation, with appropriate specimens collected from all sites of actual or attempted penetration and a second examination 2 weeks later, is recommended. Nucleic acid

amplification tests (NAATs) for *Neisseria gonorrhoea* and for *Chlamydia trachomatis* can be carried out on urine specimens, avoiding the need for invasive examinations. A positive test should be confirmed by another method for medicolegal purposes. Gram-stained slides of urethral, cervical and rectal specimens for microscopy for gonococci should be performed. Bacterial vaginosis, yeasts and *Trichomonas vaginalis* (TV) tests should be carried out on vaginal material. Syphilis serology should be requested and a serum saved. Hepatitis B, HIV and hepatitis C testing should be offered. Specimens should be identified as having potential medicolegal implications.

Management

Preventative therapy for gonorrhoea and *Chlamydia* is advised using a single dose combination of ciprofloxacin 500 mg and azithromycin 1 g. An accelerated course of Hep B vaccine should be offered and may be of value up to 3 weeks after the event. HIV prophylaxis may be offered within 72 hours of the assault, based upon risk assessment. Postcoital oral contraception may be given within 72 hours of assault. Psychological care provision and appropriate referral to support agencies should be arranged. Sexual partners should be screened and treated if necessary. Follow up at 2 weeks should be arranged to review the patient's needs and the prophylaxis regimens that are in place, with further follow-up as needed. Following sexual assault most people have a range of emotional and psychological reactions and will require varying levels of psychological support. Referral to a specialist centre is recommended.

CLINICAL SYNDROMES

HIV/AIDS

This is discussed in the section starting on page 187.

Gonorrhoea (GC)

Neisseria gonorrhoeae is a Gram-negative intracellular diplococcus (Fig. 4.39), which infects epithelium particularly of the urogenital tract, rectum, pharynx and conjunctivae. Humans are the only host and the organism is spread by intimate physical contact. It is very intolerant to drying and although occasional reports of spread by fomites exist, this route of infection is extremely rare.

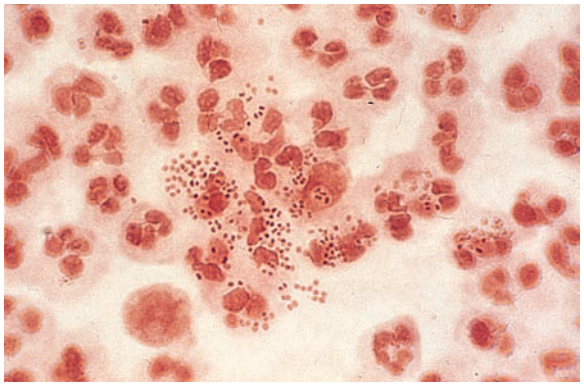


Fig. 4.39 ***Neisseria gonorrhoeae* – Gram-negative intracellular diplococci.** Courtesy of Dr B Goh.

Clinical features

Up to 50% of women and 10% of men are asymptomatic. The incubation period is 2–14 days with most symptoms occurring between days 2 and 5.

- In *men* the most common syndrome is one of anterior urethritis causing dysuria and/or urethral discharge (Fig. 4.40). Complications include ascending infection involving the epididymis or prostate leading to acute or chronic infection. In MSM rectal infection may produce proctitis with pain, discharge and itch.
- In *women* the primary site of infection is usually the endocervical canal. Symptoms include an increased or altered vaginal discharge, pelvic pain due to ascending infection, dysuria and intermenstrual bleeding. Complications include Bartholin's abscesses and in rare cases a perihepatitis (Fitzhugh–Curtis syndrome) can develop. On a global basis GC is one of the most common causes of female infertility. Rectal infection, due to local spread, occurs in women and is usually asymptomatic, as is pharyngeal infection. Conjunctival infection is seen in neonates born to infected mothers and is one cause of ophthalmia neonatorum.

Disseminated GC leads to arthritis (usually monoarticular or pauciarticular) (see p. 535) and characteristic papular or pustular rash with an erythematous base in association with fever and malaise. It is more common in women.

Diagnosis

N. gonorrhoeae can be identified from infected areas by culture on selective media with a sensitivity of at least 95%. Nucleic acid amplification tests (NAATs) using urine specimens are non-invasive and highly sensitive, although may give false positive results. Microscopy of Gram-stained secretions may demonstrate intracellular, Gram-negative diplococci, allowing rapid diagnosis. The sensitivity ranges from 90% in urethral specimens from symptomatic men to 50% in endocervical specimens. Microscopy should not be used for pharyngeal specimens. Blood culture and synovial fluid investigations should be performed in cases of disseminated GC. Coexisting pathogens such as *Chlamydia*, *Trichomonas* and syphilis must be sought.

Fig. 4.40 ***Neisseria gonorrhoeae* – purulent urethral discharge.** Courtesy of Dr B Goh.

Treatment is indicated in those patients who have a positive culture for GC, positive microscopy or a positive NAAT (this should be repeated because of false positives and all positives should be confirmed by cultures). Treatment is given to patients who have had recent sexual intercourse with someone with confirmed GC infection. Although *N. gonorrhoeae* is sensitive to a wide range of antimicrobial agents, antibiotic-resistant strains have shown a recent significant increase. In the UK, up to 12% of strains show resistance to penicillin, 27% to ciprofloxacin and over 40% resistance to tetracycline. Immediate therapy based on Gram-stained slides is usually initiated in the clinic, prior to culture and sensitivity results. Antibiotic choice is influenced by travel history or details known from contacts.

Single-dose oral therapy with cefixime (400 mg), ceftriaxone i.m. (250 mg) or spectinomycin 2 g i.m. successfully treats uncomplicated anogenital infection. Single-dose oral amoxicillin 3 g with probenecid 1 g, ciprofloxacin (500 mg) or ofloxacin (400 mg) may be used in areas with low prevalence of antibiotic resistance. Quinolones are no longer recommended for therapy in the USA due to recent increases in resistance.

Longer courses of antibiotics are required for complicated infections. There should be at least one follow-up assessment, and culture tests should be repeated at least 72 hours after treatment is complete. All sexual contacts should be notified and then examined and treated as necessary.

Chlamydia trachomatis (CT)

The organism has a world-wide distribution. Genital infection with CT is common, with up to 14% of sexually active people below the age of 25 years in the UK infected. It is regularly found in association with other pathogens: 20% of men and 40% of women with gonorrhoea have been found to have coexisting *Chlamydia* infections. In men 40% of non-gonococcal and postgonococcal urethritis is due to *Chlamydia*. As CT is often asymptomatic much infection goes unrecognized and untreated, which sustains the infectious pool in the population. The long-term complications associated with *Chlamydia* infection, especially infertility, impose significant morbidity in the UK. The UK government has introduced a national *Chlamydia* screening programme in an attempt to decrease the rates of asymptomatic infection.

Clinical features

In men CT gives rise to an anterior urethritis with dysuria and discharge; infection is asymptomatic in up to 50% and detected by contact tracing. Ascending infection leads to epididymitis. Rectal infection leading to proctitis occurs in men practising anoreceptive intercourse. *In women* the most common site of infection is the endocervix where it may go unnoticed; up to 80% of infection in women is asymptomatic. Symptoms include vaginal discharge, postcoital or intermenstrual bleeding and lower abdominal pain. Ascending infection causes acute salpingitis. Reactive arthritis (see p. 535) has been related to infection with *C. trachomatis*. Neonatal infection, acquired from the birth canal, can result in mucopurulent conjunctivitis and pneumonia.

Diagnosis

CT is an obligate intracellular bacterium, which complicates diagnosis. The nucleic acid amplification test (NAAT) is now the diagnostic test of choice and is used as the 'standard of care' with a 90–95% sensitivity. Cell culture techniques provide the 'gold standard' and are 100% specific, but are expensive and require considerable expertise. Indirect diagnostic tests include direct fluorescent antibody (DIF) tests, enzyme immunoassays (EIA), but they are less reliable and none is diagnostic.

In men first-voided urine samples are tested, or urethral swabs obtained. In women endocervical swabs are the best specimens, and up to 20% additional positives will be detected if urethral swabs are also taken. Urine specimens are less reliable than endocervical swabs in women but are useful if a speculum examination is not possible.

Ligase chain reaction (LCR) can be used as an alternative to PCR.

Treatment

Tetracyclines or macrolide antibiotics are most commonly used to treat *Chlamydia*. Doxycycline 100 mg 12-hourly for 7 days or azithromycin 1 g as a single dose are both effective for uncomplicated infection. Tetracyclines are contraindicated in pregnancy. Other effective regimens include erythromycin 500 mg four times daily. Routine test of cure is not necessary after treatment with doxycycline or azithromycin, although if symptoms persist or re-infection is suspected then further tests should be taken. NAATs may remain positive for up to 5 weeks after treatment as they pick up material from non-viable organisms. Sexual contacts must be traced, notified and treated, particularly as so many infections are clinically silent.

Urethritis

Urethritis is usually characterized in men by a discharge from the urethra, dysuria and varying degrees of discomfort within the penis. In 10–15% of cases there are no symptoms. A wide array of aetiologies can give rise to the clinical picture and are divided into two broad bands: gonococcal or non-gonococcal urethritis (NGU). NGU occurring shortly after infection with gonorrhoea is known as postgonococcal urethritis (PGU). Gonococcal urethritis and *Chlamydia* urethritis (a major cause of NGU) are discussed above.

Trichomonas vaginalis, *Mycoplasma genitilium*, *Ureaplasma urealyticum* and *Bacteroides* spp. are responsible for a proportion of cases. HSV can cause urethritis in about 30% of cases of primary infection, considerably fewer in recurrent episodes. Other causes include syphilitic chancres and warts within the urethra. Non-sexually transmitted NGU may be due to urinary tract infections, prostatic infection, foreign bodies and strictures.

Clinical features

The urethral discharge is often mucoid and worse in the mornings. Crusting at the meatus or stains on underwear occur. Dysuria is common but not universal. Discomfort or itch within the penis may be present. The incubation period is 1–5 weeks with a mean of 2–3 weeks. Asymptomatic urethritis is a major reservoir of infection. *Reactive arthritis* (see p. 535) causing conjunctivitis with arthritis occurs, particularly in HLA B27-positive individuals.

Diagnosis

NAATs for gonorrhoea and *Chlamydia* should be performed in all men with symptoms of urethritis. In those who are negative, smears should be taken from the urethra when the

patient has not voided urine for at least 4 hours and should be Gram stained and examined under a high-power (×1000) oil-immersion lens. The presence of five or more polymorphonuclear leucocytes per high-power field or a gram-stained preparation from a centrifuged sample of a first passed urine specimen, containing >10 PMNL per high-power is diagnostic. Men who are symptomatic but have no objective evidence of urethritis should be re-examined and tested after holding urine overnight. Cultures for gonorrhoea must be taken together with specimens for *Chlamydia* testing.

Treatment

Therapy for NGU is with either azithromycin 1 g orally as a single dose or doxycycline 100 mg 12-hourly for 7 days. Sexual intercourse should be avoided. The vast majority of patients will show partial or total response. Sexual partners must be traced and treated; *C. trachomatis* can be isolated from the cervix in 50–60% of the female partners of men with PGU or NGU, many of whom are asymptomatic. This causes long-term morbidity in such women, acts as a reservoir of infection for the community, and may lead to re infection in the index case if not treated.

Recurrent/persistent NGU

This is a common and difficult clinical problem which is empirically defined as persistent or recurrent symptomatic urethritis occurring 30–90 days following treatment of acute NGU. It can occur in 20–60% of men treated for acute NGU. The usual time for patients to re-present is 2–3 weeks following treatment. Tests for organisms, e.g. *Mycoplasma*, *Chlamydia* and *Ureaplasma*, are usually negative. It is necessary to document objective evidence of urethritis, check adherence to treatment and establish any possible contact with untreated sexual partners. A further 1 week's treatment with erythromycin 500 mg four times a day for 2 weeks plus metronidazole 400 mg twice daily for 5 days may be given, and any specific additional infection treated appropriately. If symptoms are mild and all partners have been treated, patients should be reassured and further antibiotic therapy avoided. In cases of frequent recurrence and/or florid unresponsive urethritis, the prostate should be investigated and urethroscopy or cystoscopy performed to investigate possible strictures, periurethral fistulae or foreign bodies.

Lymphogranuloma venereum (LGV)

Chlamydia trachomatis types LGV 1, 2 and 3 (different biovars or variant prokaryotic strain) are responsible for this sexually transmitted infection. It is endemic in the tropics, with the highest incidences in Africa, India and South East Asia. There has been a recent upsurge in UK-acquired LGV, particularly amongst HIV-infected men who have sex with men (MSM), presenting with rectal symptoms and associated genital ulceration. Many have also been hepatitis C co-infected. The L2 serovar (different antigenic properties) has been the predominant strain. The Health Protection Agency has enhanced LGV surveillance to track the current UK outbreak.

Clinical features

There are three characteristic stages. The primary lesion is a painless ulcerating papule on the genitalia occurring 7–21 days following exposure. It is frequently unnoticed. A few days to weeks after this heals, regional lymphadenopathy develops. The lymph nodes are painful and fixed and the overlying skin develops a dusky erythematous appearance. Finally, nodes may become fluctuant (buboes) and can rupture. Acute LGV also presents as proctitis with perirectal abscesses, the appearances sometimes resembling anorectal Crohn's disease. The destruction of local lymph nodes can lead to lymphoedema of the genitalia.

Diagnosis

The diagnosis is often made on the basis of the characteristic clinical picture after other causes of genital ulceration or inguinal lymphadenopathy have been excluded. Syphilis and genital herpes must be excluded. The laboratory investigations have become more sensitive and specific:

- Detection of nucleic acid (DNA). Nucleic acid amplification tests for LGV serovar are now available. Positive samples should be confirmed by real-time PCR for LGV-specific DNA.
- Isolation of *C. trachomatis* from clinical lesions in tissue culture remains the most specific test, however sensitivity is only 75–85%.
- *Chlamydia trachomatis* serology. Complement fixation (CF) tests, single L-type immunofluorescence test and micro-immunofluorescence test (micro-IF) (the latter being the most accurate). A fourfold rise in antibody titre in the course of the illness or a single point titre of >1/64 is considered to be diagnostic.

Treatment

Early treatment is critical to prevent the chronic phase. Doxycycline (100 mg twice daily for 21 days) or erythromycin (500 mg four times daily for 21 days) is efficacious. Follow-up should continue until signs and symptoms have resolved, usually 3–6 weeks. Chronic infection may result in extensive scarring and abscess and sinus formation. Surgical drainage or reconstructive surgery is sometimes required. HIV and hepatitis C screening should be recommended. Sexual partners in the 30 days prior to onset should be notified, examined and treated if necessary.

Syphilis

Syphilis is a chronic systemic disease, which is acquired or congenital. In its early stages diagnosis and treatment are straightforward but untreated it can cause complex sequelae in many organs and eventually lead to death. There has been a marked increase in the incidence of syphilis over the past decade. A majority of these infections occur in men who have sex with men, many of whom are co-infected with HIV.

The causative organism, *Treponema pallidum* (TP), is a motile spirochaete that is acquired either by close sexual contact or can be transmitted transplacentally. The organism enters the new host through breaches in squamous or columnar epithelium. Primary infection of non-genital sites may occasionally occur but is rare.

Both acquired and congenital syphilis have early and late stages, each of which has classic clinical features (Table 4.49).

- **Primary**. Between 10 and 90 days (mean 21 days) after exposure to the pathogen a papule develops at the site of inoculation. This ulcerates to become a painless, firm chancre. There is usually painless regional lymphadenopathy in association. The primary lesion

Table 4.49	Classification and clinical features of syphilis
	Clinical features
Acquired	
Early stages	
Primary	Hard chancre
	Painless, regional lymphadenopathy
Secondary	*General*: Fever, malaise, arthralgia, sore throat and generalized lymphadenopathy
	Skin: Red/brown maculopapular non-itchy, sometimes scaly rash; condylomata lata
	Mucous membranes: Mucous patches, 'snail-track' ulcers in oropharynx and on genitalia
Late stages	
Tertiary	*Late benign*: Gummas (bone and viscera)
	Cardiovascular: Aortitis and aortic regurgitation
	Neurosyphilis: Meningovascular involvement, general paralysis of the insane (GPI) and tabes dorsalis
Congenital	Stillbirth or failure to thrive
Early stages	'Snuffles' (nasal infection with discharge)
	Skin and mucous membrane lesions as in secondary syphilis
Late stages	*'Stigmata'*: Hutchinson's teeth, 'sabre' tibia and abnormalities of long bones, keratitis, uveitis, facial gummas and CNS disease

Fig. 4.41 **Rash of secondary syphilis on the palms.** Courtesy of Dr B Goh.

may go unnoticed, especially if it is on the cervix or within the rectum. Healing occurs spontaneously within 2–3 weeks.

- **Secondary**. Between 4 and 10 weeks after the appearance of the primary lesion constitutional symptoms with fever, sore throat, malaise and arthralgia appear. Any organ may be affected – leading, for example, to hepatitis, nephritis, arthritis and meningitis. In a minority of cases the primary chancre may still be present and should be sought.

Signs include:

- Generalized lymphadenopathy (50%).
- Generalized skin rashes involving the whole body including the palms and soles but excluding the face (75%) – the rash, which rarely itches, may take many different forms, ranging from pink macules, through coppery papules, to frank pustules (Fig. 4.41).
- Condylomata lata – warty, plaque-like lesions found in the perianal area and other moist body sites.
- Superficial confluent ulceration of mucosal surfaces – found in the mouth and on the genitalia, described as 'snail track ulcers'.
- Acute neurological signs in less than 10% of cases (e.g. aseptic meningitis).

Untreated early syphilis in pregnant women leads to fetal infection in at least 70% of cases and may result in stillbirth in up to 30%.

- **Latent**. Without treatment, symptoms and signs abate over 3–12 weeks, but in up to 20% of individuals may recur during a period known as early latency, a 2-year period in the UK (1 year in USA). Late latency is based on reactive syphilis serology with no clinical manifestations for at least 2 years. This can continue for many years before the late stages of syphilis become apparent.
- **Tertiary**. Late benign syphilis, so called because of its response to therapy rather than its clinical manifestations, generally involves the skin and the bones. The characteristic lesion, the gumma (granulomatous, sometimes ulcerating, lesions), can occur anywhere in the skin, frequently at sites of trauma. Gummas are commonly found in the skull, tibia, fibula and clavicle, although any bone may be involved. Visceral gummas occur mainly in the liver (hepar lobatum) and the testes.

Cardiovascular and neurosyphilis are discussed on pages 767 and 1157.

Congenital syphilis

Congenital syphilis usually becomes apparent between the second and sixth week after birth, early signs being nasal discharge, skin and mucous membrane lesions, and failure to thrive. Signs of late syphilis generally do not appear until after 2 years of age and take the form of 'stigmata' relating to early damage to developing structures, particularly teeth and long bones. Other late manifestations parallel those of adult tertiary syphilis.

Investigations for diagnosis

Treponema pallidum is not amenable to in vitro culture – the most sensitive and specific method is identification by dark-ground microscopy. Organisms may be found in variable numbers, from primary chancres and the mucous patches of secondary lesions. Individuals with either primary or secondary disease are highly infectious.

Serological tests used in diagnosis are either treponemal-specific or non-specific (cardiolipin test) (Table 4.50):

- **Treponemal specific**. The *T. pallidum* enzyme immunoassay (EIA). *T. pallidum* haemagglutination or particle agglutination assay (TPHA/TPPA) and fluorescent treponemal antibodies absorbed (FTA-abs) test are both highly specific for treponemal disease but will not differentiate between syphilis and other treponemal infection such as yaws. These tests usually remain positive for life, even after treatment.
- **Treponemal non-specific**. The Venereal Disease Research Laboratory (VDRL) and rapid plasma reagin (RPR) tests are non-specific, becoming positive within

Table 4.50 Syphilis serology

Stage of infection	Results			
	EIA	FTA-abs	TPHA/TPPA	VDRL/RPR
Very early primary	–	–	–	–
Early primary	+ (IgM)	+	–	–
Primary	+ (IgM)	+	±	+
Secondary or latent	+	+	+	+
Late latent	+	+	+	±
Treated	+	+	+	±
Biological false–positive	–	–	–	+

EIA, enzyme immunoassay; FTA-abs, fluorescent *Treponema* antibodies absorbed; TPHA/TPPA, *Treponema pallidum* haemagglutination/particle agglutination assay; VDRL, Venereal Disease Research Laboratory; RPR, rapid plasma reagin.

3–4 weeks of the primary infection. They are quantifiable tests which can be used to monitor treatment efficacy and are helpful in assessing disease activity. They generally become negative by 6 months after treatment in early syphilis. The VDRL may also become negative in untreated patients (50% of patients with late-stage syphilis) or remain positive after treatment in late stage. False-positive results may occur in other conditions – particularly infectious mononucleosis, hepatitis, *Mycoplasma* infections, some protozoal infections, cirrhosis, malignancy, autoimmune disease and chronic infections.

The EIA is the screening test of choice and can detect both IgM and IgG antibodies. A positive test is then confirmed with the TPHA/TPPA and VDRL/RPR tests. All serological investigations may be negative in early primary syphilis; the EIA IgM and the FTA-abs being the earliest tests to be positive. The diagnosis will then hinge on positive dark-ground microscopy. Treatment should not be delayed if serological tests are negative in such situations.

Examination of the CSF for evidence of neurosyphilis is indicated in those patients with syphilis who demonstrate neurological signs and symptoms, have a high titre of non-treponemal tests (usually taken to be greater than 1:32), those who have any evidence of treatment failure, and HIV co-infected patients with late latent syphilis or syphilis of unknown duration. A chest X-ray should also be performed.

Treatment

Treponemocidal levels of antibiotic must be maintained in serum for at least 7 days in early syphilis to cover the slow division time of the organism (30 hours). In late syphilis treponemes may divide even more slowly requiring longer therapy.

- *Early syphilis* (primary or secondary) should be treated with long-acting penicillin such as procaine benzylpenicillin (procaine penicillin) (e.g. Jenacillin A which also contains benzylpenicillin) 600 mg intramuscularly daily for 10 days.
- For *late-stage syphilis*, particularly when there is cardiovascular or neurological involvement, the treatment course should be extended for a further week.

For patients sensitive to penicillin, either doxycycline 200 mg daily or erythromycin 500 mg four times daily is given orally for 2–4 weeks depending on the stage of the infection. Non-compliant patients can be treated with a single dose of benzathine penicillin G 2.4 g intramuscularly. Azithromycin is not recommended following evidence of resistance. The diagnosis and management of syphilis in HIV co-infected patients remains unaltered, however if untreated it may advance more rapidly than in HIV-negative patients and have a higher incidence of neurosyphilis. The *Jarisch–Herxheimer reaction*, which is due to release of TNF-α, IL-6 and IL-8, is seen in 50% of patients with primary syphilis and up to 90% of patients with secondary syphilis. It occurs about 8 hours after the first injection and usually consists of mild fever, malaise and headache lasting several hours. In cardiovascular or neurosyphilis the reaction, although rare, may be severe and exacerbate the clinical manifestations. Prednisolone given for 24 hours prior to therapy may ameliorate the reaction but there is little evidence to support its use. Penicillin should not be withheld because of the Jarisch–Herxheimer reaction; since it is not a dose-related phenomenon, there is no value in giving a smaller dose.

The *prognosis* depends on the stage at which the infection is treated. Early and early latent syphilis have an excellent outlook but once extensive tissue damage has occurred in the later stages the damage will not be reversed although the process may be halted. Symptoms in cardiovascular disease and neurosyphilis may therefore persist.

All patients treated for early syphilis must be followed up at regular intervals for the first year following treatment. Serological markers with a fall in titre of the VDRL/RPR of at least fourfold is consistent with adequate treatment for early syphilis. The sexual partners of all patients with syphilis and the parents and siblings of patients with congenital syphilis must be contacted and screened. Babies born to mothers who have been treated for syphilis in pregnancy are retreated at birth.

Chancroid

Chancroid or soft chancre is an acute STI caused by *Haemophilus ducreyi*. Although a less common cause of genital ulceration than HSV-2, it is prevalent in parts of Africa and Asia. It is rare in the UK with cases usually associated with travel to or partners from endemic areas. Epidemiological studies in Africa have shown an association between genital ulcer disease, frequently chancroid and the acquisition of HIV infection. A new urgency to control chancroid has resulted from these observations.

Clinical features

The incubation period is 3–10 days. At the site of inoculation an erythematous papular lesion forms which then breaks down into an ulcer. The ulcer frequently has a necrotic base, a ragged edge, bleeds easily and is painful. Several ulcers may merge to form giant serpiginous lesions. Ulcers appear most commonly on the prepuce and frenulum in men and can erode through tissues. In women the most commonly affected site is the vaginal entrance and the perineum and lesions sometimes go unnoticed.

At the same time, inguinal lymphadenopathy develops (usually unilateral) and can progress to form large buboes which suppurate.

Diagnosis and treatment

Chancroid must be differentiated from other genital ulcer diseases (see Table 4.47). Co-infection with syphilis and herpes simplex is common. Isolation of *H. ducreyi*, a fastidious organism, in specialized culture media is definitive but difficult. Swabs should be taken from the ulcer and material aspirated from the local lymph nodes for culture. Polymerase chain reaction (PCR) techniques are available. Gram stains of clinical material may show characteristic coccobacilli, but this is an insensitive test. Detection of antibody to *H ducreyi* using EIA may be useful for population surveillance but, at an individual level, lacks sensitivity and specificity. A 'probable diagnosis', may be made if the patient has the appropriate clinical picture, without evidence of *T. pallidum* or herpes simplex infection.

Single-dose regimens include azithromycin 1 g orally or ceftriaxone 250 mg i.m. Other regimens include ciprofloxacin 500 mg twice daily for 3 days, erythromycin 500 mg four times daily for 7 days. Clinically significant plasmid-mediated antibiotic resistance in *H. ducreyi* is developing.

Patients should be followed up at 3–7 days, when the ulcers should be responding if the treatment is successful.

Sexual partners should be notified, examined and treated epidemiologically, as asymptomatic carriage has been reported.

HIV-infected patients should be closely monitored, as healing may be slower. Multiple-dose regimens are needed in HIV patients since treatment failures have been reported with single-dose therapy.

Donovanosis

Donovanosis is the least common of all STIs in North America and Europe, but is endemic in the tropics and subtropics, particularly the Caribbean, South East Asia and South India. Infection is caused by *Klebsiella granulomatis*, a short, encapsulated Gram-negative bacillus. The infection was also known as granuloma inguinale. Although sexual contact appears to be the most usual mode of transmission, the infection rates are low, even between sexual partners of many years' standing.

Clinical features

In the vast majority of patients, the characteristic, heaped-up ulcerating lesion with prolific red granulation tissue appears on the external genitalia, perianal skin or the inguinal region within 1–4 weeks of exposure. It is rarely painful. Almost any cutaneous or mucous membrane site can be involved, including the mouth and anorectal regions. Extension of the primary infection from the external genitalia to the inguinal regions produces the characteristic lesion, the 'pseudo-bubo'.

Diagnosis and treatment

The clinical appearance usually strongly suggests the diagnosis but *K. granulomatis* (Donovan bodies) can be identified intracellularly in scrapings or biopsies of an ulcer. Successful culture has only recently been reported and PCR techniques and serological methods of diagnosis are being developed, but none is routinely available.

Antibiotic treatment should be given until the lesions have healed. A minimum of 3 weeks' treatment is recommended. Regimens include doxycycline 100 mg twice daily, co-trimoxazole 960 mg twice daily, azithromycin 500 mg daily or 1 g weekly, or ceftriaxone 1 g daily.

Sexual partners should be notified, examined and treated if necessary.

Herpes simplex (also see p. 1231)

Genital herpes is one of the most common STIs world-wide and the most common ulcerative STI in the UK. Rates have risen substantially, particularly in young adults. Between 2005 and 2006 there has been a 16% increase in diagnoses of genital herpes in 16–19 year olds in the UK. The peak incidence is in 16- to 24-year-olds of both sexes, with women having higher rates of diagnosis than men. Infection, which is lifelong, may be either primary or recurrent. Transmission occurs during close contact with a person who is shedding virus (who may well be asymptomatic). Most genital herpes is due to type 2. Genital contact with oral lesions caused by HSV-1 can also produce genital infection.

Susceptible mucous membranes include the genital tract, rectum, mouth and oropharynx. The virus has the ability to establish latency in the dorsal root ganglia by ascending peripheral sensory nerves from the area of inoculation. It is this ability which allows for recurrent attacks.

Clinical features

Primary genital herpes is usually accompanied by systemic symptoms of varying severity including fever, myalgia and headache. Multiple painful shallow ulcers develop which may coalesce (Fig. 4.42). Atypical lesions are common. Tender inguinal lymphadenopathy is usual. Over a period of 10–14 days the lesions develop crusts and dry. In women with

Fig. 4.42 **Herpes simplex rash on the penis.** Courtesy of Dr B Goh.

vulval lesions the cervix is almost always involved. Rectal infection may lead to a florid proctitis. Neurological complications can include aseptic meningitis and/or involvement of the sacral autonomic plexus leading to retention of urine. Asymptomatic primary infection has been reported but is rare.

Recurrent attacks occur in a significant proportion of people following the initial episode and are more likely with HSV type 2 infections. Precipitating factors vary, as does the frequency of recurrence. A symptom prodrome is present in some people prior to the appearance of lesions. Systemic symptoms are rare in recurrent attacks.

The clinical manifestations in immunosuppressed patients (including those with HIV) may be more severe, asymptomatic shedding increased, and recurrences occurring with greater frequency. Systemic spread has been documented (see p. 202).

Diagnosis

The history and examination can be highly suggestive of HSV infection.

- A firm diagnosis can routinely be made only on the basis of detection of virus from within lesions.
- Culture of virus from lesions used to be the routine approach but is rapidly being replaced by genome detection anyalysis. For either assay modality, swabs should be taken from the base of lesions and placed in viral transport medium. Virus is most easily isolated from new lesions.
- HSV DNA detection by polymerase chain reaction (PCR) has increased the rates of detection compared with virus culture.
- Real-time PCR assays are highly specific although not available in all laboratories.

Type-specific immune responses can be found 8–12 weeks following primary infection and indicate HSV infection at some time. False negatives may be found early after infection. IgM assays are unreliable.

Management

- **Primary.** Saltwater bathing or sitting in a warm bath is soothing and may allow the patient to pass urine with some degree of comfort. Aciclovir 200 mg five times daily, famciclovir 250 mg three times daily or valaciclovir 500 mg twice daily, all for 5 days, are useful if patients are seen whilst new lesions are still forming. If lesions are already crusting, antiviral therapy will do little to change the clinical course. Secondary bacterial infection occasionally occurs and should be treated. Rest, analgesia and antipyretics should be advised. In rare instances when encephalitis is present, patients are admitted to hospital and aciclovir given intravenously.
- **Recurrence.** Recurrent attacks tend to be less severe and can be managed with simple measures such as saltwater bathing. Psychological morbidity is associated with recurrent genital herpes and frequent recurrences impose strains on relationships; patients need considerable support. Long-term suppressive therapy is given in patients with frequent recurrences. An initial course of aciclovir 400 mg twice daily or valaciclovir 250 mg twice daily for 6–12 months significantly reduces the frequency of attacks, although there may still be some breakthrough. Therapy should be discontinued after 12 months and the frequency of recurrent attacks reassessed.

HSV in pregnancy

The potential risk of infection to the neonate needs to be considered in addition to the health of the mother. Infection occurs either transplacentally (very rare) or via the birth canal. If HSV is acquired for the first time during pregnancy, transplacental infection of the fetus may, rarely, occur. Management of primary HSV in the first or second trimester will depend on the woman's clinical condition and aciclovir can be prescribed in standard doses. Aciclovir, given at a dose of 400 mg three times a day to accommodate altered pharmacokinetics of the drug in late pregnancy, during the last 4 weeks of pregnancy may prevent recurrence at term.

Primary acquisition in the third trimester or at term with high levels of viral shedding usually leads to delivery by caesarean section.

For women with previous infection, the risk for the baby acquiring HSV from the birth canal is very low in recurrent attacks. Only those with genital lesions at the onset of labour are delivered by caesarean section. Sequential cultures during the last weeks of pregnancy to predict viral shedding at term are no longer indicated.

Prevention and control

Patients must be advised that they are infectious when lesions are present; sexual intercourse should be avoided during this time or during prodromal stages. Condoms may not be effective as lesions may occur outside the areas covered. Sexual partners should be notified, examined and given information on avoiding infection. Asymptomatic viral shedding is a cause of onward transmission. It is most common in HSV-2, during the first 12 months following infection and in those with frequent symptomatic HSV. Antiviral drugs reduce shedding and onward transmission.

Human papillomavirus (HPV) warts

Anogenital warts are amongst the most common sexually acquired infections. The causative agent is human papillomavirus (HPV) especially types 6 and 11 with types 16 and 18 causing a majority of cases of cervical carcinoma (see p. 108). HPV is acquired by direct sexual contact with a person with either clinical or subclinical infection. Genital HPV infection is common, with only a small proportion of those infected being symptomatic. Neonates may acquire HPV from an infected birth canal, which may result either in anogenital warts or in laryngeal papillomas. The incubation period ranges from 2 weeks to 8 months or even longer.

Clinical features

Warts develop around the external genitalia in women, usually starting at the fourchette, and involve the perianal region. The vagina may be infected. Flat warts may develop on the cervix and are not easily visible on routine examination. Such lesions are associated with cervical intraepithelial neoplasia. In men the penile shaft and subpreputial space are the most common sites, although warts involve the urethra and meatus. Perianal lesions are more common in men who practise anoreceptive intercourse but they can be found in any patient. The rectum may become involved.

Ask the authors
www.studentconsult.com

Warts become more florid during pregnancy or in immunosupressed patients.

Diagnosis

The diagnosis is essentially clinical. It is critical to differentiate condylomata lata of secondary syphilis. Unusual lesions should be biopsied if the diagnosis is in doubt. Up to 30% of patients have coexisting infections with other STIs and a full screen must be performed.

Treatment

Significant failure and relapse rates are seen with current treatment modalities. Choice of treatment will depend on the number and distribution of lesions. Local agents, including podophyllin extract (15–25% solution, once or twice weekly), podophyllotoxin (0.5% solution or 1.5% cream in cycles) and trichloroacetic acid, are useful for non-keratinized lesions. Those that are keratinized respond better to physical therapy, e.g. cryotherapy, electrocautery or laser ablation. Imiquimod, an immune response modifier, induces a cytokine response when applied to skin infected with HPV (5% cream used three times a week) and is indicated in both types of lesion. Podophyllin, podophyllotoxin and imiquimod are not advised in pregnancy. Patients co-infected with HIV may have a poorer response to treatment and higher rates of intraepithelial neoplasia. Sexual contacts should be examined and treated if necessary. In view of the difficulties of diagnosing subclinical HPV, condoms should be used for up to 8 months after treatment. Because of the association of HPV with cervical intraepithelial neoplasia, women with warts and female partners of men with warts are advised to have regular cervical screening, i.e. every 3 years. Colposcopy may be useful in women with vaginal and cervical warts.

Prevention/vaccination

Two vaccines now exist against HPV. One is effective against HPV types 16 and 18 and the second against types 6, 11, 16 and 18. Both vaccines, based on a DNA-free HPV capsid, are strongly immunogenic and act by inducing a neutralizing antibody response and enhancing CD4 memory cell function. They are given over 6 months in three divided doses, with an excellent safety profile. Up to 98% serological responses, maintained for over 4 years, have been reported from early trials. In placebo-controlled clinical trials of the quadrivalent vaccine a highly significant reduction in low- and high-grade cervical dysplasia, vulval pre-cancers and external genital warts has been demonstrated. Vaccination is most beneficial in those who have not yet been exposed to HPV infection, with a recommendation that it be given before people become sexually active. In the UK there is evidence of increasing sexual activity at ever younger ages and vaccination of children is the logical approach. However this has caused considerable ethical debate. Although initially considered to be most useful for women and girls, vaccination of boys to prevent both genital warts and the transmission of oncogenic HPV strains would be a logical public health approach. Routine HPV vaccination commenced in the UK in late 2008, initially for girls of about 12 years of age, with a second catch-up opportunity at about 18 years of age.

Hepatitis B

This is discussed in Chapter 7. Sexual contacts should be screened and given vaccine if they are not immune (see p. 338).

Trichomoniasis

Trichomonas vaginalis (TV) is a flagellated protozoon which is predominantly sexually transmitted. It is able to attach to squamous epithelium and can infect the vagina and urethra. *Trichomonas* may be acquired perinatally in babies born to infected mothers.

Infected women may, unusually, be asymptomatic. Commonly the major complaints are of vaginal discharge, which is offensive and of local irritation. Men usually present as the asymptomatic sexual partners of infected women, although they may complain of urethral discharge, irritation or urinary frequency.

Examination often reveals a frothy yellowish vaginal discharge and erythematous vaginal walls. The cervix may have multiple small haemorrhagic areas which lead to the description 'strawberry cervix'. *Trichomonas* infection in pregnancy has been associated with preterm delivery and low birth weight.

Diagnosis and treatment

Phase-contrast, dark-ground microscopy of a drop of vaginal discharge shows TV swimming with a characteristic motion in 40–80% of female patients. Similar preparations from the male urethra will only be positive in about 30% of cases. Many polymorphonuclear leucocytes are also seen. Culture techniques are good and confirm the diagnosis. *Trichomonas* is sometimes observed on cervical cytology with 60–80% accuracy in diagnosis. New, highly sensitive and specific tests based on polymerase chain reactions are in development.

Metronidazole is the treatment of choice, either 2 g orally as a single dose or 400 mg twice-daily for 7 days. There is some evidence of metronidazole resistance and nimorazole may be effective in these cases. Topical therapy with intravaginal tinidazole can be effective, but if extravaginal infection exists this may not be eradicated and vaginal infection reoccurs. Male partners should be treated, especially as they are likely to be asymptomatic and more difficult to detect.

Candidiasis

Vulvovaginal infection with *Candida albicans* is extremely common. The organism is also responsible for balanitis in men. *Candida* can be isolated from the vagina in a high proportion of women of childbearing age, many of whom will have no symptoms.

The role of *Candida* as pathogen or commensal is difficult to disentangle and it may be changes in host environment which allow the organism to produce pathological effects. Predisposing factors include pregnancy, diabetes, and the use of broad-spectrum antibiotics and corticosteroids. Immunosuppression can result in more florid infection.

Clinical features

In women, pruritis vulva is the dominant symptom. Vaginal discharge is present in varying degree. Many women have only one or occasional isolated episodes. Recurrent candidiasis (four or more symptomatic episodes annually) occurs in up to 5% of healthy women of reproductive age. Examination reveals erythema and swelling of the vulva with broken skin in severe cases. The vagina may contain adherent curdy discharge. Men may have a florid balanoposthitis. More commonly, self-limiting burning penile irritation immediately after

sexual intercourse with an infected partner is described. Diabetes must be excluded in men with balanoposthitis.

Diagnosis

Microscopic examination of a smear from the vaginal wall reveals the presence of spores and mycelia. Culture of swabs should be undertaken but may be positive in women with no symptoms. *Trichomonas* and bacterial vaginosis must be considered in women with itch and discharge.

Treatment

Topical. Pessaries or creams containing one of the imidazole antifungals such as clotrimazole 500 mg single dose used intravaginally are usually effective. Nystatin is also useful.

Oral. The triazole drugs such as fluconazole 150 mg as a single dose or itraconazole 200 mg twice in 1 day are used systemically where topical therapy has failed or is inappropriate. Recurrent candidiasis may be treated with fluconazole 100 mg weekly for 6 months, or clotrimazole pessary 500 mg weekly for 6 months.

The evidence for sexual transmission of *Candida* is slight and there is no evidence that treatment of male partners reduces recurrences in women.

Bacterial vaginosis

Bacterial vaginosis (BV) is a disorder characterized by an offensive vaginal discharge. The aetiology and pathogenesis are unclear but a mixed flora of *Gardnerella vaginalis*, anaerobes including *Bacteroides*, *Mobiluncus* spp. and *Mycoplasma hominis* replaces the normal lactobacilli of the vagina. Amines and their breakdown products from the abnormal vaginal flora are thought to be responsible for the characteristic odour associated with the condition. As vaginal inflammation is not part of the syndrome the term vaginosis is used rather than vaginitis. The condition has been shown to be more common in black women than in white. It is not regarded as a sexually transmitted disease.

Clinical features

Vaginal discharge and odour are the most common complaints although a proportion of women are asymptomatic. A homogeneous, greyish white, adherent discharge is present in the vagina, the pH of which is raised (greater than 5). Associated complications are ill-defined but may include chorioamnionitis and an increased incidence of premature labour in pregnant women. Whether BV predisposes non-pregnant women to upper genital tract infection is unclear.

Diagnosis

Different authors have differing criteria for making the diagnosis of BV. In general it is accepted that three of the following should be present for the diagnosis to be made:

- characteristic vaginal discharge
- the amine test: raised vaginal pH using narrow-range indicator paper (>4.7)
- a fishy odour on mixing a drop of discharge with 10% potassium hydroxide
- the presence of 'clue' cells on microscopic examination of the vaginal fluid.

Clue cells are squamous epithelial cells from the vagina, which have bacteria adherent to their surface, giving a granular appearance to the cell. A Gram stain gives a typical reaction of partial stain uptake.

Treatment

Metronidazole given orally in doses of 400 mg twice daily for 5–7 days is usually recommended. A single dose of 2 g metronidazole is less effective. Topical 2% clindamycin cream 5 g intravaginally once daily for 7 days is effective.

Recurrence is high, with some studies giving a rate of 80% within 9 months of completing metronidazole therapy. There is debate over the treatment of asymptomatic women who fulfil the diagnostic criteria for BV. The diagnosis should be fully discussed and treatment offered if the woman wishes. Until the relevance of BV to other pelvic infections is elucidated, the treatment of asymptomatic women with BV is not to be recommended. There is no convincing evidence that simultaneous treatment of the male partner influences the rate of recurrence of BV, and routine treatment of male partners is not indicated.

INFESTATIONS (see also p. 1235)

Pediculosis pubis

The pubic louse (*Phthirus pubis*) is a blood-sucking insect which attaches tightly to the pubic hair. They may also attach to eyelashes and eyebrows. It is relatively host-specific and is transferred only by close bodily contact. Eggs (nits) are laid at hair bases and usually hatch within a week. Although infestation may be asymptomatic the most common complaint is of itch.

Diagnosis

Lice may be seen on the skin at the base of pubic and other body hairs. They resemble small scabs or freckles but if they are picked up with forceps and placed on a microscope slide will move and walk away. Blue macules may be seen at the feeding sites. Nits are usually closely adherent to hairs. Both are highly characteristic under the low-power microscope.

As with all sexually transmitted infections, the patient must be screened for coexisting pathogens.

Treatment

Both lice and eggs must be killed with 0.5% malathion, 1% permethrin or 0.5% carbaryl. The preparation should be applied to all areas of the body from the neck down and washed off after 12 hours. In a few cases a further application after 1 week may be necessary. For severe infestations, antipruritics may be indicated for the first 48 hours. All sexual partners should be seen and screened.

Scabies

This is discussed on page 1234.

FURTHER READING

British Association for Sexual Health and HIV (BASHH). *BASHH Clinical Effectiveness Guidelines*. London: BASHH, 2008. Available from: http://www.bashh.org/guidelines.asp.

Fenton KA, Breban R, Vardavas R et al. Infectious syphilis in high-income settings in the 21st century. *Lancet Infectious Diseases* 2008; **8**: 244–253.

Gupta R, Warren T, Wald A. Genital herpes. *Lancet* 2007; **370**(9605): 2127–2137.

Low N, Broutet N, Adu-Sarkodie Y et al. Global control of sexually transmitted infections. *Lancet* 2006; **368**: 2001–2016.

Robertson C, Jayasuriya A, Allan PS. Sexually transmitted infections: where are we now? *Journal of the Royal Society of Medicine* 2007; **100**: 414–417.

Stoner B. Current controversies in the management of adult syphilis. *Clinical Infectious Diseases* 2007; **44**(s3): S130–S146.

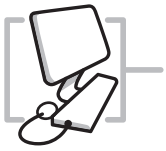
Welch J, Mason F. Rape and sexual assault. *British Medical Journal* 2007; **334**: 1154–1158.

HUMAN IMMUNE DEFICIENCY VIRUS (HIV) AND AIDS

Epidemiology

In 2007 UNAIDS (Joint United Nations Programme on HIV/AIDS) estimated that over 33 million people world-wide were living with HIV infection, 2.5 million people were newly infected during that year and 2.1 million died. On a daily basis world-wide 6800 people acquire HIV and 5700 die of AIDS. Sub-Saharan Africa remains the most seriously affected, but in some areas the numbers of new diagnoses has stabilized. However, in Eastern Europe and parts of central Asia infection rates are rising exponentially. The human, societal and economic costs are huge – 33% of 15-year-olds in high-prevalence countries in Africa will die of HIV. Effective therapy for HIV has reduced mortality and morbidity but not in the poorer parts of the world. Demographics of the epidemic have varied greatly, influenced by social, behavioural, cultural and political factors.

www.studentconsult.com More online



In the UK death rates from HIV have fallen (down from 4.7% in 1997 to 0.95% in 2006), but new diagnoses continue to rise and total numbers have tripled over the past decade. In 2006 73 000 people were HIV-positive of whom 30% were undiagnosed and unaware of their infection. Men who have sex with men (MSM) and culturally diverse heterosexual populations from sub-Saharan Africa, are the two largest groups of people living with HIV in the UK and accessing treatment and care.

The pattern is changing, with over half the new infections with HIV being acquired heterosexually, 75% of which are associated with high seroprevalence countries. The proportion of new diagnoses in the UK in heterosexual women has risen from 25% to 40% since 1997. As mortality rates fall so the population of people with HIV is becoming older, further changing the clinical picture.

Routes of acquisition

Despite the fact that HIV can be isolated from a wide range of body fluids and tissues, the majority of infections are transmitted via semen, cervical secretions and blood.

- *Sexual intercourse (vaginal and anal).* Globally, heterosexual intercourse accounts for the vast majority of infections, and coexistent STIs, especially those causing genital ulceration, enhance transmission. Passage of HIV appears to be more efficient from men to women, and to the receptive partner in anal intercourse, than vice versa. In the UK, sex between men still accounts for over half the infections reported, but there is an increasing rate of heterosexual transmission. In central and sub-Saharan Africa the epidemic has always been heterosexual and more than half the infected adults in these regions are women. South East Asia and the Indian subcontinent are experiencing an explosive epidemic, driven by heterosexual intercourse and a high incidence of other sexually transmitted diseases.
- *Mother to child (transplacentally, perinatally, breast-feeding).* Vertical transmission is the most common route of HIV infection in children. European studies suggest that, without intervention, 15% of babies born to HIV-infected mothers are likely to be infected, although rates of up to 40% have been reported from Africa and the USA. Increased vertical transmission is associated with advanced disease in the mother, maternal viral load, prolonged and premature rupture of membranes, and chorioamnionitis. Transmission can occur in utero although the majority of infections take place perinatally. Breast-feeding has been shown to double the risk of vertical transmission. In the developed world interventions to reduce vertical transmission, including the use of antiretroviral agents, delivery by caesarean section and the avoidance of breast-feeding have led to a dramatic fall in the numbers of infected children. The lack of access to these interventions in resource-poor countries in which 90% of infections occur is a major issue.
- *Contaminated blood, blood products and organ donations.* Screening of blood and blood products was introduced in 1985 in Europe and North America. Prior to this, HIV infection was associated with the use of clotting factors (for haemophilia) and with blood transfusions. In some parts of the world where blood products may not be screened, and in areas where the rate of new HIV infections is very high, transfusion-associated transmission continues to occur.
- *Contaminated needles (intravenous drug misuse, injections, needle-stick injuries).* The practice of sharing needles and syringes for intravenous drug use continues to be a major route of transmission of HIV in both developed countries and parts of South East Asia, Latin America and the states of the former Soviet Union. In some areas, including the UK, successful education and needle exchange schemes have reduced the rate of transmission by this route. Iatrogenic transmission from needles and syringes used in developing countries is reported. Healthcare workers have a risk of approximately 0.3% following a single needle-stick injury with known HIV-infected blood.

There is no evidence that HIV is spread by social or household contact or by blood-sucking insects such as mosquitoes and bed bugs.

The virus

HIV belongs to the lentivirus group of the retrovirus family. There are two types, HIV-1 and HIV-2. HIV-1 is the most frequently occurring strain globally. HIV-2 is almost entirely confined to West Africa, although there is evidence of some spread to Europe, particularly France and Portugal and the

Indian subcontinent. HIV-2 has only 40% structural homology with HIV-1 and although it is associated with immunosupression and AIDS, appears to take a more indolent course than HIV-1. Many of the drugs that are used in HIV-1 are ineffective in HIV-2. The structure of HIV is shown in Figure 4.43.

Retroviruses are characterized by the possession of the enzyme reverse transcriptase, which allows viral RNA to be transcribed into DNA, and thence incorporated into the host cell genome. Reverse transcription is an error-prone process with a significant rate of mis-incorporation of bases. This, combined with a high rate of viral turnover, leads to considerable genetic variation and a diversity of viral subtypes or clades. On the basis of DNA sequencing, HIV-1 is divided into three groups, which probably represent three zoonotic transfers from the chimpanzee but do not differ clinically in humans.

- *Group M (major) subtypes* (95% of infections worldwide) contains at least 10 subtypes (or clades), which are denoted A–J. There is a predominance of subtype B in Europe, North America and Australia, but areas of central and sub-Saharan Africa have multiple M subtypes. Recombination of viral material generates an array of circulating recombinant forms (CRFs), which increases the genetic diversity that may be encountered.
- *Group N (new)* is mostly confined to parts of West Central Africa (e.g. Gabon).
- *Group O (outlier) subtypes* are highly divergent from group M and are largely confined to small numbers centred on Cameroon.

Connections between genetic diversity and biological effects, in particular pathogenicity, rates of transmission and response to therapy are being sought. Increasing numbers of non B subtypes and recombinant forms are seen in the UK.

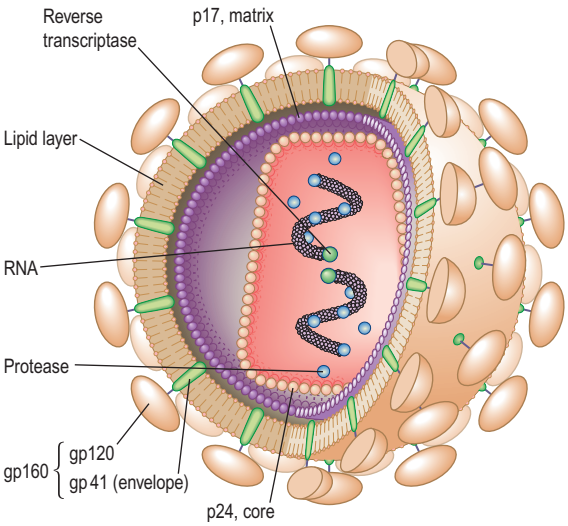


Fig. 4.43 **Structure of HIV.** Two molecules of single-stranded RNA are shown within the nucleus. The reverse transcriptase polymerase converts viral RNA into DNA (a characteristic of retroviruses). The protease includes integrase (p32 and p10). The p24 (core protein) levels can be used to monitor HIV disease. p17 is the matrix protein. gp120 is the outer envelope glycoprotein which binds to cell surface CD4 molecules. gp41, a transmembrane protein, influences infectivity and cell fusion capacity.

Pathogenesis

The interrelationship between HIV and the host immune system is the basis of the pathogenesis of HIV disease. The host cellular receptor that is recognized by HIV surface glycoprotein gp120 is the CD4 molecule, which defines the cell populations that are susceptible to infection (Fig. 4.44). The interaction between CD4 and HIV gp120 surface glycoprotein, together with chemokine co-receptors CCR5 and CXCR4 is responsible for HIV entry into cells. Auxiliary viral proteins such as those coded by the Nef gene have a role in influencing host cell membrane proteins and signal transduction pathways. CD4 receptors and HIV surface glycoprotein interactions mediate the process of syncytium formation, which is a cytopathic effect of HIV infection.

Studies of viral turnover in HIV-infected individuals have demonstrated a virus half-life in the circulation of about 6 hours. To maintain observed levels of plasma viraemia, 10^8–10^9 virus particles need to be released and cleared daily. Virus production by infected cells lasts for about 2 days and is probably limited by the death of the cell, owing to direct HIV effects, linking HIV replication to the process of CD4 destruction and depletion. Studies suggest that immunopathogenesis is a result of defective T cell homeostasis in HIV infection. The progressive and severe depletion of CD4 helper lymphocytes has profound repercussions for the functioning of the immune system (see p. 65). Cell-mediated immunodeficiency, which is the major consequence, leaves the host open to infections with intracellular pathogens, whilst the coexisting antibody abnormalities predispose to infections with capsulated bacteria. HIV also has a direct effect on certain tissues, notably the nervous system. Loss of immunological function also predisposes the host to the development of a variety of malignant tumours.

Diagnosis and natural history (Fig. 4.45)

Reduction in HIV-related morbidity and mortality with effective therapy has highlighted the necessity of testing and diagnosis. Increasing the uptake of HIV testing is a major public health objective. The pre-test counselling and consent required prior to HIV testing is straightforward and can be performed by a wide range of healthcare professionals. HIV testing is being considered on a pan European basis. There is increasing use of the newer point of care HIV testing techniques outside traditional medical settings with the aim of reaching a wider population.

The tests

HIV infection is diagnosed either by the detection of virus-specific antibodies (anti-HIV) or by direct identification of viral material.

Detection of IgG antibody to envelope components. This is the most commonly used marker of infection. The routine tests used for screening are based on ELISA techniques, which may be confirmed with Western blot assays. Up to 3 months may elapse from initial infection to antibody detection (serological latency, or window period). These antibodies to HIV have no protective function and persist for life. As with all IgG antibodies, anti-HIV will cross the placenta. All babies born to HIV-infected women will thus have the antibody at birth. In this situation, anti-HIV antibody is not a reliable marker of active infection and in uninfected babies will be gradually lost over the first 18 months of life.

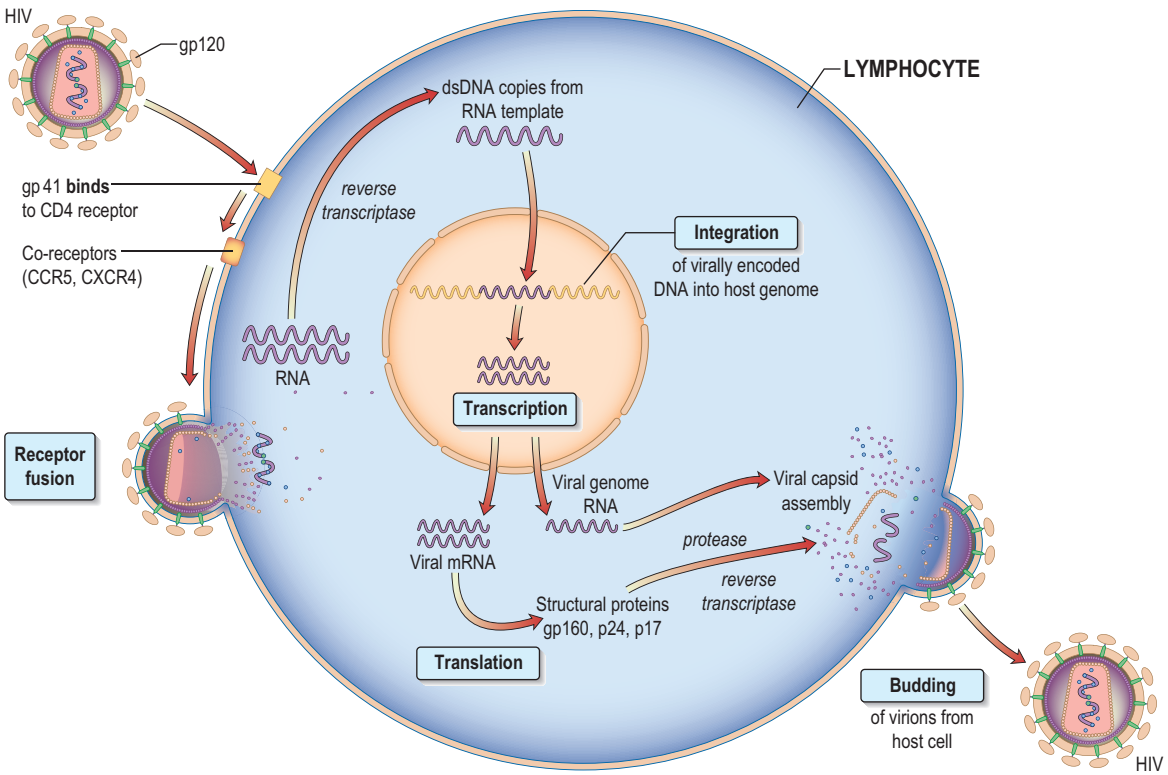


Fig. 4.44 **HIV entry and replication in CD4 T lymphocytes.** The human immune deficiency virus binds to host CD4 molecules via the envelope glycoprotein gp120. gp120 binding to secondary receptors (chemokine receptors) and subsequent conformational change results in fusion between gp41 and the cell membrane. Entry of the viral capsid is followed by uncoating of RNA. DNA copies are made from both RNA templates (reverse transcriptase). The enzyme DNA polymerase from the host cell leads to formation of dsDNA. In the nucleus the virally encoded DNA is inserted into the host genome (integration). Regulatory proteins control transcription (a process in which an RNA molecule is synthesized from a DNA template). The virus is reassembled in the cytoplasm and budded out from the host cell.

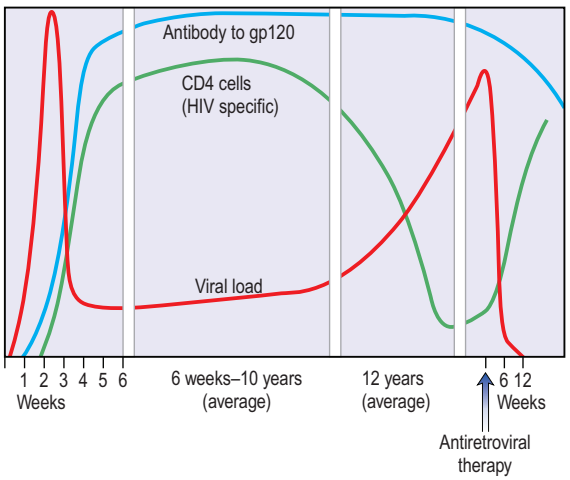


Fig. 4.45 **The immune response to HIV (seroconversion).**

Simple and rapid HIV antibody assays are increasingly available, giving results within minutes. Assays that can utilize alternative body fluids to serum/plasma such as oral fluid, whole blood and urine are now available and home testing kits are being developed. These tests are extremely sensitive and may give false positive results, making it necessary to perform a confirmatory test.

A **S**erologic **T**esting **A**lgorithm for **R**ecent **H**IV **S**eroconversions (STARHS) can be used to identify recently acquired infection. A highly sensitive ELISA that is able to detect HIV antibodies 6–8 weeks after infection is used on blood in patients with positive oral fluid test, in parallel with a less sensitive (detuned) test that identifies later HIV antibodies within 130 days. A positive result on the sensitive test and a negative 'detuned' test are indicative of recent infection, whilst positive results on both tests point to an infection that is more than 130 days old. The major application of this is in epidemiological surveillance and monitoring.

IgG antibody to p24 (anti-p24). This can be detected from the earliest weeks of infection and through the asymptomatic phase. It is frequently lost as disease progresses.

Genome detection assays. Nucleic acid-based assays that amplify and test for components of the HIV genome are available. These assays are used to aid diagnosis of HIV in the babies of HIV-infected mothers or in situations where serological tests may be inadequate, such as in early infection when antibody may not be present, or in subtyping HIV variants for medicolegal reasons. (See the discussion of viral load monitoring, p. 190.)

Viral p24 antigen (p24ag). This is detectable shortly after infection but has usually disappeared by 8–10 weeks after exposure. It can be a useful marker in individuals who have been infected recently but have not had time to mount an antibody response.

Isolation of virus in culture. This is a specialized technique available in some laboratories to aid diagnosis and as a research tool.

Box 4.16 Summary of CDC classification of HIV infection

Absolute CD4 count (/mm^3)	A: Asymptomatic OR persistent generalized lymphadenopathy OR acute seroconversion illness	B: HIV-related conditions,* not A or C	C: Clinical conditions listed in AIDS surveillance case definition (see Box 4.17)
>500	A1	B1	C1
200–499	A2	B2	C2
<200	A3	B3	C4

*Examples of category B conditions include: bacillary angiomatosis, candidiasis (oropharyngeal), constitutional symptoms, oral hairy leucoplakia, herpes zoster involving more than one dermatome, idiopathic thrombocytopenic purpura, listeriosis, pelvic inflammatory disease especially if complicated by tubo-ovarian abscess, peripheral neuropathy.

CLINICAL FEATURES OF HIV INFECTION

The spectrum of illnesses associated with HIV infection is broad and is the result of direct HIV effects, the drugs used to treat the condition, coexisting morbidity and co-infections as well as HIV-associated immune dysfunction. Several classification systems exist, the most widely used being the 1993 Centers for Disease Control (CDC) classification (Box 4.16). This classification depends to a large extent on definitive diagnoses of infection, which makes it more difficult to use in those areas of the world without sophisticated laboratory support. As immunosuppression progresses the patient is susceptible to an increasing range of opportunistic infections and tumours, certain of which meet the criteria for the diagnosis of AIDS (Box 4.17).

Since 1993 the definition of AIDS has differed between the USA and Europe. The USA definition includes individuals with CD4 counts below 200 in addition to the clinical classification based on the presence of specific indicator diagnoses shown in Box 4.17. In Europe the definition remains based on the diagnosis of specific clinical conditions with no inclusion of CD4 lymphocyte counts. Where highly active antiretroviral therapy (HAART) is available and started before the development of severe immunosuppression, progression to AIDS is now uncommon.

Box 4.17 AIDS-defining conditions

- Candidiasis of bronchi, trachea or lungs
- Candidiasis, oesophageal
- Cervical carcinoma, invasive
- Coccidioidomycosis, disseminated or extrapulmonary
- Cryptococcosis, extrapulmonary
- Cryptosporidiosis, chronic intestinal (1-month duration)
- Cytomegalovirus (CMV) disease (other than liver, spleen or nodes)
- CMV retinitis (with loss of vision)
- Encephalopathy, HIV-related
- Herpes simplex, chronic ulcers (1-month duration); or bronchitis, pneumonitis or oesophagitis
- Histoplasmosis, disseminated or extrapulmonary
- Isosporiasis; chronic intestinal (1-month duration)
- Kaposi's sarcoma
- Lymphoma, Burkitt's
- Lymphoma, immunoblastic (or equivalent term)
- Lymphoma (primary) of brain
- *Mycobacterium avium-intracellulare* complex or *M. kansasii*, disseminated or extrapulmonary
- *Mycobacterium tuberculosis*, any site
- *Mycobacterium*, other species or unidentified species, disseminated or extrapulmonary
- *Pneumocystis jiroveci* (formerly *P. carinii*) pneumonia
- Pneumonia, recurrent
- Progressive multifocal leucoencephalopathy
- *Salmonella* septicaemia, recurrent
- Toxoplasmosis of brain
- Wasting syndrome, due to HIV

Incubation

The 2–4 weeks immediately following infection are usually silent both clinically and serologically.

Seroconversion/primary illness

The majority of HIV seroconversions are usually asymptomatic. In a proportion, a self-limiting non-specific illness occurs 6–8 weeks after exposure. Symptoms include fever, arthralgia, myalgia, lethargy, lymphadenopathy, sore throat, mucosal ulcers and occasionally a transient faint pink maculopapular rash. Neurological symptoms are common, including headache, photophobia, myelopathy and neuropathy and in rare cases encephalopathy. The illness lasts up to 3 weeks and recovery is usually complete.

Laboratory abnormalities include lymphopenia with atypical reactive lymphocytes noted on blood film, thrombocytopenia and raised liver enzymes. CD4 lymphocytes may be markedly depleted and the CD4:CD8 ratio reversed. Antibodies to HIV may be absent during this early stage of infection, although the level of circulating viral RNA is high and p24 core protein may be detectable. Patients experiencing a seroconversion illness may have a more rapidly progressive course of infection.

Clinical latency

The majority of people with HIV infection are asymptomatic for a substantial but variable length of time. However, the virus continues to replicate and the person is infectious. Studies suggest a median time of 10 years from infection to development of symptomatic disease, although some patients progress much more rapidly and others have remained symptom-free for up to 15 years. Older age is associated with more rapid progression, and genetic factors may play a part. Gender and pregnancy per se do not appear to influence the rate of progression, although women may fare less well for a variety of reasons. A subgroup of patients with asymptomatic infection have *persistent generalized lymphadenopathy* (PGL), defined as lymphadenopathy (>1 cm) at two or more extra inguinal sites for more than 3 months in the absence of causes other than HIV infection. The nodes are usually symmetrical, firm, mobile and non-tender. There may be associated splenomegaly. The architecture of the nodes shows hyperplasia of the follicles and proliferation of the capillary endothelium. Biopsy is rarely indicated. Similar disease progression has been noted in asymptomatic patients with or without PGL. Nodes may disappear with disease progression.

Symptomatic HIV infection

As HIV infection progresses the viral load rises, the CD4 count falls, and the patient develops an array of symptoms and signs. The clinical picture is the result of direct HIV effects and of the associated immunosuppression.

In an individual patient the clinical consequences of HIV-related immune dysfunction will depend on at least three factors:

- The *microbial exposure of the patient throughout life*. Many clinical episodes represent reactivation of previously acquired infection, which has been latent. Geographical factors determine the microbial repertoire of an individual patient. Those organisms requiring intact cell-mediated immunity for their control are most likely to cause clinical problems.
- The *pathogenicity of organisms encountered*. High-grade pathogens such as *Mycobacterium tuberculosis*, *Candida* and the herpes viruses are clinically relevant even when immunosuppression is mild, and will thus occur earlier in the course of the disease. Less virulent organisms occur at later stages of immunodeficiency.
- The *degree of immunosuppression of the host*. When patients are severely immunocompromised (CD4 count $< 100/mm^3$) disseminated infections with organisms of very low virulence such as *M. avium-intracellulare (MAI)* and *Cryptosporidium* are able to establish themselves. These infections are very resistant to treatment, mainly because there is no functioning immune response to clear organisms. This hierarchy of infection allows for appropriate intervention with prophylactic drugs.

EFFECTS OF HIV INFECTION

Neurological disease

Infection of the nervous tissue occurs at an early stage but clinical neurological involvement increases as HIV advances. This includes *AIDS dementia complex (ADC), sensory polyneuropathy* and *aseptic meningitis* (see p. 1154). These conditions are much less common since the introduction of HAART (highly active antiretroviral therapy). The pathogenesis is thought to be due both to the release of neurotoxic products by HIV itself and to cytokine abnormalities secondary to immune dysregulation.

ADC has varying degrees of severity, ranging from mild memory impairment and poor concentration through to severe cognitive deficit, personality change and psychomotor slowing. Changes in affect are common and depressive or psychotic features may be present. The spinal cord may show vacuolar myelopathy histologically. In severe cases brain CT scan shows atrophic change of varying degrees. MRI changes consist of white matter lesions of increased density on T2-weighted sections. EEG shows non-specific changes consistent with encephalopathy. The CSF is usually normal, although the protein concentration may be raised. Patients with mild neurological dysfunction may be unduly sensitive to the effects of other insults such as fever, metabolic disturbance or psychotropic medication, any of which may lead to a marked deterioration in cognitive functioning.

Sensory polyneuropathy is seen in advanced HIV infection, mainly in the legs and feet, although hands may be affected. Severe forms cause intense pain, usually in the feet, which disrupts sleep, impairs mobility and generally reduces the quality of life.

Autonomic neuropathy may also occur with postural hypotension and diarrhoea. Autonomic nerve damage is found in the small bowel. Didanosine and stavudine produce a similar neuropathy as a major toxic side-effect and are best avoided in patients with HIV neuropathy.

HAART has shown startling improvements in cognitive function in many patients with ADC. It may also have a neuroprotective role.

Eye disease

Eye pathology is usually seen in the later stages. The most serious is cytomegalovirus retinitis (see p. 107), which is sight-threatening. Retinal cotton wool spots due to HIV per se are rarely troublesome but they can be confused with CMV retinitis. Anterior uveitis can present as acute red eye associated with rifabutin therapy for mycobacterial infections in HIV. Steroids used topically are usually effective but modification of the dose of rifabutin is required to prevent relapse. Pneumocystis, toxoplasmosis, syphilis and lymphoma can all affect the retina and the eye may be the site of first presentation.

Mucocutaneous manifestations

(see Table 4.51)

The skin is a common site for HIV-related pathology as the function of dendritic and Langerhans' cells, both target cells for HIV, is disrupted. Delayed-type hypersensitivity, a good indicator of cell-mediated immunity, is frequently reduced or absent even before clinical signs of immunosuppression appear. Pruritus is a common complaint at all stages of HIV. Generalized dry, itchy, flaky skin is typical and the hair may become thin and dry. An intensely pruritic papular eruption favouring the extremities may be found, particularly in patients from African backgrounds. Eosinophilic folliculitis presents with urticarial lesions particularly on the face, arms and legs.

Table 4.51 Some mucocutaneous manifestations of HIV infection (see also Ch. 23)

Skin	Mucous membranes
Dry skin and scalp	Candidiasis:
Onychomycosis	oral
Seborrhoeic dermatitis	vulvovaginal
Tinea:	Oral hairy leucoplakia
cruris	Aphthous ulcers
pedis	Herpes simplex:
Pityriasis:	genital
versicolor	oral
rosea	labial
Folliculitis	Periodontal disease
Acne	Warts:
Molluscum contagiosum	oral
Warts	genital
Herpes zoster:	
mutidermatomal	
disseminated	
Papular pruritic eruption	
Scabies	
Icthyosis	
Kaposi's sarcoma	

Drug reactions with cutaneous manifestations are frequent, with rashes developing notably to sulphur-containing drugs amongst others (see Fig. 23.39). Recurrent aphthous ulceration, which is severe and slow to heal, is common and can impair the patient's ability to eat. Biopsy may be indicated to exclude other causes of ulceration. Topical steroids are useful and resistant cases may respond to thalidomide. In addition to the above the skin is a common site of opportunistic infections (see p. 200).

Haematological complications

These are common in advanced HIV infection.

- *Lymphopenia* – this progresses as the CD4 count falls.
- *Anaemia of chronic HIV infection* is usually mild, normochromic and normocytic.
- *Neutropenia* is common and usually mild.
- *Isolated thrombocytopenia* may occur early in infection and be the only manifestation of HIV for some time. Platelet counts are often moderately reduced but can fall dramatically to 10–20 $\times$ 10^9/L producing easy bleeding and bruising. Circulating antiplatelet antibodies lead to peripheral destruction. Megakaryocytes are increased in the bone marrow but their function is impaired. Effective antiretroviral therapy usually produces a rise in platelet count. Thrombocytopenic patients undergoing dental, medical or surgical procedures may need therapy with human immunoglobulin, which gives a transient rise in platelet count, or be given platelet transfusion. Steroids are best avoided.
- *Pancytopenia* occurs because of underlying opportunistic infection or malignancies, in particular *Mycobacterium avium-intracellulare*, disseminated cytomegalovirus and lymphoma.
- *Other complications*. Myelotoxic drugs include zidovudine (megaloblastic anaemia, red cell aplasia, neutropenia), lamivudine (anaemia, neutropenia), ganciclovir (neutropenia), systemic chemotherapy (pancytopenia) and co-trimoxazole (agranulocytosis).

Gastrointestinal effects (see p. 309)

Weight loss and diarrhoea are common in HIV-infected patients. Wasting is a common feature of advanced HIV infection, which although originally attributed to direct HIV effects on metabolism, is usually a consequence of anorexia. There is a small increase in resting energy expenditure in all stages of HIV, but weight and lean body mass usually remain normal during periods of clinical latency when the patient is eating normally.

Gastrointestinal infections are common. An HIV enteropathy has been described with chronic diarrhoea when no other pathogen has been found. Villous atrophy is seen in the small bowel.

Hypochlorhydria is reported in patients with advanced HIV disease and may have consequences for drug absorption and bacterial overgrowth in the gut.

Rectal lymphoid tissue cells are the targets for HIV infection during penetrative anal sex and may be a reservoir for infection to spread through the body.

Renal complications

HIV-associated nephropathy (HIVAN) (see p. 597), although rare, can cause significant renal impairment, particularly in more advanced disease. It is most frequently seen in black male patients and can be exacerbated by heroin use.

Nephrotic syndrome subsequent to focal glomerulosclerosis is the usual pathology, which may be a consequence of HIV cytopathic effects on renal tubular epithelium. The course is usually relentlessly progressive and dialysis may be required.

Many *nephrotoxic drugs* are used in the management of HIV-associated pathology, particularly foscarnet, amphotericin B, pentamidine and sulfadiazine. Tenofovir has been associated with Fanconi's syndrome.

Respiratory complications

The upper airway and lungs serve as a physical barrier to airborne pathogens and any damage will decrease the efficiency of protection, leading to an increase in upper and lower respiratory tract infections. The sinus mucosa may also function abnormally in HIV infection and is frequently the site of chronic inflammation. Response to antibacterial therapy and topical steroids is usual but some patients require surgical intervention. A similar process is seen in the middle ear, which can lead to chronic otitis media.

Lymphoid interstitial pneumonitis (LIP) is well described in paediatric HIV infection but is uncommon in adults. There is an infiltration of lymphocytes, plasma cells and lymphoblasts in alveolar tissue. Epstein–Barr virus may be present. The patient presents with dyspnoea, and a dry cough, which may be confused with pneumocystis infection (see p. 861). Reticular nodular shadowing is seen on chest X-ray. Therapy with steroids may produce clinical and histological benefit in some patients.

Endocrine complications

Various endocrine abnormalities have been reported, including reduced levels of testosterone and abnormal adrenal function. The latter assumes clinical significance in advanced disease when intercurrent infection superimposed upon borderline adrenal function precipitates clear adrenal insufficiency requiring replacement doses of gluco- and mineralocorticoid. CMV is also implicated in adrenal-deficient states.

Cardiac complications

Cardiovascular pathology is increasingly recognized as a cause of morbidity in people with HIV. Although lipid dysregulation has been associated with antiretroviral medication (ARV), the observation has been made that HDL levels are lower in those with untreated HIV infection than in HIV-negative controls. In a large international study (SMART), ischaemic heart disease was more common in those who took intermittent ARV therapy than in those who maintained viral suppression. The pathogenesis is still unclear. Cardiomyopathy, although rare, is associated with HIV, and may lead to congestive cardiac failure. Lymphocytic and necrotic myocarditis have been described. Ventricular biopsy should be performed to ensure other treatable causes of myocarditis are excluded.

CONDITIONS ASSOCIATED WITH HIV IMMUNODEFICIENCY (see p. 200)

Immunodeficiency allows the development of opportunistic infections (OI) (Table 4.52). These are diseases caused by organisms that are not usually considered pathogenic,

Table 4.52 Major HIV-associated pathogens

Protozoa
Toxoplasma gondii
Cryptosporidium parvum
Microsporidia spp.
Leishmania donovani
Isospora belli

Viruses
Cytomegalovirus
Herpes simplex
Varicella zoster
Human papillomavirus
JC polyoma virus

Fungi and yeasts
Pneumocystis jiroveci (formerly *P. carinii*)
Cryptococcus neoformans
Candida spp.
Dermatophytes (*Trichophyton*)
Aspergillus fumigatus
Histoplasma capsulatum
Coccidioides immitis

Bacteria
Salmonella spp.
Mycobacterium tuberculosis
M. avium-intracellulare
Streptococcus pneumoniae
Staphylococcus aureus
Haemophilus influenzae
Moraxella catarrhalis
Rhodococcus equii
Bartonella quintana
Nocardia

unusual presentations of known pathogens, and the occurrence of tumours that may have an oncogenic viral aetiology. Susceptibility increases as the patient becomes more immunosupressed. CD4 T lymphocyte numbers are used as markers to predict the risk of infection. Patients with CD4 counts above 200 are at low risk for the majority of AIDS-defining OIs. A hierarchy of thresholds for specific infectious risks can be constructed. Mechanisms include defective T cell function against protozoa, fungi and viruses, impaired macrophage function against intracellular bacteria such as *Mycobacteria* and *Salmonella* and defective B cell immunity against capsulated bacteria such as *Strep. pneumoniae* and *Haemophilus*. Many of the organisms causing clinical disease are ubiquitous in the environment or are already carried by the patient.

Diagnosis in an immunosuppressed patient may be complicated by a lack of typical signs, as the inflammatory response is impaired. Examples are lack of neck stiffness in cryptococcal meningitis or minimal clinical findings in early *Pneumocystis jiroveci* pneumonia. Multiple pathogens may coexist. Indirect serological tests are frequently unreliable. Specimens should be obtained from the appropriate site for examination and culture in order to make a diagnosis.

INVESTIGATIONS AND MONITORING OF HIV-INFECTED PATIENTS

Initial assessment

A full medical history, physical examination and laboratory evaluation should be undertaken in all newly diagnosed patients to determine the stage of infection, the presence of co-morbidities and co-infections, and to assess overall physical, mental and sexual health. The initial assessment should also include details of socio-economic situation, relationships, family and social support networks, and substance misuse. Baseline investigations will depend on the clinical setting, but those appropriate for an asymptomatic person in the UK are shown in Box 4.18.

Box 4.18 Baseline investigations in a newly diagnosed asymptomatic patient with HIV infection

Haematology

Full blood count, differential count and film

Biochemistry

Serum, liver and renal function
Fasting serum lipid profile
Fasting blood glucose

Immunology

Lymphocyte subsets
HLA B*5701 status

Virology

HIV antibody (confirmatory)
HIV viral load
HIV genotype
Hepatitis serology (A, B and C)
Cytomegalovirus antibody

Microbiology

Toxoplasmosis serology
Syphilis serology
Screen for other sexually transmitted infections

Other

Cervical cytology
Chest X-ray if indicated

Monitoring

Patients are regularly monitored to assess the progression of the infection and its treatment. Clinical examination will identify signs of immunosuppression (such as oral hairy leucoplakia) and detect early evidence of major opportunistic events. Decisions about appropriate intervention can be made.

Immunological monitoring

CD4 lymphocytes. The absolute CD4 count and its percentage of total lymphocytes falls as HIV progresses. These figures bear a relationship to the risk of the occurrence of HIV-related pathology, with patients with counts below 200 cells at greatest risk. Rapidly falling CD4 counts and those below 350 are an indication for HAART. Factors other than HIV (e.g. smoking, exercise, intercurrent infections and diurnal variation) also affect CD4 numbers. CD4 counts are performed at approximately 3-monthly intervals unless values are approaching critical levels for intervention, in which case they are performed more frequently.

Virological monitoring

Viral load (HIV RNA)

HIV replicates at a high rate throughout the course of infection, with many billion new virus particles being produced daily. The rate of viral clearance is relatively constant in any individual and thus the level of viraemia is a reflection of the rate of virus replication. This has both prognostic and therapeutic value.

The commonly used term 'viral load' encompasses viraemia and HIV RNA levels. Three *HIV RNA assays for viral load* are in current use:

- branched-chain DNA (bDNA)
- reverse transcription polymerase chain reaction (RT-PCR)
- nucleic acid sequence-based amplification (NASBA).

Results are given in copies of viral RNA per millilitre of plasma, or converted to a logarithmic scale, and there is good correlation between tests. The most sensitive test is able to detect as few as 20 copies of viral RNA per millilitre. Transient increases in viral load are seen following immunizations (e.g. for influenza and *Pneumococcus*) or during episodes of acute intercurrent infection (e.g. tuberculosis); and viral load measurements should not be carried out within a month of these events.

By about 6 months after seroconversion to HIV, the viral set-point for an individual is established and there is a correlation between HIV RNA levels and long-term prognosis, independent of the CD4 count. Those patients with a viral load consistently greater than 100 000 copies/mL have a 10 times higher risk of progression to AIDS over the ensuing 5 years than those consistently below 10 000 copies/mL.

HIV RNA is the standard marker of treatment efficacy (see below). Both duration and magnitude of virus suppression are pointers to clinical outcome. None of the currently available therapies appears to be able to suppress viral replication indefinitely, and a rising viral load, in a patient where compliance is assured, indicates drug failure.

Various guidelines exist for viral load monitoring in clinical practice. Baseline measurements are followed by repeat estimations at intervals of 3–4 months, ideally in conjunction with CD4 counts to allow both pieces of evidence to be used together in decision-making. Following initiation of antiretroviral therapy or changes in therapy, effects on viral load should be seen by 4 weeks, reaching a maximum at 10–12 weeks, when repeat viral load testing should be carried out (Fig. 4.45).

Genotype determination

Clear genotype variations exist within HIV, not only between viral subtypes but also with well-identified point mutations associated with antiretroviral drugs. New infections with drug-resistant variants of HIV have been increasing. Viral genotype analysis is recommended for all newly diagnosed patients with HIV. The most appropriate sample is the one closest to the time of diagnosis.

Box 4.19 An approach to sick HIV-positive patients

Potential problems include

Adverse drug reactions
Acute opportunistic infections
Presentation or complication of malignancy
Immune reconstitution phenomenon
Infection in an immunocompromised host
Organic or functional brain disorders
Non-HIV-related pathology must not be forgotten

Full medical history

Remember:
- Antiretroviral drugs, prophylaxis, travel, previous HIV-related pathology, potential source of infectious agents (food hygiene, pets, contacts with acute infections, contact with TB, sexually transmitted infections)
- Secure confidentiality. Check with patient who is aware of HIV diagnosis

Full physical examination

Remember:
- Signs of adverse drug reactions, e.g. skin rashes, oral ulceration
- Signs of disseminated sepsis
- Clinical evidence of immunosuppression, e.g. oral candida, oral hairy leucoplakia
- Focal neurological signs and/or meningism
- Evidence of altered mental state – organic or functional
- Examine:
  - the genitalia, e.g. herpes simplex, syphilis, gonorrhoea
  - the fundi, e.g. CMV retinitis
  - the mouth
- Lymphadenopathy

Immediate investigations

Full blood count and differential count
Liver and renal function tests
Plasma glucose
Blood gases including acid–base balance
Blood cultures, including specimens for mycobacterial culture
Microscopy and culture of available/appropriate specimens: stool, sputum, urine, CSF
Malaria screen in recent travellers from malaria areas
Serological tests for cryptococcal antigen, toxoplasmosis: save serum for viral studies
Chest X-ray
CT/MR scan of brain if focal neurological signs, and ALWAYS before lumbar puncture
NB: Lymphocyte subsets and HIV viral load assays may yield misleading results during intercurrent illness

MANAGEMENT OF THE HIV-INFECTED PATIENT (Box 4.19)

Since the advent of highly active antiretroviral therapy (HAART) in 1997, management in the resource-rich world has moved away from treating opportunistic conditions in immunosuppressed patients towards delivering long-term, effective suppressive therapy.

Nevertheless, there is still no cure for HIV and patients must live with a chronic, infectious and unpredictable condition. Limitations to efficacy include the inability of current drugs to clear HIV from certain intracellular pools, the occurrence of serious drug side-effects, strict adherence requirements, complex drug–drug interactions and the ongoing emergence of resistant viral strains.

The *aims of management* in HIV infection are to maintain physical and mental health, improve the quality of life, increase survival rates, restore and improve immune function, avoid onward transmission of the virus and to provide appropriate palliative support as needed. This requires long-term, maximal suppression of HIV activity using antiretroviral medication and management via a multidisciplinary team approach. Confidentiality must be strictly observed and care taken over establishing who is aware of the patient's diagnosis and who is excluded from that knowledge. Psychological support is needed not only for the patient but also for family, friends and carers. Dietary assessment and advice should be freely accessible. Clear advice on reducing the risk of HIV transmission must be provided and future sexual practices discussed. Information must be available to allow people to make informed choices about childbearing. The implications for existing family members should be considered. General health promotion advice on smoking, alcohol, diet, drug misuse and exercise should be given, particularly in the light of the cardiovascular, metabolic and hepatotoxicity risks associated with HAART (see p. 192).

Antiretroviral drugs (ARVs)

(see Table 4.53)

The treatment of HIV using antiretroviral therapy (HAART) continues to evolve and improve. Increased potency, reduced toxicity, greater convenience of formulation, compounds with different mechanisms of action coupled with improved understanding of drug resistance have combined to make HIV a manageable condition. However, increasing numbers of compounds, the array of drug–drug interactions, for example, combine to make HIV treatment complex, and better clinical outcomes have been linked closely to physician expertise and the numbers of patients under direct care. Regularly updated treatment guidelines are produced in the UK by the British HIV Association (www.BHIVA.org.uk) and in the USA by the Department of Health and Human Services (http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf). The most up-to-date versions can be found on these websites and the current version must be used.

The key practical principles of prescribing ARVs are given in Box 4.20.

Reverse transcriptase inhibitors

- *Nucleoside/nucleotide analogues.* Nucleoside reverse transcriptase inhibitors (NRTIs) inhibit the synthesis of DNA by reverse transcription and also act as DNA chain terminators. NRTIs need to be phosphorylated intracellularly for activity to occur. These were the first group of agents to be used against HIV, initially as monotherapy and later as dual drug combinations. Usually two drugs of this class are combined to provide the 'backbone' of a HAART regimen. Several fixed dose NRTI combinations are available, which helps reduce the pill burden. NRTIs have been associated with mitochondrial toxicity (see p. 198), a consequence of their effect on the human mitochondrial DNA polymerase. Lactic acidosis is a recognized complication of this group of drugs. Nucleotide analogues (nucleotide reverse transcriptase inhibitors (NtRTIs)) have a similar mechanism of action but only require two intracellular phosphorylation steps for activity (as opposed to the three steps for nucleoside analogues). Tenofovir, a monophosphorylated thymidine derivative, is the only licensed compound of this group.
- *Non-nucleoside analogues.* Non-nucleoside reverse transcriptase inhibitors (NNRTIs) interfere with reverse transcriptase by direct binding to the enzyme. They are generally small molecules that are widely disseminated throughout the body and have a long half-life. NNRTIs affect cytochrome P450. They are ineffective against HIV-2. The level of cross-resistance across the class is very high. All have been associated with rashes and elevation of liver enzymes. A second-generation NNRTI, etravirine, has now been developed with activity against viruses resistant to other compounds of the NNRTI class.

Protease inhibitors (PIs)

These act competitively on the HIV aspartyl protease enzyme, which is involved in the production of functional viral proteins, and enzymes. As a consequence, viral maturation is impaired and immature dysfunctional viral particles are produced. Most of the protease inhibitors are active at very low

Table 4.53 Antiretroviral drugs available in the UK

Drug name (alphabetical order)	Dose/regimen	Metabolism	Recognized side-effects
Reverse transcriptase inhibitors			
1. Nucleoside/nucleotide* reverse transcriptase inhibitors (NRTIs)			
Abacavir	600 mg, (1 × 600 mg tablet or 2 × 300 mg tablets) once-daily	No food restrictions	Hypersensitivity reaction, linked with HLAB*5701 positive allele. Fever, rash
Didanosine, DDI	400 mg once daily (>60 kg), 250 mg once daily (<60 kg) 1 capsule/day	30–60 minutes before food	Nausea, diarrhoea, polyneuropathy, pancreatitis. Association with mitochondrial dysfunction and lactic acidosis
Emtricitabine, FTC	200 mg (1 tablet) once daily	No food effects. Renal excretion	Headache, nausea, skin pigmentation. Modify dose if creatinine clearance is <50 mL/min
Lamivudine, 3TC	1 × 300 mg or 2 × 150 mg tablets, once-daily	No food effects, well absorbed high bioavailability	Nausea, headache, rash, peripheral neuropathy, myelosuppression. Association with mitochondrial dysfunction and lactic acidosis
Stavudine, D4T	40 mg (1 tablet) twice daily. Reduce to 30 mg twice daily for persons less than 60 kg	High bioavailability. Competes with zidovudine for intracellular activation by phosphorylation so do not use together	Polyneuropathy. May be able to tolerate reduced dosage. Association with lipoatrophy, megaloblastic changes. Association with mitochondrial dysfunction and lactic acidosis
Tenofovir*	1 × 300 mg tablet, once-daily	After food	Hypophosphataemia, renal toxicity
Zidovudine, AZT	250–300 mg (1 capsule) twice daily	Well absorbed with good bioavailability. No food effects	Nausea, headache, insomnia, skin and nail pigmentation, myelosuppression, megaloblastic changes. Lipoatrophy. Association with mitochondrial dysfunction and lactic acidosis

Table 4.53 Antiretroviral drugs available in the UK—cont'd

Drug name (alphabetical order)	Dose/regimen	Metabolism	Recognized side-effects
2. Non-nucleoside reverse transcriptase inhibitors (NNRTIs)			
Efavirenz	600 mg (1 tablet) once daily, ideally at night	Metabolized by cytochrome P4503A (mixed inducer and inhibitor)	Contraindicated in pregnancy. Rash, toxic epidermal necrolysis, central nervous system effects (vivid dreams, agitation, hallucinations, depression)
Etravirine (TMC125). US and UK licence	200 mg (two 100 mg tablets) twice daily	After food. Substrate of and inducer of CYP3A4, and substrate and inhibitor of CYP2C9, and CYP2C19	Hypersensitivity reaction, erythema multiforme, nausea
Nevirapine	200 mg (1 tablet) once daily for first 14 days then 200 mg (1 tablet) twice daily	High bioavailability, long half-life, wide tissue distribution. Induces its own metabolism, hence dose escalation. No food effects	Rash, toxic epidermal necrolysis, hepatic toxicity
Protease inhibitors (PI)			
Atazanavir	300 mg once daily (two 150 mg capsules) + 100 mg ritonavir once daily	Cytochrome P450 3A4 inhibitor and substrate. Take with food; no concomitant antacids	Hyperbilirubinaemia. Fewer lipid abnormalities than some other PI
Darunavir	600 mg (two 300 mg tablets) + 100 mg ritonavir twice daily	Take with food. Cytochrome P450 3A4 inhibitor and substrate	Rash, toxic epidermal necrolysis and erythema multifome, hyperlipidaemia. Diarrhoea, nausea, headache, lipodystrophy
Fosamprenavir	700 mg (1 tablet) + 100 mg ritonavir twice daily	With food. Cytochrome P450 3A4 inhibitor, inducer, and substrate	Rash, diarrhoea, nausea, vomiting. Hyperlipidaemia, lipodystrophy
Indinavir	800 mg (two 400 mg tablets) + 100 mg ritonavir twice daily	With food. Cytochrome P450 3A4 inhibitor	Nephrolithiasis, hyperbilirubinaemia, alopecia, headache, hyperlipidaemia, lipodystrophy
Lopinavir/ritonavir	400 mg + 100 mg ritonavir (2 tablets) twice daily	With food. Cytochrome P450 3A4 inhibitor and substrate	GI intolerance, nausea, vomiting and diarrhoea. Hyperlipidaemia, lipodystrophy
Ritonavir	600 mg (six 100 mg capsules) twice daily	With food. Cytochrome P450 3A4 – powerful inhibitor	GI intolerance, nausea, vomiting, diarrhoea, peri-oral paraesthesia, hyperlipidaemia, lipodystrophy, metallic taste
Saquinavir	1000 mg (two 500 mg tablets) + 100 mg ritonavir twice daily	Take with food. Cytochrome P450 inhibitor and substrate	GI intolerance, nausea, diarrhoea, headache, hyperlipidaemia, lipodystrophy
Tipranavir	500 mg (two 250 mg capsules) + 200 mg ritonavir twice daily	With food	Rare cases of intracranial bleeding, rash, hyperlipidaemia, lipodystrophy
Fusion inhibitors			
Enfurvitide, T20	90 mg (1 mL) twice daily by s.c. injection		Reaction at the injection site. Hypersensitivity
Co-receptor blockers			
Maraviroc	150–600 mg twice daily, depending on co-administered medications	Blocks the CCR5 chemokine co-receptor. Take with or without food. Cytochrome P450 substrate	Hepatotoxicity, pyrexia, rash, hypotension
Integrase inhibitors			
Raltegravir	400 mg (1 tablet) twice daily	Take with or without food. Metabolism via UGT1A1 glucuronidation	Diarrhoea, nausea, headache, myopathy and rhabdomyolysis
Fixed dose combinations			
Combivir: AZT 300 mg/3TC 150 mg Kivexa: Abacavir 600 mg/3TC 300 mg Truvada: Tenofovir 300 mg/FTC 200 mg Trizivir: Abacavir 300 mg/AZT 300 mg/3TC 150 mg Atripla: Tenofovir 300 mg/FTC 200 mg/Efavirenz 600 mg			

Box 4.20 Prescribing antiretroviral drugs: practice points

Characteristics of antiretroviral agents	Practice points
Must be taken exactly as prescribed	Check for any factors that may compromise accurate adherence Ensure that the proposed drug regimen fits with lifestyles Clarify that the patient is fully conversant with the requirements and understands the reasons for strict adherence Check access to appropriate storage conditions for some agents
Should not be stopped suddenly	Make sure that mechanisms are in place to ensure adequate drug supplies, e.g. regular clinic appointments, repeat prescriptions, home delivery of medications Beware unexpected time away from home, e.g. holidays, intercurrent hospital admissions, immigration detention, police detention If there is an urgent medical indication to stop, obtain advice from specialist physician or pharmacist
Can be compromised by the introduction of other medications, including other ARVs and vice versa	Be careful with enzyme inducers, e.g. rifampicin, rifabutin, warfarin, nevirapine, which will reduce the effective levels of some ARVs Methadone levels may be reduced by efavirenz Some ARVs block the metabolism of other agents which may reach toxic levels Check potential interactions before adding new agents at www.hiv-druginteractions.org Consider therapeutic drug monitoring
Can adversely interact with some herbal, complementary and recreational agents	Herbal remedies that induce cytochrome P450, e.g. St Johns' wort (Hypericum); Chinese herbal remedies will reduce levels of some ARVs Check potential interactions before adding new agents at www.hiv-druginteractions.org Consider therapeutic drug monitoring
May produce additive toxicities when given with other medications	E.g. hepatotoxicity with anti-tuberculosis medication, myelosuppression with chemotherapy or high-dose co-trimoxazole
Are associated with a range of adverse drug reactions which may be confused with other pathology	E.g. rash, fever, nausea, diarrhoea may all be caused by intercurrent pathology and/or ARVs
May exacerbate co-morbidities	Immune reconstitution inflammatory syndrome (IRIS). Examples include hepatic dysfunction due to hepatitis B and C, cardiovascular risk, osteoporosis

concentrations and in vitro are found to have synergy with reverse-transcriptase inhibitors. However there are differences in toxicity, pharmacokinetics, resistance patterns and also cost, which influence prescribing. Cross-resistance occurs across the PI group, which can make it difficult to use the drugs sequentially. There appears to be no activity against human aspartyl proteases (e.g. renin), although there are clinically significant interactions with the cytochrome P450 system. This is used to therapeutic advantage, 'boosting' blood levels of PI by blocking drug breakdown with small doses of ritonavir. PIs have been linked with abnormalities of fat metabolism and control of blood sugar, and some have been associated with deterioration in clotting function in people with haemophilia. Second generation PIs (e.g. darunavir, tipranavir) with activity against viruses resistant to the first generation drugs are now available.

Fusion inhibitors

Enfurvitide is the only licensed compound in this class of agents. It is an injectable peptide derived from HIV gp41 that inhibits gp41-mediated fusion of HIV with the target cell. It is synergistic with NRTIs and PIs. Although resistance to enfurvitide has been described, there is no evidence of cross-resistance with other drug classes. Because it has an extracellular mode of action there are few drug–drug interactions. Side-effects relate to the subcutaneous route of administration in the form of injection site reactions.

Co-receptor blockers

The only drug available so far is maraviroc (Celsentri), a chemokine receptor antagonist which blocks the cellular CCR5 receptor entry by CCR5 tropic strains of HIV. These strains are found in earlier HIV infection and, with time adaptations (against which maraviroc is ineffective) allow the CXCR4 receptor to become the more dominant form. The drug is metabolized by CYP p450 (3A), giving the potential for drug–drug interactions.

Integrase inhibitors

These drugs act as a selective inhibitor of HIV integrase, which blocks viral replication by preventing insertion of HIV DNA into the human DNA genome. Three compounds, have so far been developed, although only one, raltegravir, has a European licence. Raltegravir is metabolized by glucuronidination, and does not require retroviral drug boosting. Toxicity profile appears favourable in the short term and long-term durability data is awaited. It is effective in both treatment of experienced and naïve patients.

Starting therapy

Treatment regimens for HIV infection require a long-term commitment to high levels of adherence. Risks of therapy include serious side-effects, drug interactions and the potential for development of resistant viral strains. The full involvement of patients in therapeutic decision-making is essential for success. Various national guidelines and treatment frameworks exist (e.g. BHIVA Guidelines, DHHS Guidelines and IAS Recommendations). A combination of clinical assessment and laboratory marker data, including viral load and CD4 counts, together with individual circumstances, should guide therapeutic decision-making. The current UK recommendations are shown in Table 4.54. In situations where therapy is recommended but the patient elects not to start, then more intensive clinical and laboratory monitoring is advisable.

Clear clinical benefit has been demonstrated with the use of antiretroviral drugs in advanced HIV disease. Treatment should be recommended for all patients with symptomatic HIV disease, AIDS or a CD4 count that is consistently below

FURTHER READING

Dolin R. A new class of anti-HIV therapy. *New England Journal of Medicine* 2008; **359**: 1509–1511.

Table 4.54 When to start antiretroviral therapy

Primary HIV infection	Treatment in clinical trial *or* if there is neurological involvement *or* if the CD4 < 200 for >3/12 *or* if an AIDS-defining illness is present
Established HIV infection	
CD4 < 250	Treat
CD4 251–350	Treat as soon as patient ready
CD4 > 350	Consider enrolment into 'when to start' trial
AIDS diagnosis/CDC stage C	Treat (except for TB when CD4 > 350)

Modified from Gazzard BG. British HIV Association guidelines for the treatment of HIV-1-infected adults with antiretroviral therapy 2008. *HIV Medicine* 2008; **9**: 563–608. http://www3.interscience.wiley.com/journal/118491579/home.

200 cells/mm^3. In such situations there is a significant risk of serious HIV-associated morbidity and mortality.

In asymptomatic patients the absolute CD4 count is the key investigation used to guide treatment decisions. Counts below 200 cells/mm^3 are strongly associated with disease progression and death. The outcome for patients initiating therapy below 200 CD4 cells/mm^3 is not as good as for those who start at higher counts. The UK recommendation is that therapy should be started before the count falls below 200. Due to late presentation of HIV in the UK almost two-thirds of patients initiating ARVs in 2006 had a CD4 count below 200.

There are no randomized trials of when to start therapy in patients with higher CD4 cell counts. The risk of disease progression for individuals with a count greater than 350 is low, and has to be balanced against ARV therapy toxicity and development of resistance. Earlier intervention at higher CD4 counts is being debated but the recommendation remains to begin therapy at a CD4 count between 350 and 201 cells/mm^3 and those with a higher risk of disease progression, e.g. with high viral loads (>60 000 copies/mL) or rapidly falling CD4 count (losing more than 80 cells/year). Co-infection with hepatitis C virus is also a factor for earlier intervention (see p. 340).

Treatment for primary HIV infection is only recommended either within a clinical trial or to alleviate symptoms.

Special situations (seroconversion, pregnancy, post-exposure prophylaxis) in which antiretroviral agents may be used are described on page 198.

Choice of drugs

The drug regimen used for starting therapy must be individualized to suit each patient's needs. Treatment is initiated with three drugs, two NRTIs in combination, with either a NNRTI or a boosted protease inhibitor (Tables 4.53 and 4.55).

Nucleoside reverse transcriptase inhibitor (NRTI)

Two NRTIs to form a backbone is increasingly influenced by both efficacy and ease of administration. The availability of the once daily one-tablet fixed-dose combinations, *Truvada* (TDF/FTC) and *Kivexa* (ABC/3TC) has led to the majority of patients who are naïve to medication being prescribed one of these as their 2NRTI backbone. Kivexa should only be used in those who are HLA-B*5701 negative. *Combivir* (ZDV/3TC), with a twice-daily dosing schedule and poorer toxicity profile, is now less frequently prescribed. Comparison of Truvada with Combivir showed more virological failure in the Combivir arm. Head to head efficacy data comparing Truvada and Kivexa in naïve patients has demonstrated non-inferiority of Kivexa at viral levels below 100 000 copies/mL. In patients with high viral levels Kivexa should be reserved for use when Truvada is contraindicated. *Stavudine* is associated with lipodystrophy and peripheral neuropathy and is rarely used as first line therapy in the UK, although it continues to be used widely in other parts of the world.

Table 4.55 Initial HAART regimens – choice of initial therapy. Preferred regimens

Choose one drug from columns A, B and C			
Regimen	**A**	**B**	**C**
Preferred			
	Efavirenz	Tenofovir*†	Lamivudine‡§
		Abacavir§	Emtricitabine†
Alternative			
	Lopinavir/r	Didanosine	
	Fosamprenavir	Zidovudine‡	
	Atazanavir/r‖		
	Saquinavir/r		
Specific groups			
	Nevirapine*		
	Atazanavir‖¶		

*Only when CD4 < 250 cells/mm^3 in females, <400 cells/mm^3 in males.
†Co-formulated as Truvada®.
‡Co-formulated as Combivir®.
§Co-formulated as Kivexa®.
‖Currently unlicensed for naïve patients in the UK.
¶Where established cardiovascular disease risk factors and a PI required.
Modified from Gazzard BG. British HIV Association guidelines for the treatment of HIV-1-infected adults with antiretroviral therapy 2008. *HIV Medicine* 2008; **9**: 563–608. http://www3.interscience.wiley.com/journal/118491579/home.

Non-nucleoside reverse transcriptase inhibitor (NNRTI)

The decision about use of NNRTI or boosted PI will depend on the particular circumstances of each patient but in the UK, based on clinical trial data, an NNRTI-based regimen is most commonly prescribed to treat naïve patients.

Efavirenz is the recommended option in the UK, having demonstrated good durability over time, potency at low CD4 counts and in high viral loads. Efavirenz has the advantage of once-daily dosing but is associated with CNS side-effects such as dysphoria and insomnia, and is contraindicated in pregnancy. A new fixed dose preparation of efavirenz co-formulated with Truvada (Atripla) allows for a 'one pill once a day' regimen.

Nevirapine is of equivalent potency to efavirenz but has a higher incidence of hepatoxicity and rash and is taken twice daily. Toxicity is greater in women and in those with higher CD4 counts. It is contraindicated in women with CD4 counts above 250 cells and in men with counts above 400. However it a useful alternative to efavirenz if CNS side-effects are troublesome and in women with lower CD4 counts who wish to conceive.

Etravirine is a new, second generation NNRTI, active against drug-resistant strains and useful in treatment of experienced patients.

Ask the authors
www.studentconsult.com

Box 4.21 Monitoring patients on highly active antiretroviral therapy (HAART)

Clinical history including:

- Assessment of adherence to medication
- Evidence of antiretroviral therapy (ART) toxicity

Physical examination:

- Blood pressure
- Weight
- HIV viral load
- Lymphocyte subsets
- Full blood count
- Urinalysis if taking tenofovir
- Liver and renal function
- Fasting lipid profile
- Fasting blood glucose

Additional investigations (depending on the clinical picture) may include:

- HIV genotype if there is evidence of virological failure
- Therapeutic drug level monitoring

Protease inhibitor

This class of drugs have demonstrated excellent efficacy in clinical practice. PIs combined with a low dose of ritonavir (a boosting PI) provides a pharmacokinetic advantage and is now most commonly used in naïve patients. Using this approach, the half-life of the active drug is increased, allowing greater drug exposure, fewer pills with enhanced potency and the risk of resistance minimized. The disadvantage of this approach is a possible risk of greater lipid abnormalities, particularly raised fasting triglycerides.

Either *lopinavir/ritonavir* or *fosamprenavir/ritonavir* (both taken twice daily) are recommended for first line therapy, with *saquinavir/ritonavir* as an alternative. All three regimens commonly cause GI disturbance and lipid abnormalities. *Atazanavir/ritonavir* is a once-daily regimen, but atazanavir (not currently licensed for naïve use in the UK) produces an isolated, unconjugated hyperbilirubinaemia. *Tipranavir* and *darunavir* both active against some PI-resistant strains are currently only used in treatment-experienced patients although trial data is emerging on use of darunavir in naïve patients.

In summary, recent comparative data show that three NRTIs, although able to produce good levels of viral suppression in treatment-naïve patients, are less efficacious than combining two NRTIs with either an NNRTI or a PI and are no longer recommended. Despite interest in the use of PI monotherapy in some situations, the current data does not support its use in treatment-naïve patients outside a clinical trial.

Monitoring therapy (Box 4.21)

Virological success of HAART is taken to be a viral load of less than 50 copies/mL within 3–6 months of starting therapy. Once therapy has been initiated the viral load should be measured at 4 weeks to assess efficacy. The durability of the virological response is linked both to the rate of viral load fall and the absolute nadir attained. Viral load and CD4 count should be measured at 12 weeks and then at 3-monthly intervals. Regular clinical assessment should include review of adherence to and tolerability of the regimen, weight, blood pressure and urinalysis. Patients should be monitored for drug toxicity, including full blood count, liver and renal function and fasting lipids and glucose levels.

Box 4.22 Mechanisms and implications of HIV drug resistance

1. HIV replicates rapidly and inaccurately. Replication in the presence of antiretroviral drugs leads to a selection pressure for those mutations which can survive, i.e. selects for drug resistance
2. Specific point mutations in the viral reverse transcriptase, protease and integrase genes correlate with reduced drug sensitivity and can be identified by genotyping the virus
3. Inadequate antiretroviral (ARV) drug levels both fail to suppress viral replication/viral load and precipitate drug resistance
4. Inadequate drug levels can result from poor adherence, altered GI tract absorption, increased drug breakdown and drug–drug interactions
5. Some ARVs, especially NNRTIs, have a low genetic barrier, i.e. a small number of mutations that occur rapidly can quickly result in high levels of resistance
6. Stopping drugs with a long half-life can leave a subtherapeutic drug tail for long enough for resistant strains to develop
7. In patients on stable ARV therapy, the introduction of new drugs that have an impact on cytochrome P450 pathways can lead to dangerous alterations in ARV drug levels
8. Therapeutic drug monitoring (TDM) may be a useful investigation for some drugs in some circumstances
9. Without drug selection pressure wild type virus reasserts itself and resistant variants no longer make up the major circulating viral strains and may not be found on investigation. This means that genotyping should if possible be carried out on specimens obtained whilst on therapy
10. Resistant variants survive and are archived. They reappear if the drug selection pressure is reintroduced. This means that all previous genotypes need to be considered when assessing virological failure and planning new therapy

Drug resistance (Box 4.22)

Resistance to ARVs results from mutations in the protease and reverse transcriptase genes of the virus. HIV has a rapid turnover with 10^8 replications occurring per day. The error rate is high, resulting in genetic diversity within the population of virus in an individual, which will include drug-resistant mutants. When drugs only partially inhibit virus replication there will be a selection pressure for the emergence of drug-resistant strains. The rate at which resistance develops depends on the frequency of pre-existing variants and the number of mutations required. Resistance to most NRTIs and PIs occurs with an accumulation of mutations, whilst a single-point mutation will confer high-level resistance to NNRTIs. There is evidence for the transmission of HIV strains that are resistant to all or some classes of drugs. Studies of primary HIV infection have shown prevalence rates between 2% and 20%. Prevalence of primary mutations associated with drug resistance in chronically infected patients not on treatment ranges from 3% to 10% in various studies.

HIV antiretroviral drug resistance testing has become routine clinical management in patients at diagnosis/before starting therapy and for whom therapy is failing. Genotypic assays to determine the genetic structure of the reverse transcriptase (RT) and protease genes of HIV are available. The tests are based on PCR amplification of virus and give an indirect measure of drug susceptibility in the predominant variants. Such assays are limited both by the starting concentration of virus, most assays requiring at least 1000 copies/mL of blood, and by their poor ability to detect minority strains.

For results to be useful in situations where therapy is failing, samples must be analysed when the patient is on therapy, as once the selection pressure of therapy is withdrawn, wild type virus becomes the predominant strain and resistance mutations present earlier may no longer be detectable.

Databases containing nearly all published HIV (amongst others) and protease sequences, and associated resistance patterns, are maintained in real time by Stanford University and can be accessed at http://hivdb.stanford.edu.

Phenotypic assays provide a more direct measure of susceptibility but the complexity of the assays limits availability.

Drug interactions

Drug therapy in HIV is highly complex and the potential for clinically relevant drug interactions is substantial. Both PIs and NNRTIs are able to variably inhibit and induce cytochrome P450, influencing both their own and other drug metabolic rates. Both inducers and inhibitors of cytochrome P450 are sometimes prescribed simultaneously. Induction of metabolism may result in sub-therapeutic antiretroviral drug levels with the risk of treatment failure and development of viral resistance, whilst inhibition can raise drug levels to toxic values and precipitate adverse reactions.

Conventional (e.g. rifamycins) and complementary therapies (e.g. St John's Wort) affect cytochrome P450 activity and may precipitate substantial drug interactions. Therapeutic drug monitoring (TDM) indicating peak and trough plasma levels may be useful in certain settings.

Potential interactions can be checked using the online tool maintained by Liverpool University at www.hiv-druginteractions.org.

Adherence

Adherence to treatment is pivotal to success. Levels of adherence below 95% have been associated with poor virological and immunological responses. Poor absorption and low bioavailability mean that for some compounds trough levels are barely adequate to suppress viral replication and missing even a single dose will result in plasma drug levels falling dangerously low. Patchy adherence facilitates the emergence of drug-resistant variants, which in time will lead to virological treatment failure.

Factors implicated in poor adherence may be associated with the medication, with the patient or with the provider. The former include side-effects associated with medications, the degree of complexity and pill burden and inconvenience of the regimen. Patient factors include the level of motivation and commitment to the therapy, psychological well-being, the level of available family and social support, and health beliefs. Supporting adherence is a key part of clinical care and specific guidelines are available (BHIVA 2004). Education of patients about their condition and treatment is a fundamental requirement for good adherence, as is education of clinicians in adherence support techniques. Provision of acute and ongoing multidisciplinary support for adherence within clinical settings should be universal. Medication-alert devices may be useful for some patients.

Treatment failure

Failure of antiretroviral treatment, i.e. persistent viral replication causing immunological deterioration and eventual clinical evidence of disease progression, is caused by a variety of factors, e.g. limited drug potency and food or other medication may compromise drug absorption. There may be drug interactions or limited penetration of drug into sanctuary sites such as the CNS, permitting viral replication. Side-effects and other patient-related elements contribute to poor adherence.

Changing therapy

A rise in viral load, a falling CD4 count or new clinical events that imply progression of HIV disease are all reasons to review therapy. Reasons for treatment failure include the emergence of resistant viral strains, poor patient adherence or intolerance/adverse drug reactions. Virological failure, i.e. two consecutive viral loads of greater than 400 copies/mL in a previously fully suppressed patient requires investigation. Viral genotyping should be used to help select future therapy. If a new suitable treatment option is available it should be started as soon as possible.

Treatment failure in highly treatment-experienced patients poses considerable challenges, but new classes of antiretroviral agents with activity against drug-resistant strains of HIV, makes long-term virological suppression a realistic objective, even in heavily pre-treated patients. However, in some situations it may be better to hold back a new drug and await development of another new agent to give the maximum chance of success.

If the patient has a viral load below the limit of detection and a change needs to be made because of intolerance of a particular drug, then a switch to another sensitive drug within the same class should be made. Simplification of complex regimens may be considered if adherence is problematic.

Stopping therapy

Antiretroviral drugs may have to be stopped in, for example, cumulative toxicity, or potential drug interactions with medications needed to deal with another more pressing problem. If adherence is poor, stopping completely may be preferable to continuing with inadequate dosing, in order to reduce the development of viral resistance. Poor quality of life and the view of the patient should be discussed.

NNRTIs efavirenz and nevirapine have long half-lives and therefore should be stopped *before* the other drugs in the mixture to reduce the risk of drug resistance. If this is not possible, lopinavir/ritonavir may be used either in substitution of the NNRTI or as monotherapy for several weeks to cover the period of subtherapeutic levels.

Structured treatment interruptions (STIs) were tried to enhance HIV-specific immune responses or to reduce the exposure to drugs and limit adverse events. However, rapid viral rebound, the adverse effects of viral rebound on the immune system and falling CD4 count in already immunosuppressed patients is a major problem.

New data from a large international trial, the SMART study, on intermittent antiretroviral therapy (even with CD4 counts above 250 cells/mL) suggests these patients do less well than those on continuous therapy. Current British guidance does not support treatment interruption as a standard of care.

Complications of antiretroviral therapy

Side-effects are a common problem in HAART (Table 4.53). Some are acute and associated with initiation of medication, whilst others emerge after longer-term exposure to drugs.

Allergic reactions

Allergic reactions occur with greater frequency in HIV infection and have been documented with all the antiretroviral drugs. Abacavir is associated with a potentially fatal hypersensitivity reaction, strongly associated with the presence of HLAB*5701, usually within the first 6 weeks of treatment. There may be a discrete rash and often a fever coupled with general malaise and gastrointestinal and respiratory symptoms. The diagnosis is clinical and symptoms resolve when abacavir is withdrawn. Re-challenge with abacavir can be fatal and is contraindicated. In the UK routine screening for the HLAB*5701 allele has reduced the incidence of abacavir hypersensitivity. Allergies to NNRTIs (often in the second or third week of treatment) usually present with a widespread maculopapular pruritic rash, often with a fever and disordered liver biochemical tests. Reactions can resolve even with continuing therapy but drugs should be stopped immediately in any patient with mucous membrane involvement or severe hepatic dysfunction.

Lipodystrophy and metabolic syndrome

A syndrome of lipodystrophy occurs in patients with HIV on HAART comprising characteristic morphological changes and metabolic abnormalities. The main characteristics include a loss of subcutaneous fat in the arms, legs and face (lipoatrophy), deposition of visceral, breast and local fat, raised total cholesterol, HDL cholesterol and triglycerides, and insulin resistance with hyperglycaemia. The syndrome is potentially associated with increased cardiovascular morbidity. The aetiology is unclear although PIs and NRTIs have been implicated. The highest incidence occurs in those taking combinations of NRTIs and PIs. Stavudine and zidovudine are associated with the lipoatrophy component of the process. The syndrome is difficult to treat, and switching or stopping therapy may not always reverse the processes, although switching away from a stavudine and/or PI-based therapy is recommended. Dietary advice and increasing exercise may improve some of the metabolic problems and help body shape. The addition of glitazones, metformin and growth hormone have been tried. Statins and fibrates are recommended to reduce circulating lipids. Pravastatin has fewer interactions with protease inhibitors but simvastatin is contraindicated as it has high levels of drug interactions with PIs. Surgical approaches are useful in improving facial appearance following lipoatrophy. Polylactic acid injections (New-Fill) have provided improvement in some patients.

Mitochondrial toxicity and lactic acidosis

Mitochondrial toxicity, mostly involving the nucleoside analogue class, leads to raised lactate and lactic acidosis, which has in some cases been fatal. NRTIs inhibit gamma DNA polymerase, and other enzymes that are necessary for normal mitochondrial function. Symptoms are often vague and insidious and include anorexia, nausea, abdominal pain and general malaise. Venous lactate is raised, and the anion gap is typically widened. This is a serious condition requiring immediate cessation of antiretroviral therapy and provision of appropriate supportive measures until normal biochemistry is restored. All patients should be alerted to possible symptoms and encouraged to attend hospital promptly.

Bone metabolism

A variety of bone disorders have been reported in HIV, in particular osteopenia, osteoporosis and avascular necrosis. The prevalence of these conditions has varied widely in different studies. Antiretroviral agents, particularly PIs, have been implicated in the aetiology, although untreated HIV is believed to have a direct impact on bone metabolism.

IRIS

Paradoxical inflammatory reactions (immune reconstitution inflammatory syndrome, IRIS) may occur on initiating HAART. This occurs usually in people who have been profoundly immunosuppressed and begin therapy. As their immune system recovers, they are able to mount an inflammatory response to a range of pathogens, which can include exacerbation of symptoms with new or worsening of clinical signs. Examples include unusual mass lesions or lymphadenopathy associated with mycobacteria, including deteriorating radiological appearances associated with TB infection. Inflammatory retinal lesions in association with cytomegalovirus, deterioration in liver function in chronic hepatitis B carriers and vigorous vesicular eruptions with herpes zoster have also been described.

Specific therapeutic situations

Acute seroconversion

Antiretroviral therapy in patients presenting with an acute seroconversion illness is controversial. This stage of disease may represent a unique opportunity for therapy as there is less viral diversity, and the host immune capacity is still intact. There is evidence to show that the viral load can be reduced substantially by aggressive therapy at this stage, although it rises when treatment is withdrawn. The longer-term clinical sequelae of treatment at this stage remain uncertain. People with severe symptoms during primary HIV infection may gain a clinical improvement on antiretrovirals. If treatment is contemplated in this situation, entry into a clinical trial is sensible.

Pregnancy

Management of HIV-infected pregnant women requires close collaboration between obstetric, medical and paediatric teams. The aim of HIV management is to deliver a healthy, uninfected baby to a healthy mother without prejudicing the future treatment options of the mother. Although considerations of pregnancy must be factored into clinical decision-making, pregnancy per se should not be a contraindication to providing optimum HIV-related care for the woman. HIV-positive women should be advised against breast-feeding, which doubles the risk of vertical transmission. Delivery by caesarean section reduces the risk of transmission but if the patient is on HAART is unnecessary. For women eligible for treatment of their own HIV disease, whether pregnant or not, triple therapy is the regimen of choice. Risk of vertical transmission increases with viral load. Although the fetus will be exposed to more drugs, the chances of reducing the viral load and hence preventing infection are greatest with a potent triple therapy regimen in the mother.

Treatment should start from 12–14 weeks of pregnancy and continue during delivery. Zidovudine is the usual choice with the baby receiving zidovudine for 6 weeks postpartum. The woman should remain on treatment with appropriate monitoring and support.

Women who do not need treatment for themselves should be prescribed a short course of antiretroviral therapy initiated

between 22 and 28 weeks of pregnancy to reduce vertical transmission. Treatment of the mother with zidovudine monotherapy is at variance with the data on combination therapy for adults as described above, but remains a viable option for a selected group of patients (see BHIVA pregnancy guidelines). Efavirenz has been associated with developmental abnormalities in primate models and several retrospective cases of neural tube defects in babies born to women taking efavirenz have been reported. Efavirenz should not be initiated in pregnancy and women who have conceived on the drug should discuss the risks and benefits of continuing or switching with an expert clinician.

Details of adverse effects associated with ARVs in pregnancy is maintained by the antiretroviral pregnancy registry which holds prospective international data and is regularly updated (www.apregistry.com).

Post-exposure prophylaxis

Healthcare workers following occupational exposure to HIV may need antiretroviral therapy. The British recommendation is zidovudine, lamivudine and lopinavir/ritonavir for 4 weeks. Prophylaxis after sexual exposure may be appropriate in certain situations, in particular rape or in HIV-discordant relationships. The choice of drugs will depend on the clinical setting and what is known of the HIV source.

Box 4.23 Use of vaccines in HIV-infected adults

Vaccines that can be used in all HIV-infected adults (all inactivated)	Vaccines that are contraindicated in all HIV-infected adults	Vaccines that can be used only in asymptomatic HIV-infected adults with a current CD4 count >200 cells/mm^3
Anthrax	Cholera – CVD103-HgR (live)	Measles, mumps, rubella (MMR)
Cholera – WC/rBS	Influenza – intranasal (live)	Varicella
Hepatitis A	Poliomyelitis – oral (OPV) (live)	Yellow fever
Hepatitis B	Typhoid – Ty21a (live)	
Haemophilus influenzae b (Hib)	Tuberculosis (BCG) (live)	
Influenza – parenteral	Smallpox (vaccinia) (live)	
Japanese encephalitis		
Meningococcus – MenC		
Meningococcus – ACWY		
Pneumococcus – PPV23		
Poliomyelitis – parenteral (IPV)		
Rabies		
Tetanus-Diphtheria (Td)		
Tick-borne encephalitis		
Typhoid – ViCPS		

Based on BHIVA guidelines 2006, www.bhiva.org.

Opportunistic infections in the HAART era

Since the introduction of HAART in 1996, the USA and Europe have witnessed a sustained fall in opportunistic infection (OI) and mortality. In the USA deaths that included AIDS-related causes decreased from 3.79/100 person-years in 1996 to 0.32/100 person-years in 2004. Death rates across Europe fell fivefold between 1995 and 1998. European data showed a fall in AIDS-defining illnesses (ADI) from 30.7 per 100 patient-years of observation to 2.5 per 100 patient-years between 1994 and 1998, with patients on HAART having a lower rate of ADIs than patients not on HAART. This trend has continued – recent data from Europe demonstrate a 50% lower incidence of AIDS in 1998–2001 than in 1996–1998, irrespective of CD4 count. Potential long-term adverse effects associated with HAART have not altered its effectiveness in treating AIDS. However the decline has been steeper for some OIs than others, suggesting that the immune reconstitution due to HAART may not be functionally equal against all HIV-associated complications. Not all patients have adequate responses to these medications, even if they are available and tolerable. Immune reconstitution with HAART may produce unusual responses to opportunistic pathogens and confuse the clinical picture. Thus prevention and treatment of OIs remains an integral part of the management of HIV infection.

Prevention of opportunistic infection in HIV-infected patients

Avoid infection

Exposure to certain organisms can be avoided in those known to be HIV infected. Attention to food hygiene will reduce exposure to salmonella, toxoplasmosis and *Cryptosporidium*, and protected sexual intercourse will reduce exposure to herpes simplex virus (HSV), hepatitis B and papillomaviruses. Cytomegalovirus (CMV)-negative patients should be given CMV-negative blood products.

Immunization strategies (Box 4.23)

Guidance on the appropriate use of vaccines in HIV is available from www.bhiva.org/files.

Immunization may not be as effective in HIV-infected individuals.

Hepatitis A and B vaccines should be given for those without natural immunity who are at risk, particularly if there is coexisting liver pathology, e.g. hepatitis C.

Chemoprophylaxis

In the absence of a normal immune response, many OIs are hard to eradicate using antimicrobials, and the recurrence rate is high. Primary and secondary chemoprophylaxis has reduced the incidence of many OIs. Advantages must be balanced against the potential for toxicity, drug interactions and cost, with each medication added to what are often complex drug regimens.

- Primary prophylaxis is effective in reducing the risk of *Pneumocystis jiroveci*, toxoplasmosis and *Mycobacterium avium-intracellulare*.
- Primary prophylaxis is **not** normally recommended against cytomegalovirus, herpesviruses or fungi.

With the introduction of HAART and immune reconstitution, ongoing chemoprophylaxis can be discontinued in those patients with CD4 counts that remain consistently above 200 and who have a low viral load. In developing countries unable to obtain HAART, long-term secondary prophylaxis is still advocated. Other less severe but recurrent infections may also warrant prophylaxis (e.g. herpes simplex, candidiasis).

SPECIFIC CONDITIONS ASSOCIATED WITH HIV INFECTION

Fungal infections

Pneumocystis jiroveci (formerly P. carinii) (see p. 861)

This organism most commonly causes pneumonia (PCP) but can cause disseminated infection. It is not usually seen until patients are severely immunocompromised with a CD4 count below 200. The infection remains common although the use of primary prophylaxis in patients with CD4 < 200 has reduced the incidence. The organism damages alveolar epithelium, which impedes gas exchange and reduces lung compliance.

The onset is often insidious over a period of weeks, with a prolonged period of increasing shortness of breath (usually on exertion), non-productive cough, fever and malaise. *Clinical examination* reveals tachypnoea, tachycardia, cyanosis and signs of hypoxia. Fine crackles are heard on auscultation, although in mild cases there may be no auscultatory abnormality. In early infection the chest X-ray is normal but the typical appearances are of bilateral perihilar interstitial infiltrates, which can progress to confluent alveolar shadows throughout the lungs. High-resolution CT scans of the chest demonstrate a characteristic ground-glass appearance even when there is little to see on the chest X-ray. The patient is usually hypoxic and desaturates on exercise. Definitive diagnosis rests on demonstrating the organisms in the lungs via bronchoalveolar lavage or by PCR amplification of the fungal DNA from a peripheral blood sample. As the organism cannot be cultured in vitro it must be directly observed either with silver staining or immunofluorescent techniques.

Treatment should be instituted as early as possible. First-line therapy is with intravenous co-trimoxazole (120 mg/kg daily in divided doses) for 21 days. Up to 40% of patients receiving this regimen will develop some adverse drug reaction, including a typical allergic rash. If the patient is sensitive to co-trimoxazole, intravenous pentamidine (4 mg/kg per day) or dapsone and trimethoprim are given for the same duration. Atovaquone or a combination of clindamycin and primaquine is also used. In severe cases (P_aO_2 less than 9.5 kPa), systemic corticosteroids have been shown to reduce mortality and should be added. Continuous positive airways pressure (CPAP) or mechanical ventilation (see p. 915) is required if the patient remains severely hypoxic or becomes too tired. Pneumothorax not uncommonly complicates the clinical course in an already severely hypoxic patient.

Secondary prophylaxis is required in patients whose CD4 count remains below 200, to prevent relapse, the usual regimen being co-trimoxazole 960 mg three times a week. Patients sensitive to sulphonamide are given dapsone, pyrimethamine or nebulized pentamidine. The last only protects the lungs and does not penetrate the upper lobes particularly efficiently; hence if relapses occur on this regimen they may be either atypical or extrapulmonary.

Cryptococcus (see p. 151)

The most common presentation of cryptococcus infection in the context of HIV is meningitis, although pulmonary and disseminated infections can also occur. The organism, *C. neoformans*, is widely distributed – often in bird droppings – and is usually acquired by inhalation. The onset may be insidious with non-specific fever, nausea and headache. As the infection progresses the conscious level is impaired and changes in affect may be noted. Fits or focal neurological presentations are uncommon. Neck stiffness and photophobia may be absent as these signs depend on the inflammatory response of the host, which in this setting is abnormal.

Diagnosis is made on examination of the CSF (perform CT scan before lumbar puncture to exclude space-occupying pathology). Indian ink staining shows the organisms directly and CSF cryptococcal antigen is positive at variable titre. It is unusual for the cryptococcal antigen to become negative after treatment, although the levels should fall substantially. Cryptococci can also be cultured from CSF and/or blood.

Factors associated with a poor prognosis include a high organism count in the CSF, a low white cell count in the CSF, and an impaired consciousness level at presentation.

Treatment. Initial treatment is usually with intravenous amphotericin B (0.7 mg/kg per day) ± flucytosine as induction, although intravenous fluconazole (400 mg daily) is useful if renal function is impaired or if amphotericin side-effects are troublesome. Oral fluconazole 200 mg daily can be substituted if the organism is shown to be fully sensitive and the patient is responding, and should be continued lifelong unless the CD4 count is >100 for a 6-month period.

The mortality from a first episode of cryptococcal meningitis is up to 20%. Relapse is very common, possibly from a prostatic reservoir in men.

Candida (see p. 149)

Mucosal infection with *Candida* is very common in HIV-infected patients. Oral *Candida* is one of the most common conditions. *C. albicans* is the usual organism, although *C. krusei* and *C. glabrata* occur. Pseudomembranous candidiasis consisting of creamy plaques in the mouth and pharynx is easily recognized but *erythematous Candida* appears as reddened areas on the hard palate or as atypical areas on the tongue. Angular cheilitis can occur in association with either form. Vulvovaginal *Candida* is often problematic.

Oesophageal Candida infection produces odynophagia (see p. 255). Fluconazole or itraconazole are the agents of choice. With prolonged exposure to these agents in HIV-infected patients, azole-resistant *C. albicans* is becoming an increasing problem. Switching azoles may produce a response. *Disseminated Candida* is uncommon in the context of HIV infection but if present, fluconazole is the preferred drug with amphotericin, voriconazole or capsofungin also used. Itraconazole was formerly used. *C. krusei* may colonize patients who have been treated with fluconazole, as it is fluconazole-resistant. Amphotericin is useful in the treatment of this infection and an attempt to type *Candida* from clinically azole-resistant patients should be made.

Aspergillus (see p. 151)

Infection with *Aspergillus fumigatus* occurs in advanced HIV disease. Patients with long-standing neutropenia (often due to chemotherapy) and those on ganciclovir therapy for CMV and myelotoxic antiretrovirals, are prone to this infection. Spores are airborne and ubiquitous. Following inhalation, lung infection proceeds to haematogenous spread to other organs. Sinus infection occurs.

The prognosis is very poor, with amphotericin B or voriconazole being the mainstay of therapy. Capsofungin is also effective. It is almost impossible to treat *Aspergillus* unless the neutrophil count can be sustained, often with granulocyte colony-stimulating factors.

Histoplasmosis (see p. 150)

This infection is a well-recognized complication of HIV in the USA where it is endemic in soil. The most common manifestation is with pneumonia, which may be confused with *Pneumocystis jiroveci* in its presentation (see above).

Mucocutaneous dermatophyte infections

(Table 4.51)

These infections are common (see p. 1232).

Protozoal infections

Toxoplasmosis (see p. 160)

Toxoplasma gondii most commonly causes encephalitis and cerebral abscess in the context of HIV, usually as a result of reactivation of previously acquired infection. The incidence depends on the rate of seropositivity to toxoplasmosis in the particular population. High antibody levels are found in France (90% of the adult population). About 25% of the adult UK population is seropositive to toxoplasma.

Clinical presentation is of a focal neurological lesion with convulsions, fever, headache and possible confusion. Examination reveals focal neurological signs in more than 50% of cases. Eye involvement with chorioretinitis may also be present. In most but not all cases of *Toxoplasma* serology is positive. Typically CT scan of the brain shows multiple ring-enhancing lesions. A single lesion on CT may be found to be one of several on MRI. A solitary lesion on MRI, however, makes a diagnosis of toxoplasmosis unlikely.

Diagnosis is by brain biopsy, but in most cases an empirical trial of anti-toxoplasmosis therapy is instituted and if this leads to radiological improvement within 3 weeks this is considered diagnostic. The differential diagnosis includes cerebral lymphoma, tuberculoma or focal cryptococcal infection.

Treatment is with pyrimethamine for at least 6 weeks (loading dose 200 mg, then 50 mg daily) combined with sulfadiazine and folinic acid. Clindamycin and pyrimethamine may be used in patients allergic to sulphonamide. Anticonvulsants should be given. Lifelong maintenance is required to prevent relapse unless the CD4 count can be restored by HAART.

Cryptosporidiosis (see p. 162)

Cryptosporidium parvum can cause a self-limiting acute diarrhoea in an immunocompetent individual. In HIV infection it can cause severe and progressive watery diarrhoea which may be associated with anorexia, abdominal pain, nausea and vomiting. Cysts attach to the epithelium of the small bowel wall, causing secretion of fluid into the gut lumen and failure of fluid absorption. It is also associated with sclerosing cholangitis (see p. 372). The cysts are seen on stool specimen microscopy using Kinyoun acid-fast stain and are readily identified in small bowel biopsy specimens.

HAART is associated with resolution of symptoms, otherwise treatment is largely supportive. There are no effective antimicrobial agents available other than paromamycin, which may have a limited effect on diarrhoea.

Microsporidiosis

Enterocytozoon bieneusi and *Septata intestinalis* are a cause of diarrhoea. Spores can be detected in stools using a trichrome or fluorescent stain that attaches to the chitin of the spore surface. HAART and immune restoration can have a dramatic effect and albendazole eradicates the infection.

Leishmaniasis (see p. 158)

This occurs in immunosuppressed HIV-infected individuals who have been in endemic areas, which include South America, tropical Africa and much of the Mediterranean. Symptoms are frequently non-specific with fever, malaise, diarrhoea and weight loss. Splenomegaly, anaemia and thrombocytopenia are significant findings. Amastigotes may be seen on bone marrow biopsy or from splenic aspirates. Serological tests exist for *Leishmania* but they are not reliable in this setting.

Treatment is based on sodium stibogluconate (pentavalent antimony) but liposomal amphotericin is the drug of choice if resources permit. Relapse without HAART is common unless long-term secondary prophylaxis is given.

Viral infections

Hepatitis B and C viruses (see pp. 335 and 340)

Because of the comparable routes of transmission of hepatitis viruses and HIV, co-infection is common, particularly in MSM, drug users and those infected by blood products. A higher prevalence of hepatitis viruses is found in those with HIV infection than in the general population. With the striking improvement in prognosis in HIV since the introduction of HAART, morbidity and mortality of hepatitis co-infection has become increasingly significant and may limit life expectancy in HIV. In co-infected patients the hepatotoxicity associated with certain antiretroviral agents may be potentiated. HBV vaccination (and HAV) should be given.

Hepatitis B infection does not appear to influence the natural history of HIV; however, in HIV co-infected patients there is a significantly reduced rate of hepatitis B e antigen (HBeAg) clearance, and the risk of developing chronic infection is increased. HBV reactivation and reinfection is also seen. Liver disease occurs most commonly in those with high HBV DNA levels indicative of continuing replication.

Treatment options for HBV are limited. The ideal time to start treatment is not clear; however, response rates for co-infected patients may be better at higher CD4 counts. Treatments for HBV include some agents with concomitant anti-HIV activity, including lamivudine, tenofovir and emtricitabine. These must be used within an effective anti-HIV regimen.

Hepatitis C is associated with more rapid progression of HIV infection, and the CD4 responses to HAART in co-infected patients may be blunted. Hepatitis C progression is both more likely and more rapid in the presence of HIV infection, and the hepatitis C viral load tends to be elevated. The hepatotoxicity associated with HAART is worse in those with HCV co-infection.

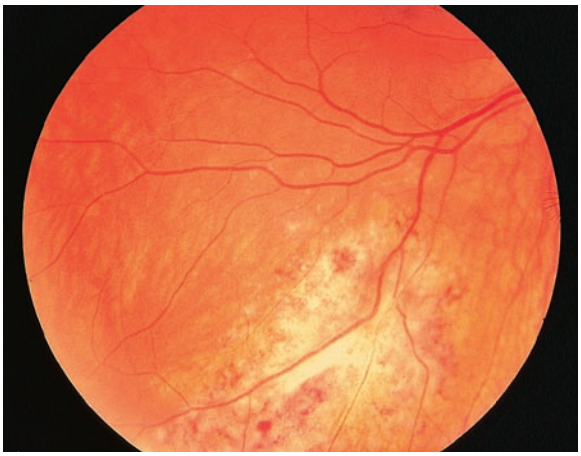


Fig. 4.46 **Untreated CMV retinitis.**

Assessment of co-infected patients requires full clinical and laboratory evaluation and staging of both infections. For HCV both viral load and genotype will influence therapeutic decision-making. In HBV infection, detection and quantification of HBV DNA acts as a marker of viral activity, whilst the significance of viral genotype is still uncertain. Liver biopsy is useful to obtain a histological staging of disease.

Treatment options in HCV co-infection are similar for those infected with HCV alone, depending on stage of disease and HCV genotype. Pegylated interferon has greater efficacy than standard interferon and is combined with ribavirin. In HIV/HCV co-infected patients those with a CD4 count above 200 have a better chance of success. In general it is preferable to treat HCV first if HIV infection is stable, as this minimizes hepatotoxicity associated with HAART. However, if the CD4 count is low or the patient is at risk of HIV progression, HIV therapy should be instituted first.

Cytomegalovirus (see p. 107)

CMV can be a cause of considerable morbidity in HIV-infected individuals, especially in the later stages of disease. The major problems encountered are retinitis, colitis, oesophageal ulceration, encephalitis and pneumonitis. CMV infection is associated with an arteritis, which may be the major pathogenic mechanism. CMV also causes polyradiculopathy and adrenalitis.

CMV retinitis

This tends to occur once the CD4 count is below 100 and has been found in up to 30% of AIDS cases. It is the most common cause of eye disease and blindness. Although usually unilateral to begin with, the infection frequently progresses to involve both eyes. Presenting features depend on the area of retina involved (loss of vision being most common with macular involvement) and include floaters, loss of visual acuity, field loss and scotomata, orbital pain and headache.

Examination of the fundus (Fig. 4.46) reveals haemorrhages and exudates, which follow the vasculature of the retina (so called 'pizza pie' appearances). The features are highly characteristic and the diagnosis is made clinically. Retinal detachment and papillitis may occasionally occur. If untreated, retinitis spreads within the eye, destroying the retina within its path. Routine fundoscopy should be carried out on all HIV-infected patients to look for evidence of early infection. Any patient with symptoms of visual disturbance should have a thorough examination with pupils dilated, and if no evident pathology is seen a specialist ophthalmologic opinion should be sought.

Treatment for CMV should be started as soon as possible with either ganciclovir (10 mg/kg daily) or foscarnet (60 mg/kg 8-hourly) given intravenously for at least 3 weeks, or until retinitis is quiescent. Reactivation is common, leading to blindness. The major side-effect of ganciclovir is myelosuppression, and foscarnet is nephrotoxic. Maintenance therapy may be required until HAART is instituted and has improved immune competence. Valganciclovir, an oral prodrug of ganciclovir, is available which has some long-term benefit when used as maintenance therapy, but has a lower efficacy than intravenous ganciclovir. Ganciclovir can be given directly into the vitreous cavity but regular injections are required. A sustained-release implant of ganciclovir can be surgically inserted into the affected eye. Cidofovir is available for use when the above drugs are contraindicated. It has renal toxicity.

CMV gastrointestinal conditions

CMV colitis usually presents with abdominal pain, often generalized or in the left iliac fossa, diarrhoea which may be bloody, generalized abdominal tenderness and a low-grade fever. Dilated large bowel may be seen on abdominal X-ray. Sigmoidoscopy shows a friable or ulcerated mucosa; histology shows the characteristic 'owls eye' cytoplasmic inclusion bodies (Fig. 4.18).

Treatment with either intravenous ganciclovir or foscarnet for 3 weeks improves symptoms and the histological changes are reversed. However, relapse is common once therapy is stopped, unless HAART is available. The issue of long-term maintenance therapy in this situation is controversial since, unlike CMV retinitis, repeated attacks of colitis do not necessarily have significant long-term sequelae. Given the complications associated with maintenance anti-CMV therapy there may be no overall benefit.

Other sites along the gastrointestinal tract are also prone to CMV infection, e.g. ulceration of the oesophagus, usually in the lower third, causes odynophagia. CMV can also cause hepatitis.

CMV neurological conditions

CMV polyradiculopathy usually affects the lumbosacral roots, leading to muscle weakness and sphincter disturbance. The CSF has an increase in white cells, which surprisingly are almost all neutrophils. Although progression may be arrested by anti-CMV medication, functional recovery may not occur.

CMV encephalopathy has clinical similarities to that caused by HIV itself. It tends to respond poorly to therapy.

Herpesviruses (see pp. 104 and 1230)

Herpes simplex primary infection occurs with greater frequency and severity, presenting in an ulcerative rather than vesicle form in profoundly immunosuppressed individuals. Genital, oral and occasionally disseminated infection is seen. Viral shedding may be prolonged in comparison with immunocompetent patients.

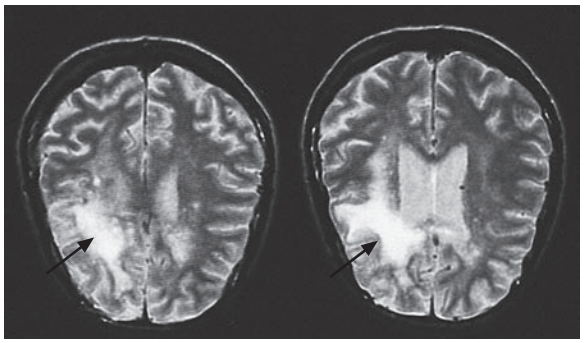


Fig. 4.47 **MRI scans showing progressive multifocal leucoencephalopathy.**

Varicella zoster virus

Varicella zoster can occur at any stage of HIV but tends to be more aggressive and longer-lasting in the more immunosuppressed patient. Multidermatomal zoster may occur.

Therapy with aciclovir is usually effective. Frequent recurrences need suppressive therapy. Aciclovir-resistant strains (usually due to thymidine kinase-deficient mutants) in HIV-infected patients have become more common. Such strains may respond to foscarnet.

Herpesvirus 8 (HHV-8) is the causative agent of Kaposi's sarcoma (see p. 108).

Epstein–Barr virus (see p. 107)

Patients with HIV have been shown to have high levels of EBV colonization. There are increased EBV titres in oropharyngeal secretions and high levels of EBV-infected B cells. The normal T cell response to EBV is depressed in HIV. EBV is strongly associated with primary cerebral lymphoma and non-Hodgkin's lymphoma (see below). Oral hairy leucoplakia caused by EBV is a sign of immunosuppression first noted in HIV but now also recognized in other conditions. It appears intermittently on the lateral borders of the tongue or the buccal mucosa as a pale ridged lesion. Although usually asymptomatic, patients may find it unsightly and occasionally painful. The virus can be identified histologically and on electron microscopy. There is a variable response to aciclovir.

Human papillomavirus (see p. 108)

HPV produces genital, plantar and occasionally oral warts, which may be slow to respond to therapy and recur repeatedly. HPV is associated with the more rapid development of cervical and anal intraepithelial neoplasia, which in time may progress to squamous cell carcinoma of the cervix or rectum in HIV-infected individuals. HPV vaccination is now available (see p. 108).

Polyomavirus (see p. 108)

JC virus, a member of the papovavirus family, which infects oligodendrocytes, causes progressive multifocal leucoencephalopathy (PML). This leads to demyelination particularly within the white matter of the brain. The features are of progressive neurological and/or intellectual impairment, often including hemiparesis or aphasia. The course is usually inexorably progressive but a stuttering course may be seen. Radiologically the lesions are usually multiple and confined to the white matter. They do not enhance with contrast and do not produce a mass effect. MRI (Fig. 4.47) is more sensitive than CT and reveals enhanced signal on T2-weighted images of the lesions. Definitive diagnosis is made on histological and viral examination of brain tissue obtained at biopsy. There is no specific therapy. HAART which enhances the immune response has, however, produced both clinical and radiological remission in a number of cases. Addition of cidofovir to HAART may improve outcome in some patients.

Bacterial infections

Bacterial infection in HIV is common. Cell-mediated immune responses normally control infection against intracellular bacteria, e.g. *Mycobacteria*. The abnormalities of B cell function associated with HIV lead to infections with encapsulated bacteria, as reduced production of IgG2 cannot protect against the polysaccharide coat of such organisms. These functional abnormalities may be present well before there is a significant decline in CD4 numbers and so bacterial sepsis may be seen at early stages of HIV infection. *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Moraxella catarrhalis* infections are examples. Bacterial infection is often disseminated and, although usually amenable to standard antibiotic therapy, may reoccur. Long-term prophylaxis is required if recurrent infection is frequent.

Mycobacteria

Mycobacterium tuberculosis (see p. 863)

Many parts of the world with a high prevalence of TB also have high rates of HIV infection, e.g. Africa, both of which are increasing. The respiratory transmission of TB means both HIV-positive and negative people are being infected. TB can cause disease when there is only minimal immunosuppression and thus often appears early in the course of HIV infection. HIV-related TB frequently represents reactivation of latent TB, but there is also clear evidence of newly acquired infection and nosocomial spread in HIV-infected populations.

The pattern of disease differs with immunosuppression.

- Patients with relatively well preserved CD4 counts have a clinical picture similar to that seen in HIV-negative patients with pulmonary infection.
- In more advanced HIV disease, atypical pulmonary presentations without cavitation and prominent hilar lymphadenopathy, or extrapulmonary TB affecting lymph nodes, bone marrow or liver occur. Bacteraemia may be present.

Diagnosis depends on demonstrating the organisms in appropriate tissue specimens. The response to tuberculin testing is blunted in HIV-positive individuals and is unreliable. Sputum microscopy may be negative even in pulmonary infection and culture techniques are the best diagnostic tool.

Treatment. M. tuberculosis infection usually responds well to standard treatment regimens, although the duration of therapy may be extended, especially in extrapulmonary infection. Multidrug resistance and extensive drug resistant TB (see p. 866) is becoming a problem, particularly in the USA where it is becoming a nosocomial danger. Cases from HIV units in the UK have been reported. Compliance with antituberculous therapy needs to be emphasized. Treatment of TB (see below) is not curative and long-term isoniazid prophylaxis may be given. In patients from TB-endemic areas, primary prophylaxis may prevent emergence of infection.

Drug–drug interactions between antiretroviral and antituberculous medications are complex and are a consequence of enzyme induction or inhibition. Rifampicin is a potent inducer of cytochrome P450, which is also the route for metabolism of HIV protease inhibitors. Using both drugs together results in a reduction in circulating protease inhibitor with reduced efficacy and increased potential for drug resistance. Some protease inhibitors themselves block cytochrome P450, which leads to potentially toxic levels of rifampicin and problems such as uveitis and hepatotoxicity. The non-nucleoside reverse transcriptase inhibitor class also interacts variably with rifamycins, requiring dose alterations. Additionally, there are overlapping toxicities between HAART regimens and antituberculous drugs, in particular hepatotoxicity, peripheral neuropathy and gastrointestinal side-effects. Rifabutin has a weaker effect on cytochrome P450 and may be substituted for rifampicin. Dose adjustments must be made for drugs used in this situation to take account of these interactions.

Paradoxical inflammatory reactions (immune reconstitution inflammatory syndrome, IRIS) which can include exacerbation of symptoms, new or worsening clinical signs and deteriorating radiological appearances have been associated with the improvement of immune function seen in HIV-infected patients starting HAART in the face of *M. tuberculosis* infection. They are most commonly seen in the first few weeks after initiation of HAART in patients recovering from TB and can last several weeks or months. The syndrome does not reflect inadequate TB therapy and is not confined to any particular combination of antiretroviral agents. It is vital to exclude new pathology in this situation. However delaying antiretroviral therapy increases the risks of further opportunistic events. Allowing at least two weeks of antituberculous therapy before commencing HAART allows some reduction in the burden of mycobacteria. If the CD4 count is less than 100 then antiretrovirals should be started at about 2 weeks of anti-TB medication, If the CD4 count is above 200 then initiation of HAART may wait for at least 6 weeks after the start of antituberculous therapy.

Mycobacterium avium-intracellulare

Atypical mycobacteria, particularly *M. avium-intracellulare (MAI)*, generally appear only in the later stages of HIV infection when patients are profoundly immunosuppressed. It is a saprophytic organism of low pathogenicity that is ubiquitous in soil and water. Entry may be via the gastrointestinal tract or lungs with dissemination via infected macrophages.

The major *clinical features* are fevers, malaise, weight loss, anorexia and sweats. Dissemination to the bone marrow causes anaemia. Gastrointestinal symptoms may be prominent with diarrhoea and malabsorption. At this stage of disease patients frequently have other concurrent infections, so differentiating MAI is difficult on clinical grounds. Direct examination and culture of blood, lymph node, bone marrow or liver give the diagnosis most reliably.

Treatment. MAI is typically resistant to standard antituberculous therapies, although ethambutol may be useful. Drugs such as rifabutin in combination with clarithromycin or azithromycin reduce the burden of organisms and in some ameliorate symptoms. A common combination is ethambutol, rifabutin and clarithromycin. Addition of amikacin to a drug regimen may produce a good symptomatic response. Primary prophylaxis with rifabutin or azithromycin may delay the appearance of MAI, but no corresponding increase in survival has been shown.

Infections due to other organisms

Salmonellae (non-typhoidal) (see p. 129) are frequent pathogens in HIV infection. Salmonellae are able to survive within macrophages, this being a major factor in their pathogenicity. Organisms are usually acquired orally and frequently result in disseminated infection. Gastrointestinal disturbance may be disproportionate to the degree of dissemination, and once the pathogen is in the bloodstream any organ may be infected. Salmonella osteomyelitis and cystitis have been reported. Diagnosis is from blood and stool cultures.

Response to standard antibiotic therapy, depending on laboratory sensitivities, is usually good. Recurrent infection is, however, common and long-term prophylaxis may be required.

Education on food hygiene should be provided.

Skin conditions such as folliculitis, abscesses and cellulitis are common and are usually caused by *Staph. aureus*. *Periodontal disease*, which may be necrotizing, causes pain and damage to the gums. It is more common in smokers, but no specific causative agent has been identified. Therapy is with local debridement and systemic antibiotics.

Strongyloides (see p. 167), a nematode found in tropical areas, may produce a hyperinfection syndrome in HIV-infected patients. Larvae are produced which invade through the bowel wall and migrate to the lung and occasionally to the brain. Albendazole or ivermectin may be used to control infection. Gram-negative septicaemia can develop (see p. 899).

Scabies (see p. 1234) may be much more severe in HIV infection. It may be widely disseminated over the body and appear as atypical, crusted papular lesions known as ‘Norwegian scabies’, from which mites are readily demonstrated. Superadded staphylococcal infection may occur. Treatment with conventional agents such as lindane may fail, and ivermectin has been used to good effect in some patients.

Neoplasms

The mortality and morbidity associated with neoplasia in HIV is substantial, non-Hodgkin’s lymphoma being the most significant tumour now, ahead of Kaposi’s sarcoma.

Kaposi’s sarcoma (see p. 1260)

Kaposi’s sarcoma (KS) in association with HIV (epidemic KS) behaves more aggressively than that associated with HIV-negative populations (endemic KS). The incidence has fallen significantly since the introduction of HAART. The tumour is most common in men who have sex with men and others who have acquired HIV sexually, particularly from a partner who has KS, implicating a sexually transmitted cofactor in the pathogenesis. Human herpesvirus 8 (HHV-8) is involved in pathogenesis. KS skin lesions are characteristically pigmented, well circumscribed and occur in multiple sites. It is a multicentric tumour consisting of spindle cells and vascular endothelial cells, which together form slit-like spaces in which red blood cells become trapped. This process is responsible for the characteristic purple hue of the tumour. In addition to the skin lesions, KS affects lymphatics and lymph nodes, the lung and gastrointestinal tract, giving rise to a wide range of symptoms and signs. Most patients with visceral involvement also have skin or mucous membrane

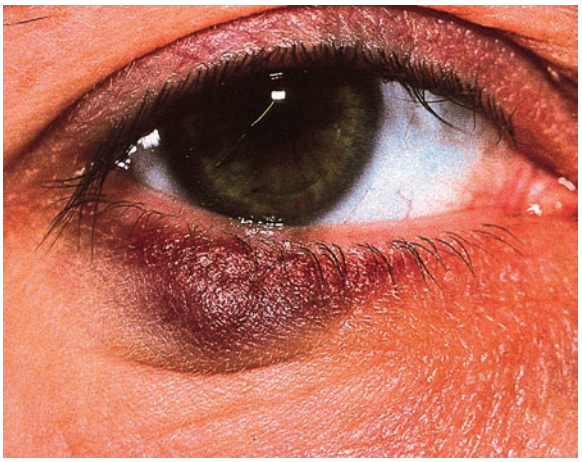


Fig. 4.48 **Kaposi's sarcoma of the eyelid.**

lesions. Visceral KS carries a worse prognosis than that confined to the skin. Kaposi's sarcoma is seen around the eye (Fig. 4.48), particularly in the conjunctivae, which can lead to periorbital oedema.

Initiation of HAART may cause regression of lesions and prevent new ones emerging. Treatment with local radiotherapy gives good results in skin lesions and is helpful in lymph node disease. For patients with aggressive disease, systemic chemotherapy is indicated using combinations of vincristine and bleomycin or the newer liposomal preparations of doxorubicin. Response is often very good although of uncertain duration. Alpha-interferon is also effective.

Lymphoma

A significant proportion of patients with HIV develop lymphoma, mostly of the non-Hodgkin's, large B cell type. These are frequently extranodal, often affecting the brain, lung and gastrointestinal tract. Many of these tumours are strongly associated with Epstein–Barr virus (EBV), with evidence of expression of latent gene nuclear antigens such as EBNA 1–6, some of which are involved in the immortalization of B cells and drive a neoplastic pathway.

HIV-associated lymphomas are frequently very aggressive. Patients often present with systemic 'B' syndromes and progress rapidly despite chemotherapy. Primary cerebral lymphoma is variably responsive to radiotherapy but overall carries a poor prognosis. Lymphomas occurring early in the course of HIV infection tend to respond better to therapy and carry a better prognosis, occasionally going into complete remission.

Squamous cell carcinoma

Squamous cell carcinoma, especially of the cervix and anus, is associated with HIV. Human papillomavirus infection has a role in the pathogenesis of these malignancies. Women with HIV infection should have yearly cervical cytology for detection of premalignant change.

PREVENTION AND CONTROL

As the numbers of new infections continue to rise, appropriate prevention interventions are fundamental to the control of the epidemic. Vaccine development has been hampered by the genetic variability of the virus and the complex immune response that is required from the host. Antiretroviral drugs are not a cure and are accessible only to a privileged minority of the world's population. However data from discordant couples demonstrate that as well as a treatment strategy, effective antiretroviral medication is able to reduce onward transmission. Pre-exposure prophylaxis with antiretroviral agents has been considered but clinical trials have so far been challenging to design and implement. Use of a tenofovir-based regimen is being considered in a trial setting. Topical microbicides for intravaginal use are proving very slow to develop.

Strategies that have been shown to be effective include treatment of sexually transmitted infections, consistent use of condoms, use of clean needles and syringes for drug users and antiretroviral drugs to reduce mother-to-child transmission.

Medically performed circumcision has been shown in African studies of HIV-negative heterosexual men to reduce the female to male transmission of HIV by at least 50%. Male circumcision in HIV-positive men however has not been demonstrated to reduce male to female transmission.

Screening of blood products has reduced iatrogenic infection in developed countries but is expensive and not globally available.

Partner notification schemes are helpful but are sensitive and controversial. Availability and accessibility of confidential HIV testing provides an opportunity for individual health education and risk reduction to be discussed.

Understanding and changing behaviour is crucial but notoriously difficult, especially in areas that carry as many taboos as sex, HIV and AIDS. Poverty, social unrest and war all contribute to the spread of HIV. Political will, not always readily available, is required if progress in these areas is to occur.

FURTHER READING

Barnett T, Whiteside A. *AIDS in the Twenty-First Century, Disease and Globalisation*, 2nd edn. London: Palgrave, 2006.

CDC. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in healthcare settings. *Mortality and Morbidity Weekly Report* 2006; **55**: 1–17.

Centers for Disease Control, Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the Use of Antiretroviral Agents in HIV-1-infected Adults and Adolescents. Department of Health and Human Services, 2008. [Updated 2008; cited 12/03/2008.] Available from: http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf.

Chadwick DR, Geretti AM. Immunization of the HIV infected traveller. *AIDS* 2007; **21**: 787–794.

Gazzard B. British HIV Association Guidelines for the Treatment of HIV-infected Adults with Anti-Retroviral Therapy. *HIV Medicine* 2008; **9**: 563. http://www.bhiva.org.

Palella FJ Jr, Baker RK, Moorman AC et al. Mortality in the highly active antiretroviral therapy era: changing causes of death and disease in the HIV outpatient study. *Journal of Acquired Immune Deficiency Syndromes* 2006; **43**(1): 27–34.

de Ruiter A, Mercey D, Anderson J et al. The British HIV Association and Children's HIV Association guidelines for the management of HIV in pregnant women. *HIV Medicine* 2008; **9**: 452–502.

Simoni JM, Pearson CR, Pantalone DW et al. Efficacy of interventions in improving highly active antiretroviral therapy adherence and HIV-1 RNA viral load. A meta-analytic review of randomized controlled trials. *Journal of Acquired Immune Deficiency Syndromes* 2006; **43**(suppl 1): S23–35.

Taylor BS, Carr JK, Salminen MO et al. The challenge of HIV-1 subtype diversity. *New England Journal of Medicine* 2008; **358**: 1590–1603.

The UK Collaborative Group for HIV and STI Surveillance. *Testing Times: HIV and other Sexually Transmitted Infections in the United Kingdom*. London: Health Protection Agency Centre for Infections, November 2007.

UK National Guidelines for HIV testing 2008. http://www.bhiva.org.

Young T, Arens F, Kennedy G et al. Antiretroviral post-exposure prophylaxis (PEP) for occupational HIV exposure. *Cochrane Database of Systematic Reviews* 2007; Jan **21**(1): CD002835 2007(1).

SIGNIFICANT WEBSITES

http://www.aidsinfo.nih.gov/HealthTopics/HealthTopicDetails.aspx?MenuItem=HealthTopics&Search=Off&HealthTopicID=100&ClassID=62
National Institutes of Health: AIDS Info-HIV Treatment Guidelines.

http://bashh.org
British Association for Sexual Health and HIV.

http://hivdb.stanford.edu/
University of Stanford. HIV Drug Resistance Database.

http://www.aidsmap.com/
National AIDS Manual. Aidsmap Information on HIV and AIDS.

http://www.bhiva.org
BHIVA. British HIV Association

http://www.hiv-druginteractions.org/
Liverpool University. HIV Drug Interactions.

http://www.i-base.info/
i-base. HIV i-Base. HIV Treatment Information.

5 Nutrition

General aspects 207
Water and electrolyte balance 207
Dietary requirements 208
Protein-energy malnutrition (PEM) 212
In developed countries 212
In developing countries 215
Vitamins 217
Fat-soluble vitamins 218
Water-soluble vitamins 220
Minerals 225
Nutrition and ageing 226
The ageing process 226
Obesity 228
Nutritional support in the hospital patient 233
Nutritional support in the home patient 236
Food allergy and food intolerance 237
Alcohol 238

GENERAL ASPECTS

In developing countries, lack of food and poor usage of the available food can result in protein–energy malnutrition (PEM); 50 million pre-school African children have PEM. In developed countries, excess food is available and the most common nutritional problem is obesity.

Diet and disease are interrelated in many ways. Excess energy intake, particularly when high in animal (saturated) fat content, is thought to contribute to a number of diseases, including ischaemic heart disease and diabetes. A relationship between food intake and cancer has been found in many epidemiological studies; an excess of energy-rich foods (i.e. fat and sugar containing), often with physical inactivity, plays a role in the development of certain cancers, while diets high in vegetables and fruits reduce the risk of most epithelial cancers. Numerous carcinogens, either intentionally added to food (e.g. nitrates for preserving foods) or accidental contaminants (e.g. moulds producing aflatoxin and fungi), may also be involved in the development of cancer.

The proportion of processed foods eaten may affect the development of disease. A number of processed convenience foods have a high sugar and fat content and therefore may predispose to dental caries and obesity, respectively. They also have a low fibre content, and dietary fibre is possibly necessary in the prevention of a number of diseases (see p. 211). Some epidemiological data suggest that there are long-term effects of undernutrition, low growth rates in utero being associated with high death rates from cardiovascular disease in adult life.

In 1991 the Department of Health published the dietary reference values for food and energy and nutrients for the United Kingdom. Values were based on the available information including data from the Food and Agriculture Organization (FAO–WHO), United Nations University (UNU) expert committee, so that there is broad agreement on the reference values given. In the UK recommended daily amounts (RDAs) are no longer used, but have been replaced by the *reference nutrient intake* (RNI) to provide more help in interpreting dietary surveys.

The RNI is roughly equivalent to the previous RDA, and is sufficient or more than sufficient to meet the nutritional needs of 97.5% of healthy people in a population. Most people's daily requirements are less than this, and so an *estimated average requirement* (EAR) is also given, which will certainly be adequate for most. A lower reference nutrient intake (LRNI) which fails to meet the requirements of 97.5% of the population is also given. The RNI figures quoted in this chapter are for the age group 19–50 years. These represent values for healthy subjects and are not always appropriate for patients with disease.

FURTHER READING

Panel on Dietary Reference Values of the Committee on Medical Aspects of Food Policy. *Dietary Reference for Food Energy and Nutrients for the United Kingdom* (Report 41DOH). London: HMSO. 1991.

WATER AND ELECTROLYTE BALANCE

Water and electrolyte balance is dealt with fully in Chapter 12. About 1 L of water is required in the daily diet to balance insensible losses, but much more is usually drunk, the kidneys being able to excrete large quantities. The daily RNI for sodium is 70 mmol (1.6 g) but daily sodium intake varies in the range 90–440 mmol (2–10 g). These are needlessly high intakes of sodium which are thought by some to play a role in causing hypertension (see p. 798).

DIETARY REQUIREMENTS

Energy

Food is necessary to provide the body with energy (Fig. 5.1). The SI unit of energy is the joule (J), and 1 kJ = 0.239 kcal. The conversion factor of 4.2 kJ, equivalent to 1.00 kcal, is used in clinical nutrition.

Energy balance

Energy balance is the difference between energy intake and energy expenditure. Weight gain or loss is a simple, but accurate, way of indicating differences in energy balance.

Energy requirements

There are two approaches to assessing energy requirements for subjects who are weight stable and close to energy balance:

- assessment of energy intake
- assessment of total energy expenditure.

Energy intake

This can be estimated from dietary surveys and in the past this has been used to decide daily energy requirements.

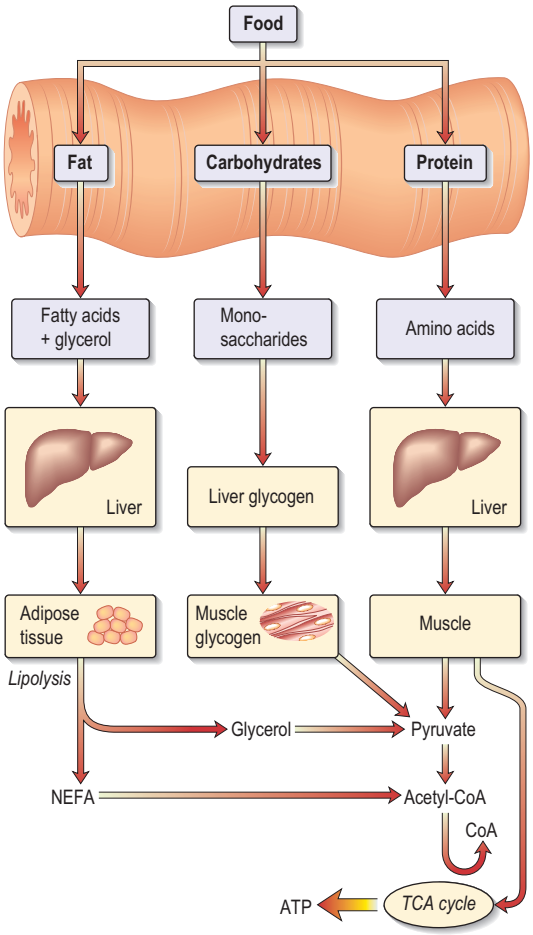


Fig. 5.1 **The production of energy from the main constituents of food.** Alcohol produces up to 5% of total calories, but the variation between individuals is wide. 1 mol of glucose produces 36 mol of ATP. NEFA, non-esterified fatty acids; ATP, adenosine triphosphate; TCA, tricarboxylic acid.

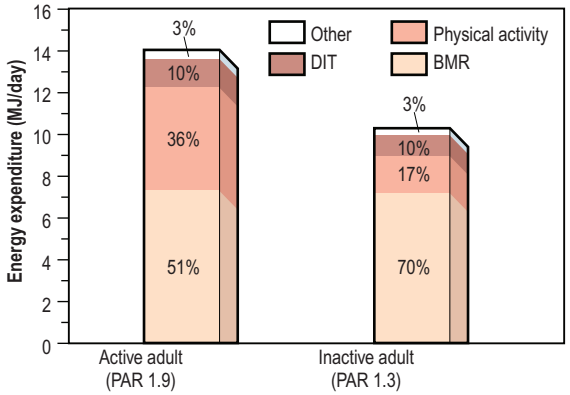


Fig. 5.2 **Daily energy expenditure in an active and a sedentary 70 kg adult.** BMR, basal metabolic rate; DIT, dietary induced thermogenesis; PAR, physical activity ratio.

However, measurement of energy expenditure gives a more accurate assessment of requirements.

Energy expenditure

Daily energy expenditure (Fig. 5.2) is the sum of:

- the basal metabolic rate (BMR)
- the thermic effect of food eaten
- occupational activities
- non-occupational activities.

Total energy expenditure can be measured using a double-labelled water technique. Water containing the stable isotopes ^{2}H and ^{18}O is given orally. As energy is expended carbon dioxide and water are produced. The difference between the rates of loss of the two isotopes is used to calculate the carbon dioxide production, which is then used to calculate energy expenditure. This can be done on urine samples over a 2- to 3-week period with the subject ambulatory. The technique is accurate, but it is expensive and requires the availability of a mass spectrometer. An alternative tracer technique for measuring total energy expenditure is to estimate CO_2 production by isotopic dilution. A subcutaneous infusion of labelled bicarbonate is administered continuously by a minipump, and urine is collected to measure isotopic dilution by urea, which is formed from CO_2. Other methods for estimating energy expenditure, such as heart rate monitors or activity monitors, are also available but are less accurate.

Basal metabolic rate. The BMR can be calculated by measuring oxygen consumption and CO_2 production, but it is more usually taken from standardized tables (Table 5.1) that require knowledge of the subject's age, weight and sex.

Physical activity. The physical activity ratio (PAR) is expressed as multiples of the BMR for both occupational and non-occupational activities of varying intensities (Table 5.2).

$$\text{Total daily energy expenditure} = \text{BMR} \times [\text{Time in bed} + (\text{Time at work} \times \text{PAR}) + (\text{Non-occupational time} \times \text{PAR})].$$

Thus, for example, to determine the daily energy expenditure of a 66-year-old, 50 kg female doctor, with a BMR of 4805 kJ per day spending one-third of a day sleeping, working or

Table 5.1 Equations for the prediction of basal metabolic rate (in MJ per day)

Age range (years)	Prediction equation (BMR =)	95% confidence limits
Men		
10–17	0.074 (wt)* + 2.754	±0.88
18–29	0.063 (wt)* + 2.896	±1.28
30–59	0.048 (wt)* + 3.653	±1.40
60–74	0.0499 (wt)* + 2.930	N/A
75+	0.0350 (wt)* + 3.434	N/A
Women		
10–17	0.056 (wt)* + 2.898	± 0.94
18–29	0.062 (wt)* + 2.036	± 1.00
30–59	0.034 (wt)* + 3.538	± 0.94
60–74	0.0386 (wt)* + 2.875	N/A
75+	0.0410 (wt)* + 2.610	N/A

Data reproduced with permission of Department of Health, 1991
*Bodyweight (wt) in kilograms

Table 5.2 Physical activity ratio (PAR) for various activities (expressed as multiples of BMR)

	PAR
Occupational activity	
Professional/Housewife	1.7
Domestic helper/Sales person	2.7
Labourer	3.0
Non-occupational activity	
Reading/Eating	1.2
Household/Cooking	2.1
Gardening/Golf	3.7
Jogging/Swimming/Football	6.9

engaged in non-occupational activities, the latter at a PAR of 2.1, the following calculation ensues:

$$(4805\,\text{kJ/day}) \times [0.3 + (0.3 \times 1.7) + (0.3 \times 2.10)] = 6919\,\text{kJ or } 1655\,\text{kcal/day}.$$

In the UK the estimated 'average' daily requirement is:

- for a 55-year-old female – 8100 kJ (1940 kcal)
- for a 55-year-old male – 10600 kJ (2550 kcal).

This is at present made up of about 50% carbohydrate, 35% fat, 15% protein plus or minus 5% alcohol. In developing countries, however, carbohydrate may be more than 75% of the total energy input, and fat less than 15% of the total energy input.

Energy requirements increase during the growing period, with pregnancy and lactation, and sometimes following infection or trauma. In general, the increased BMR associated with inflammatory or traumatic conditions is counteracted or more than counteracted by a decrease in physical activity, so that total energy requirements are not increased.

In the basal state, energy demands for resting muscle are 20% of the total energy required, abdominal viscera 35–40%, brain 20% and heart 10%. There can be more than a 50-fold increase in muscle energy demands during exercise.

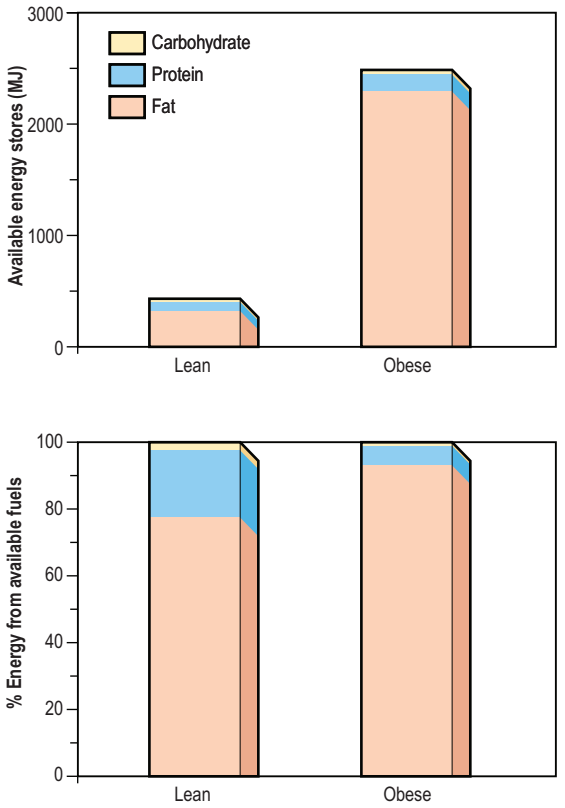


Fig. 5.3 **Available energy reserves from different fuels expressed in MJ (upper) and as a percentage of total (lower) in a lean hypothetical 70 kg adult and an obese 140 kg adult.**

Energy stores

Although virtually all body fat and glycogen are available for oxidation, less than half the protein is available for oxidation. Figure 5.3 shows that fat accounts for the largest reserves of energy in both lean and obese subjects. The size of the stores determines survival during starvation.

Bodyweight

Bodyweight depends on energy balance. Intake depends not only on food availability but also on a number of complex interrelationships that include the stimulus of good food, the role of hunger, metabolic changes (e.g. hypoglycaemia), and the pleasure and habit of eating. Some people are able to keep their bodyweight constant within a few kilograms for many years, but most gradually increase their weight owing to a small but continuous increase of intake over expenditure. A gain or loss of energy of 25–29 MJ (6000–7000 kcal) would respectively increase or decrease bodyweight by approximately 1 kg.

Protein

In the UK the adult daily RNI for protein is 0.75 g/kg, with protein representing at least 10% of the total energy intake. Most affluent people eat more than this, consuming 80–100 g of protein per day. The total amount of nitrogen excreted in the urine represents the balance between protein breakdown and synthesis. In order to maintain nitrogen balance, at least 40–50 g of protein are needed. The amount of protein oxidized can be calculated from the amount of nitrogen excreted in the urine over 24 hours using the following equation:

Grams of protein required = Urinary nitrogen × 6.25
(most proteins contain about 16% of nitrogen).

In practice, urinary urea is more easily measured and forms 80–90% of the total urinary nitrogen (N). In healthy individuals urinary nitrogen excretion reflects protein intake. However, urine N excretion does not match intake either in catabolic conditions (negative N balance) or during growth or repletion following an illness (positive N balance).

Protein contains many **amino acids**, of which nine are *indispensable (essential)*: trytophan, histidine, methionine, threonine, isoleucine, valine, phenylalanine, lysine, leucine. These amino acids cannot be synthesized and must be provided in the diet. The *dispensable (non-essential)* amino acids can be synthesized in the body, but some may still be needed in the diet unless adequate amounts of their precursors are available. Animal proteins, such as in milk, meat and eggs, are of high nutritional value as they contain a good balance of all indispensable amino acids. Conversely, many proteins from vegetables are deficient in at least one indispensable amino acid.

In developing countries, adequate protein intake is achieved mainly from vegetable proteins. By combining foodstuffs with different low concentrations of indispensable amino acids (e.g. maize with legumes), protein intake can be adequate provided enough vegetables are available.

Loss of protein from the body (negative N balance) occurs not only because of inadequate protein intake, but also owing to inadequate energy intake. When there is loss of energy from the body, more protein is directed towards oxidative pathways and eventually gluconeogenesis for energy. Of all the amino acids, glutamine is quantitatively the most significant in the circulation and in inter-organ exchange. Alanine is released from muscle; it is deaminated and converted into pyruvic acid before entering the citric acid cycle. Homocysteine is a sulphur-containing amino acid which is derived from methionine in the diet. A raised plasma concentration is an independent risk factor for vascular disease (see p. 747).

Amino acids may be utilized to synthesize products other than protein or urea. For example:

- haem requires glycine
- melanin and thyroid hormones require tyrosine
- nucleic acid bases require glutamine, aspartate and glycine
- glutathione, which is part of the defence system against free radicals, requires glutamate, cysteine and glycine.

Fat

Dietary fat is chiefly in the form of triglycerides, which are esters of glycerol and free fatty acids. Fatty acids vary in chain length and in saturation (Table 5.3). Unsaturated fatty acids are monounsaturated or polyunsaturated. The hydrogen molecules related to these double bonds can be in the *cis* or the *trans* position, most natural fatty acids in food being in the *cis* position (Box 5.1).

The *essential fatty acids* (EFAs) are linoleic and α-linolenic acid, both of which are precursors of prostaglandins. Eicosapentaenoic and docosahexaenoic are also necessary, but can be made to a limited extent in the tissues from linoleic and linolenic, and thus a dietary supply is not essential.

Table 5.3 The main fatty acids in foods

Saturated
Lauric C12:0
Myristic C14:0
Palmitic C16:0
Stearic C18:0

Monounsaturated
Oleic C18:1 (n-9)
Elaidic C18:1 (n-9 *trans**)

Polyunsaturated
Linoleic C18:2 (n-6)
α-Linolenic C18:3 (n-3)
Arachidonic C20:4 (n-6)
Eicosapentaenoic C20:5 (n-3)
Docosahexaenoic C22:6 (n-3)

The number of carbon atoms is indicated before the colon; the number of double bonds after the colon. In parentheses the positions of the double bonds (designated either n as here or ω) are shown counted from the methyl end of the molecule. All double bonds are in the *cis* position except that marked with an asterisk (*)

Box 5.1 Dietary sources of fatty acids

Type of acid	Sources
Saturated fatty acids	Mainly animal fat
n-6 fatty acids	Vegetable oils and other plant foods
n-3 fatty acids	Vegetable foods, rapeseed oil, fish oils
trans fatty acids	Hydrogenated fat or oils, e.g. in margarine, cakes, biscuits

Synthesis of triglycerides, sterols and phospholipids is very efficient, and even with low-fat diets subcutaneous fat stores can be normal.

Dietary fat provides 37 kJ (9 kcal) of energy per gram. A high fat intake has been implicated in the causation of:

- cardiovascular disease
- cancer (e.g. breast, colon and prostate)
- obesity
- type 2 diabetes.

The data on causation are largely epidemiological and disputed by many. Nevertheless, it is often suggested that the consumption of saturated fatty acids should be reduced, accompanied by an increase in monounsaturated fatty acids (the 'Mediterranean diet') or polyunsaturated fatty acids. Any increase in polyunsaturated fats should not, however, exceed 10% of the total food energy, particularly as this requires a big dietary change.

Increased consumption of hydrogenated vegetable and fish oils in margarines has led to an increased *trans* fatty acid consumption and their intake should not, on present evidence, increase more than the current estimated average of 5 g per day or 2% of the dietary energy. This is because *trans fatty acids* (also called *trans* fats) behave as if they were saturated fatty acids, increasing circulating LDL and decreasing HDL cholesterol concentrations, which in turn increase the risk of cardiovascular disease. For this reason, the US Food and Drug Administration (FDA) ruled that from 2006 the nutrition labels for all conventional foods and supplements must indicate the content of *trans* fatty acids. *Trans* fatty acids from partially hydrogenated oils appear to provide no nutritional benefit and considerable potential for harm.

Ask the authors
www.studentconsult.com

The **n-6 polyunsaturated fatty acids (PUFA)** are components of membrane phospholipids, influencing membrane fluidity and ion transport. They also have antiarrhythmic, antithrombotic and anti-inflammatory properties, all of which are potentially helpful in preventing cardiovascular disease.

The **n-3 PUFA** increase circulating high-density lipoprotein (HDL) cholesterol and lower triglycerides, both of which might reduce cardiovascular risk. Some of the actions of n-3 PUFA are mediated by a range of leukotrienes and eicosanoids, which differ in pattern and functions from those produced from n-6 PUFA.

Epidemiological studies and clinical intervention studies suggest that n-3 PUFA may have effects in the secondary prevention of cardiovascular disease and 'all-cause mortality' (e.g. 20–30% reduction in mortality from cardiovascular disease according to some studies). The benefits, which have been noted as early as 4 months after intervention, have been largely attributed to the antiarrhythmic effects of n-3 PUFA, but some work suggests that n-3 PUFA, administered as capsules, can be rapidly incorporated into atheromatous plaques, stabilizing them and preventing rupture. Whether these effects are due directly to n-3 PUFA or other changes in the diet is still debated.

The *GISSI Prevention Trial*, which followed over 11 000 patients for 3.5 years after a myocardial infarction, administered the fish oils (eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA)) in the form of capsules and demonstrated a striking benefit in reducing mortality. The effects of vitamin E (300 mg α-tocopherol/day) were also studied, but no benefit was found.

The British Nutrition Foundation and the American Heart Association presently recommend a two-fold increase of the current intake of total n-3 PUFA (several fold increase in the intake of fish oils, and a 50% increase in the intake of α-linolenic acid). Implementing this recommendation will mean either a major change in the dietary habits of populations that eat little fish, or ingestion of capsules containing fish oils. Some government agencies have warned of the hazards of eating certain types of fish, which increase the risk of mercury poisoning and possibly other toxicities.

The current recommendations for fat intake for the UK are as follows:

- saturated fatty acids should provide approximately 10% of the dietary energy
- *cis*-monounsaturated acids (mainly oleic acid) should continue to provide approximately 12% of the dietary energy
- *cis*-polyunsaturated acids should provide 6% of dietary energy, and are derived from n-6 and n-3 polyunsaturated fatty acids, which should provide ~0.5% of total energy intake
- total fat intake should be no more than 35% of the total dietary energy, and restriction to 30% is desirable.

Cholesterol is found in all animal products. Eggs are particularly rich in cholesterol, which is virtually absent from plants. The average daily intake in the UK is 300–500 mg. Cholesterol is also synthesized (see p. 321) and only very high or low dietary intakes will significantly affect blood levels.

Essential fatty acid deficiency

Essential fatty acid deficiency may accompany protein–energy malnutrition (PEM), but it has been clearly defined as a clinical entity only in patients on long-term parenteral nutrition given glucose, protein and no fat. Alopecia, thrombocytopenia, anaemia and a dermatitis occur within weeks with an increased ratio of triene (n-9) to tetraene (n-6) in plasma fatty acids.

Carbohydrate

Carbohydrates are readily available in the diet, providing 17 kJ (4 kcal) per gram of energy (15.7 kJ (3.75 kcal) per gram monosaccharide equivalent).

Carbohydrate intake comprises the polysaccharide starch, the disaccharides (mainly sucrose) and monosaccharides (glucose and fructose). Carbohydrate is cheap compared with other foodstuffs; a great deal is therefore eaten, usually more than required.

Dietary fibre, which is largely *non-starch polysaccharide* (NSP) (entirely NSP according to some authorities), is often removed in the processing of food. This leaves highly refined carbohydrates such as sucrose which contribute to the development of dental caries and obesity. Lignin is included in dietary fibre in some classification systems, but it is not a polysaccharide. It is only a minor component of the human diet.

The principal classes of NSP are:

- cellulose
- hemicelluloses
- pectins
- gums.

None of these are digested by gut enzymes. However, NSP is partly broken down in the gastrointestinal tract, mainly by colonic bacteria, producing gas and volatile fatty acids, e.g. butyrate.

All plant food, when unprocessed, contains NSP, so that all unprocessed food eaten will increase the NSP content of the diet. Bran, the fibre from wheat, provides an easy way of adding additional fibre to the diet: it increases faecal bulk and is helpful in the treatment of constipation.

The average daily intake of NSP in the diet is approximately 16 g. NSP deficiency is accepted as an entity by many authorities and it is suggested that the total NSP be increased to up to 30 g daily. This could be achieved by increased consumption of bread, potatoes, fruit and vegetables, with a reduction in sugar intake in order not to increase total calories. Each extra gram of fibre daily adds approximately 3–5 g to the daily stool weight. Pectins and gums have been added to food to slow down monosaccharide absorption, particularly in type 2 diabetes.

Eating a diet rich in plant foods (fruits, vegetables, cereals and whole grain) is generally recommended for general health promotion, including protection against ischaemic heart disease, stroke and certain types of cancers. This has been attributed to a lipid lowering effect, the presence of protective substances, such as vitamin and non-vitamin antioxidants, and other vitamins such as folic acid, which is linked to homocysteine metabolism, a risk factor for cardiovascular disease. Fermentation of fibre in the colon may protect against development of colonic cancer. However, associated lifestyle factors such as low physical activity may also help explain some of those associations.

Health promotion

Many chronic diseases – particularly obesity, diabetes mellitus and cardiovascular disease – cause premature mortality and morbidity and are potentially preventable by dietary change.

Box 5.2 suggests the composition of the 'ideal healthy diet'. The values given are based on the principle of:

- reducing total fat in the diet, particularly saturated fat
- increasing consumption of fish which contain n-3 (or ω-3) polyunsaturated fatty acids
- increasing intake of whole-grain cereals, green and orange vegetables and fruits, leading to an increase in fibre and antioxidants.

Reductions in dietary sodium and cholesterol have also been suggested. There would be no disadvantage in this, and most studies have suggested some benefit.

Fortification of foods with specific nutrients is common. In the UK *margarine and milk* are fortified with vitamins A and D, *flour* with calcium, iron, thiamin and niacin, and *breakfast cereals* with several vitamins and iron. Not all substances used in fortification have nutritive value. For example, *Olestra* is a polymer of sucrose and six or more triglycerides which has been introduced to combat obesity. It is not absorbed and is therefore used particularly in savoury snack foods (where it has FDA approval) as a 'fake fat'. Therefore, it results in a reduction in total calories. It has side-effects, e.g. loose stools, abdominal cramps, and its use is being carefully monitored.

The interests of the individual are often different from those associated with government policy. A distinction needs to be made between nutrient goals and dietary guidelines. *Nutrient goals* refer to the national intakes of nutrients that are considered appropriate for optimal health in the population, whereas *dietary guidelines* refer to the dietary methods used to achieve these goals. Since dietary habits in different countries vary, dietary guidelines may also differ, even when the nutrient goals are the same. Nutrient goals are based on scientific information that links nutrient intake to disease. Although the information is incomplete, it includes evidence from a wide range of sources, including experimental animal studies, clinical studies and both short-term and long-term epidemiological studies.

Box 5.2 Recommended healthy dietary intake

Dietary component	Approximate amounts given as % of total energy unless otherwise stated	General hints
Total carbohydrate	55 (55–75)	Increase fruit, vegetables, beans, pasta, bread
Free sugar	10 (<10)	Decrease sugary drinks
Protein	15 (10–15)	Decrease red meat (see fat below)
Total fat	30 (15–30)	Increase vegetable (including olive oil) and fish oil and decrease animal fat
Saturated fat	10 (<10)	
Unsaturated fat	20	
Approximate amounts		
Cholesterol	<300 (<300) mg/day	Decrease meat and eggs
Salt	<6 (<5) g/day	Decrease prepared meats and do not add extra salt to food
Total dietary fibre	30 (>25) g/day	Increase fruit and vegetables and wholegrain foods

Values in parentheses are goals for the intake of populations, as given by the WHO (including populations who are already on low-fat diets). Some of the extreme ranges are not realistic short-term goals for developed countries, e.g. 75% of total energy from carbohydrate and 15% fat. When total energy intake is 2500 kcal (10 500 kJ) per day, 55% of intake comes from carbohydrate (344 g, i.e. 1376 kcal (5579 kJ)) and 30% from fat (83 g, i.e. 747 kcal (3137 kJ)).

FURTHER READING

FAO/WHO/UNU Expert Consultation. Human energy requirements. In: *Food and Nutrition Technical Series 1*. Rome: FAO, 2004.

Ferguson LR, Harris PJ. The dietary fibre debate. *Lancet* 2003; **361**: 1487–1488.

Institute of Medicine. *Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Protein, and Amino Acids*. Washington, DC: The National Academies Press, 2005.

Thies F, Garry JMC, Yaqoob P et al. Association of n-3 polyunsaturated fatty acids with stability of atherosclerotic plaques: a randomised controlled trial. *Lancet* 2003; **361**: 477–485.

WHO/FAO Expert Consultation. *Diet, Nutrition and the Prevention of Chronic Diseases*. Geneva: WHO, 2003.

PROTEIN-ENERGY MALNUTRITION (PEM)

IN DEVELOPED COUNTRIES

Starvation uncomplicated by disease is relatively uncommon in developed countries, although some degree of undernourishment is seen in very poor areas. Most nutritional problems occurring in the population at large are due to eating wrong combinations of foodstuffs, such as an excess of refined carbohydrate or a diet low in fresh vegetables. Undernourishment associated with disease is common in hospitals and nursing homes, and Table 5.4 gives a list of conditions in which malnutrition is often seen. Surgical complications, with sepsis, are a common cause. Many patients are admitted to hospital undernourished, and a variety of chronic conditions predispose to this state (Table 5.5).

Table 5.4 Common conditions associated with protein–energy malnutrition

Sepsis	Dementia
Trauma	Malignancy
Surgery, particularly of GI tract with complications	Any very ill patient
GI disease, particularly involving the small bowel	Severe chronic inflammatory diseases
	Psychosocial: poverty, social isolation, anorexia nervosa, depression

The majority of the weight loss, leading to malnutrition, *is due to poor intake secondary to the anorexia associated with the underlying condition*. Disease may also contribute by causing malabsorption and increased catabolism, which is mediated by complex changes in cytokines, hormones, side-effects of drugs, and immobility. The elderly are particularly at risk of malnutrition because they often suffer from diseases and psychosocial problems such as social isolation or bereavement (Table 5.5).

Pathophysiology of starvation (Fig. 5.4)

In the first 24 hours following low dietary intake, the body relies for energy on the breakdown of hepatic glycogen to glucose. Hepatic glycogen stores are small and therefore gluconeogenesis is soon necessary to maintain glucose levels. Gluconeogenesis takes place mainly from pyruvate, lactate, glycerol and amino acids, especially alanine and glutamine. The majority of protein breakdown takes place in muscle, with eventual loss of muscle bulk.

Lipolysis, the breakdown of the body's fat stores, also occurs. It is inhibited by insulin, but the level of this hormone falls off as starvation continues. The stored triglyceride is hydrolysed by lipase to glycerol, which is used for gluconeogenesis, and also to non-esterified fatty acids that can be used directly as a fuel or oxidized in the liver to ketone bodies.

As starvation continues, *adaptive processes* take place to prevent the body's available protein being completely utilized. There is a decrease in metabolic rate and total body energy expenditure. Central nervous metabolism changes from glucose as a substrate to ketone bodies. Gluconeogenesis in the liver decreases as does protein breakdown in muscle, both of these processes being inhibited directly by ketone bodies. Most of the energy at this stage comes from adipose tissue, with some gluconeogenesis from amino acids, particularly from alanine in the liver, and glutamine in the kidney.

The *metabolic response to prolonged starvation* differs between lean and obese individuals. One of the major differences concerns the proportion of energy derived from protein oxidation, which determines the proportion of weight loss from lean tissues. This proportion may be up to three times smaller in obese subjects than lean subjects. It can be regarded as an adaptation which depends on the composition of the initial reserves (Fig. 5.3). This means that deterioration in body function is more rapid in lean subjects. Furthermore, survival time is much less in lean subjects (~2 months), compared to the obese (can be at least several months).

Following trauma or shock, some of the adaptive changes do not take place. Glucocorticoids and cytokines (see below)

Table 5.5 Nutritional consequences of disease and the underlying risk factors (physical/psychosocial problems)

Risk factors	Consequence
Underlying disease	
Almost any moderate/severe chronic disease Recovery from severe acute/subacute disease	Anorexia, increased requirements for some nutrients, and other effects indicated below (depending on condition)
Physical problems	
Muscle weakness (respiratory and peripheral muscles) and/or incoordination Severe arthritis in hands and arms	Problems with shopping, cooking and eating
Swallowing problem (neurological causes), painful or obstructive conditions of mouth and gastrointestinal tract (GIT)	Inadequate food intake, and/or risk of aspiration pneumonia
GIT symptoms (e.g. nausea, vomiting, diarrhoea, jaundice)	Food aversion, malabsorption (small bowel disease), anorexia
Sensory deficit (e.g. impaired sight, hearing and other deficits)	Difficulties in shopping, cooking and/or decreased intake of food
Psychosocial problems	
Loneliness, depression, bereavement, confusion, living alone, poverty, alcoholism, drug addiction	Self-neglect, inadequate intake of food or quality of food
Multiple drug use (polypharmacy)	Indicates severe disease or multiple physical and psychosocial problems; drugs may lead to confusion, sedation, depression and GIT side-effects (including malabsorption of nutrients)

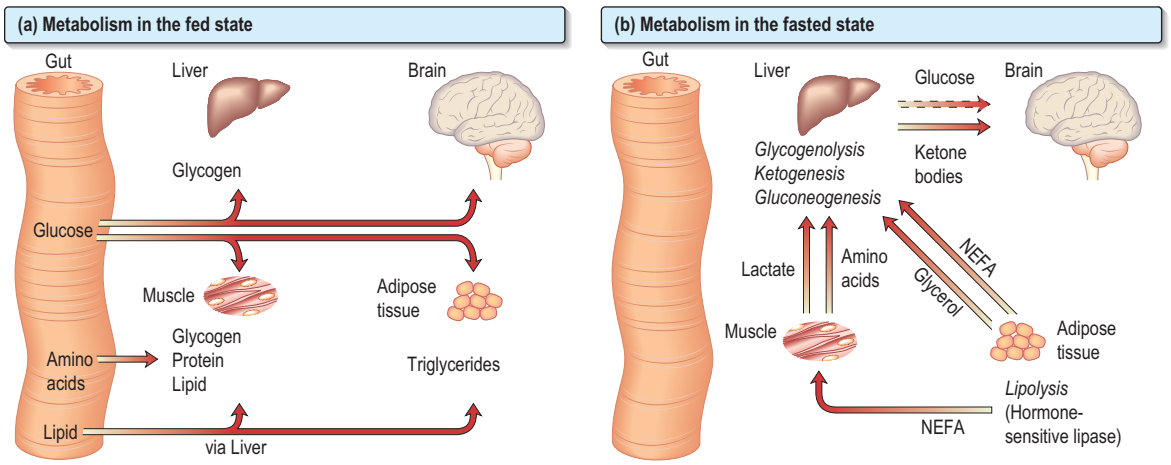


Fig. 5.4 **Metabolism in (a) the fed and (b) the fasted state.** NEFA, non-esterified fatty acids.

stimulate the ubiquitin–proteasome pathway in muscle, which is responsible for accelerated proteolysis in muscle in many catabolic illnesses. In starvation there is a decrease in BMR, whilst in inflammatory and traumatic disease the BMR is often increased. These changes all result in continuing gluconeogenesis with massive muscle breakdown, and further reduction in survival time.

Regulation of metabolism

Tissue metabolism is regulated by multiple coordinated processes. Some are rapid involving nerves, whilst others are slower involving circulating substrates and hormones. Factors include:

- *Circulating substrate concentrations*. The uptake and metabolism of ketone bodies, which serve as the major fuel for the brain during prolonged starvation, is primarily determined by the circulatory concentration which can increase up to 5 mmol/L or more. The liver is responsible for producing ketone bodies, the production of which is in turn controlled by the availability of fatty acids derived from adipose tissue. Substrates may also compete with each other for metabolism, for example glucose competes with non-esterified fatty acids for uptake and metabolism in muscle and heart (the glucose–fatty acid cycle) and this is independent of hormones.
- *Blood flow.* The delivery of substrates (and other signals) to tissues depends not only on their circulating concentration but also on the blood flow to tissues. In many tissues there is coupling between metabolic activity and blood flow, with arterioles regulating blood flow to the tissue according to demand, e.g. blood flow to muscle increases during exercise.
- *Signals.* Hormones and other signals, such as cytokines, regulate intracellular metabolism.

In the fed state, insulin/glucagon ratios are high. Insulin promotes synthesis of glycogen, protein and fat, and inhibits lipolysis and gluconeogenesis.

In the fasted state, insulin/glucagon ratios are low. Glucagon acts mainly on the liver and has no action on muscle. It increases glycogenolysis and gluconeogenesis, as well as increasing ketone body production from fatty acids. It also stimulates lipolysis in adipose tissue. Catecholamines have a similar action to glucagon but also affect muscle metabolism. These agents both act via cyclic adenosine monophosphate (cAMP) to stimulate lipolysis, producing free fatty acids that can then act as a major source of energy.

During weight loss, uncomplicated by disease, the proportion of lean to fat tissue loss (or proportion of energy derived from protein metabolism) is greater in lean than overweight/obese individuals.

During acute disease, loss of lean tissue, which is associated with protein oxidation, can be particularly rapid. Hormones such as corticosteroids, pro-inflammatory cytokines and insulin resistance are all involved.

The metabolic response to trauma, injury and inflammation depends on the balance between pro-inflammatory (e.g. tumour necrosis factor (TNF), interleukin-2 (IL-2)) and anti-inflammatory cytokines (e.g. IL-10), and the production of many of these cytokines is influenced by genetic polymorphisms. Since many chronic diseases, including atherosclerosis, have an inflammatory component, these changes have wide-reaching metabolic implications.

Cytokines such as IL-1, IL-6 and TNF play a significant role in regulating metabolism. In acute diseases they contribute to the catabolic process, glycogenolysis, and acute-phase protein synthesis. TNF, which inhibits lipoprotein lipase, is a 'cachexia factor' in patients with cancer.

It is unclear how these cytokines interact with central feeding pathways to cause anorexia. However, in animal models of both cancer and inflammatory bowel disease, many peripheral and central mediators of appetite are involved. For example, neuropeptide Y levels in the hypothalamus are often inappropriately low, so there is a reduced drive to feeding.

Clinical features

Patients are sometimes seen with loss of weight or malnutrition as the primary symptom (failure to thrive in children). Mostly, however, malnourishment is only seen as an accompaniment of some other disease process, such as malignancy. Severe malnutrition is seen mainly with advanced organic disease or after surgical procedures followed by complications. Three key features which help in the detection of chronic protein–energy malnutrition (PEM) in adults are:

1. *The body mass index (BMI):*
   - Probable chronic PEM: <18.5 kg/m^2
   - Possible chronic PEM: 18.5–20 kg/m^2
   - Little or no risk of chronic PEM: >20 kg/m^2.

In patients with oedema or dehydration the BMI may be somewhat misleading.

2. *Weight loss in previous 3–6 months:* $>10\%$, high risk; 5–10%, possible risk; $<5\%$ low/no risk of developing PEM.
3. *Acute disease effect:* Diseases that have resulted or are likely to result in no dietary intake for more than 5 days are associated with a high risk of malnutrition (e.g. prolonged unconsciousness, persistent swallowing problems after a stroke, or prolonged ileus after abdominal surgery).

Other factors:

- History of decreased food intake/loss of appetite
- Clothes becoming loosely fitting (weight loss) and a general appearance indicating obvious wasting
- Physical and psychosocial disturbances likely to have contributed to the weight loss.

These factors act as a link between detection and management (Fig. 5.5, The Malnutrition Universal Screening Tool). If the underlying physical or psychosocial problems are not adequately addressed, treatment may not be successful.

PEM leads to a depression of the immunological defence mechanism, resulting in a decreased resistance to infection. It also detrimentally affects muscle strength and fatigue, reproductive function (e.g. in anorexia nervosa, which is common in adolescent girls; p. 1219), wound healing, and psychological function (depression, anxiety, hypochondriasis, loss of libido).

Treatment (see also pp. 233 and 236)

When malnutrition is obvious and the underlying disease cannot be corrected at once, some form of nutritional support is necessary. Nutrition should be given enterally if the gastrointestinal tract is functioning adequately. This can most easily be done by encouraging the patient to eat more often

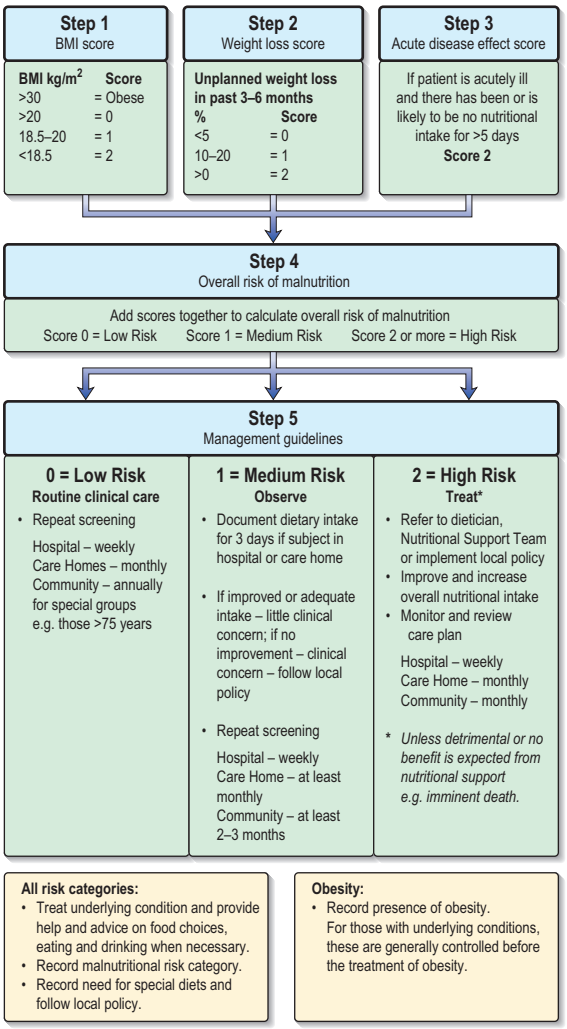


Fig. 5.5 **Malnutrition Universal Screening Tool (MUST).** Reproduced with permission from the British Association for Parenteral and Enteral Nutrition (BAPEN): http://www.bapen.org.uk.

and by giving a high-calorie supplement. If this is not possible, a liquefied diet may be given intragastrically via a fine-bore tube or by a percutaneous endoscopic gastrostomy (PEG). If both of these measures fail, parenteral nutrition is given.

(a)

(b)

Fig. 5.6 **Malnourished children: (a) marasmus and (b) kwashiorkor.** Courtesy of Dr Paul Kelly.

IN DEVELOPING COUNTRIES

The International Union of Nutritional Sciences, with support from the International Pediatric Association, launched a global Malnutrition Task Force in 2005 to ensure that an integrated system of prevention and treatment of malnutrition is actively supported.

In many areas of the world, people are on the verge of malnutrition due to extreme poverty. In addition, if events such as drought, war or changes in political climate occur, millions suffer from starvation. Although the basic condition of PEM is the same in all parts of the world from whatever cause, malnutrition resulting from long periods of near-total starvation produces unique clinical appearances in children virtually never seen in high-income countries. The term 'protein–energy malnutrition' covers the spectrum of clinical conditions seen in adults and children. Children under 5 years may present with:

- *Marasmus* is the childhood form of starvation, which is associated with obvious wasting. The child looks emaciated, there is obvious muscle wasting and loss of body fat. There is no oedema. The hair is thin and dry (Fig. 5.6a). The child is not so apathetic or anorexic as with kwashiorkor. Diarrhoea is frequently present and signs of infection must be looked for carefully.
- *Kwashiorkor* occurs typically in a young child displaced from breast-feeding by a new baby. It is often precipitated by infections such as measles, malaria and diarrhoeal illnesses. The child is apathetic and lethargic with severe anorexia. There is generalized oedema with skin pigmentation and thickening (Fig. 5.6b). The hair is

Table 5.6 Classification of childhood malnutrition

	Moderate malnutrition	Severe malnutrition*
Symmetrical oedema	No	Yes (oedematous malnutrition)†
Weight-for-height	SD score –3 to –2 (70–79%)‡	SD score below –3 (<70%)‡ (severe wasting)§
Height-for-age	SD score –3 to –2 (85–89%)‡	SD score below –3 (<85%)‡ (severe stunting)

*The diagnoses are not mutually exclusive
†This includes kwashiorkor and marasmic-kwashiorkor in the older classifications. To avoid confusion with the clinical syndrome of kwashiorkor, which includes other features (see text), the term oedematous malnutrition is preferred
‡Percentage of the median National Centre for Health Statistics/WHO reference
§This corresponds to marasmus (without oedema) in the Wellcome classification and to grade II in the Gomez classification. To avoid confusion the term 'severe' wasting is preferred

dry, sparse and may become reddish or yellow in colour. The abdomen is distended owing to hepatomegaly and/or ascites. The serum albumin is always low. The exact cause is unknown, but theories related to diet (low in protein, and high in carbohydrate) and free radical damage in the presence of inadequate antioxidant defences have been proposed.

A classification of severe malnutrition by the *World Health Organization* (Table 5.6) makes no distinction between kwashiorkor and marasmus, because their approach to treatment is similar.

The World Health Organization (WHO) classification of chronic undernutrition in children is based on standard deviation (SD) scores. Thus, children with an SD score between –2 and –3 (between 3 and 2 standard deviation scores below the median – corresponding to a value between 0.13 and 2.3 centile) can be regarded as being at moderate risk of undernutrition, and below an SD score of –3, of severe malnutrition. A low weight-for-height is a measure of thinness (wasting when pathological), and a low height-for-age is a measure of shortness (stunting when pathological). Those with oedema and clinical signs of severe malnutrition are classified as having oedematous malnutrition.

Starvation in adults may lead to extreme loss of weight depending upon the severity and duration. They may crave for food, are apathetic and complain of cold and weakness with a loss of subcutaneous fat and muscle wasting. The WHO classification is based on body mass index, with a value less than 18.5 kg/m^2 indicating malnutrition (severe malnutrition if less than 16.0 kg/m^2).

Severely malnourished adults and children are very susceptible to respiratory and gastrointestinal infections, leading to an increased mortality in these groups.

Investigations

These are not always practicable in field work.

- **Blood tests**
  a) Anaemia due to folate, iron and copper deficiency is often present, but the haematocrit may be high owing to dehydration.
  b) Eosinophilia suggests parasitic infestation.
  c) Electrolyte disturbances are common.
  d) Malarial parasites should be looked for.
  e) HIV tests.
- **Stools** should be examined for parasitic infestations.
- **Chest X-ray** – tuberculosis is common and is easily missed if a chest X-ray is not performed.

Treatment

Treatment must involve the provision of protein and energy supplements and the control of infection.

Resuscitation and stabilization

The severely ill child will require:

- Correction of fluid and electrolyte abnormalities, but intravenous therapy should be avoided if possible because of the danger of fluid overload.
- Treatment of shock with oxygen.
- Treatment of hypoglycaemia (blood glucose <3 mmol/L), hypothermia (reduce heat loss, and provide additional heat if necessary) and infection (antibiotics) – these often coexist.

The standard WHO *oral hydration solution* has a high sodium and low potassium content and is not suitable for severely malnourished children. Instead, the rehydration solution for malnutrition (ReSoMal) is recommended. It is commercially available but can also be produced by modification of the standard WHO oral hydration solution.

Infection is common. Diarrhoea is often due to bacterial or protozoal overgrowth; metronidazole is very effective and is often given routinely. Parasites are also common and, as facilities for stool examination are usually not available, mebendazole 100 mg twice daily should be given for 3 days. In high-risk areas, antimalarial therapy is given.

Large doses of vitamin A are also given because deficiency of this vitamin is common. After the initial resuscitation, further stabilization over the next few days is undertaken, as indicated in Table 5.7.

Refeeding

This needs to be planned carefully. During the initial treatment of the acute situation, a balanced diet with sufficient protein and energy is given to maintain a steady state. Large increases in energy can lead to heart failure, circulatory collapse and death (*refeeding syndrome*). Initial feeding involves administration of feeds of low osmolarity and low in lactose (WHO recommendations: 100 kcal/kg/day, 1.0–1.5 g protein/kg/day and 130 mL liquid/kg/day (100 mL/kg/day if the child has marked oedema)). Attempts should be made to give the feeds slowly and frequently (e.g. 2-hourly during days 1–2; 3-hourly during days 3–5; and 4-hourly thereafter), although anorexia is often a problem and can be exacerbated by excessive feeding. If necessary, fluids and food should be given by nasogastric tube. The child is then gradually weaned to liquids and then solids by mouth. All severely malnourished children have vitamin and mineral deficiencies. Although anaemia is common, the WHO recommends giving iron only after the child develops a good appetite and starts gaining weight, because of concern about making the infection worse (iron is a pro-oxidant). The child should be given daily micronutrient supplements for at least 2 weeks. These should include a multivitamin supplement with folic acid, zinc and copper.

Table 5.7 Time frame for the management of the child with severe malnutrition (the 10-step approach recommended by the WHO)

	Stabilization		Rehabilitation	Follow-up
	Days 1–2	Days 3–7	Weeks 2–6	Weeks 7–26
1. Treat or prevent hypoglycaemia	→			
2. Treat or prevent hypothermia	→			
3. Treat or prevent dehydration	→			
4. Correct electrolyte imbalance	→	→	→	
5. Treat infection	→	→		
6. Correct micronutrient deficiencies	Without iron →	→	With iron →	
7. Begin feeding	→	→		
8. Increase feeding to recover lost weight			→	→
9. Stimulate emotional and sensorial development	→	→	→	→
10. Prepare for discharge			→	

Rehabilitation

Gradually, as the child improves, more energy can be given, and during rehabilitation weight gain is achieved by providing extra energy and protein ('catch-up weight gain'). Children who have been severely ill need constant attention right through the convalescent period, as often home conditions are poor and feeds are refused. Sensory stimulation and emotional support is a major component of management during both the stabilization and rehabilitation phases.

Adults do not usually suffer such severe malnutrition, but the same general principles of treatment should be followed.

Care setting

There are not enough hospitals or therapeutic feeding centres to cope with the malnutrition problem (even acute malnutrition problems), which emphasizes the need for outpatient and community based programmes. These may involve the use of ready-to-use therapeutic foods, such as energy-dense pastes with minerals and vitamins, without the need to add water, which could potentially contaminate the food.

Prognosis

Children with extreme malnutrition have a mortality of over 50%. By careful management this can be reduced significantly to less than 10%, depending on the availability of facilities and trained staff. Treatment of underlying disease is essential. Brain development takes place in the first years of life, a time when severe PEM frequently occurs. There is evidence that intellectual impairment and behavioural abnormalities occur in severely affected children. Physical growth is also impaired. Probably both of these effects can be alleviated if it is possible to maintain a high standard of living with a good diet and freedom from infection over a long period.

Prevention

Prevention of PEM depends not only on adequate nutrients being available but also on education of both governments and individuals in the importance of good nutrition and immunization (Box 5.3). Short-term programmes are useful for acute shortages of food, but long-term programmes involving improved agriculture are equally necessary. Bad feeding practices and infections are more prevalent than actual shortage of food in many areas of the world. However, good surveillance is necessary to avoid periods of famine.

Box 5.3 Prevention of protein–energy malnutrition – GOBIF (a WHO priority programme)

- **G**rowth monitoring: The WHO has a simple growth chart that the mother keeps
- **O**ral rehydration, particularly for diarrhoea
- **B**reast-feeding supplemented by food after 6 months
- **I**mmunization: against measles, tetanus, pertussis, diphtheria, polio and tuberculosis
- **F**amily planning

Food supplements (and additional vitamins) should be given to 'at-risk' groups by adding high-energy food (e.g. milk powder, meat concentrates) to the diet. Pregnancy and lactation are times of high energy requirement and supplements have been shown to be beneficial.

FURTHER READING

Collins S, Dent N, Binns P et al. Management of severe acute malnutrition in children. *Lancet* 2006; **368**: 1992–2000.

Collins S, Sadler K, Dent N et al. Key issues in the success of community-based management of severe malnutrition. *Food and Nutrition Bulletin* 2006; **27**: S49–82.

Elia M (Chairman and editor). *The 'MUST' report. Nutritional screening for adults: a multidisciplinary responsibility. Development and use of the 'Malnutrition Universal Screening Tool' ('MUST') for adults*. A report by the Malnutrition Advisory Group of the British Association for Parenteral and Enteral Nutrition, 2003, p. 127.

Stratton RJ, Elia M. A review of reviews: a new look at the evidence for oral nutritional supplements in clinical practice. *Clinical Nutrition* Suppl 2007; **2**: 5–23.

VITAMINS

Deficiencies due to inadequate intake associated with PEM (Table 5.8) are commonly seen in the developing countries. This is not, however, invariable. For example, vitamin A deficiency is not seen in Jamaica, but is common in PEM in Hyderabad, India. In the West, deficiency of vitamins is rare

Table 5.8 Fat-soluble and water-soluble vitamins: reference nutrient intake (RNI) and lower reference nutrient intake (LRNI)

Vitamin	RNI/day (sufficient)	LRNI/day (insufficient)	Major clinical features of deficiency
Fat-soluble			
A (retinol)	700 μg	300 μg	Xerophthalmia, night blindness, keratomalacia, follicular hyperkeratosis
D (cholecalciferol)	No dietary intake required	10 μg (living indoors)	Rickets, osteomalacia
K	1 μg/kg bodyweight		Coagulation defects
E (α-tocopherol)	10 mg*§		Neurological disorders, e.g. ataxia
Water-soluble			
B_1 (thiamin)	0.4 mg per 1000 kcal†	0.23 mg per 1000 kcal†	Beriberi, Wernicke–Korsakoff syndrome
B_2 (riboflavin)	1.3 mg	0.8 mg	Angular stomatitis
Niacin	6.6 mg per 1000 kcal	4.4 mg per 1000 kcal	Pellagra
B_6 (pyridoxine)	15 μg per g of dietary protein	11 μg per g of dietary protein	Polyneuropathy
B_{12} (cobalamin)	1.5 μg	1.0 μg	Megaloblastic anaemia, neurological disorders
Folate	200 μg	100 μg	Megaloblastic anaemia
C (ascorbic acid)	40 mg	10 mg	Scurvy

*No official RNI because amount varies depending upon polyunsaturated fatty acid content of diet
†Thiamin requirements are related to energy metabolism
§USA RDA

Table 5.9 Some causes of vitamin deficiency in developed countries

Decreased intake

Alcohol dependency: chiefly B vitamins (e.g. thiamin)

Small bowel disease: chiefly folate, occasionally fat-soluble vitamins

Vegans: vitamin D (if no exposure to sunlight), vitamin B_{12}

Elderly with poor diet: chiefly vitamin D (if no exposure to sunlight), folate

Anorexia from any cause: chiefly folate

Decreased absorption

Ileal disease/resection: only vitamin B_{12}

Liver and biliary tract disease: fat-soluble vitamins

Intestinal bacterial overgrowth: vitamin B_{12}

Oral antibiotics: vitamin K

Miscellaneous

Long-term enteral or parenteral nutrition: usually vitamin supplements are given

Renal disease: vitamin D

Drug antagonists (e.g. methotrexate interfering with folate metabolism)

except in the specific groups shown in Table 5.9. The widespread use of vitamins as 'tonics' is unnecessary and should be discouraged. Toxicity from excess fat-soluble vitamins is occasionally seen.

FAT-SOLUBLE VITAMINS

Vitamin A

Vitamin A (retinol) is part of the family of retinoids which is present in food and the body as esters combined with long-chain fatty acids. The richest food source is liver, but it is also found in milk, butter, cheese, egg yolks and fish oils. Retinol or carotene is added to margarine in the UK and other countries.

Beta-carotene is the main carotenoid found in green vegetables, carrots and other yellow and red fruits. Other carotenoids, lycopene and lutein, are probably of little quantitative importance as dietary precursors of vitamin A.

Beta-carotene is cleaved in the intestinal mucosa by carotene dioxygenase, yielding retinaldehyde which can be reduced to retinol. Between a quarter and a third of dietary vitamin A in the UK is derived from retinoids. Nutritionally, 6 μg of β-carotene is equivalent to 1 μg of preformed retinol; vitamin A activity in the diet is given as retinol equivalents.

Function

Retinol is stored in the liver and is transported in plasma bound to an α-globulin, retinol-binding protein (RBP). Vitamin A has several metabolic roles:

- Retinaldehyde in its *cis* form is found in the opsin proteins in the rods (rhodopsin) and cones (iodopsin) of the retina (p. 1086). Light causes retinaldehyde to change to its *trans* isomer, and this leads to changes in membrane potentials that are transmitted to the brain.
- Retinol and retinoic acid are involved in the control of cell proliferation and differentiation.
- Retinyl phosphate is a cofactor in the synthesis of most glycoproteins containing mannose.

Vitamin A deficiency

World-wide, vitamin A deficiency and xerophthalmia (see below) is the major cause of blindness in young children despite intensive preventative programmes.

Xerophthalmia has been classified by the WHO (Table 5.10). Impaired adaptation followed by night blindness is the first effect. There is dryness and thickening of the conjunctiva and the cornea (xerophthalmia occurs as a result of keratinization). Bitot's spots – white plaques of keratinized epithelial cells – are found on the conjunctiva of young children with vitamin A deficiency. These spots can, however, be seen without vitamin A deficiency, possibly caused by exposure. Corneal softening, ulceration and dissolution (keratomalacia)

Table 5.10 Classification of xerophthalmia by ocular signs

Night blindness (XN)
Conjunctival xerosis (XIA)
Bitot's spot (X2)
Corneal xerosis (X2)
Corneal ulceration/keratomalacia <1/3 corneal surface (X3A)
Corneal ulceration/keratomalacia ≥1/3 corneal surface (X3B)
Corneal scar (XS)
Xerophthalmic fundus (XF)

Reproduced with permission from WHO/UNICEF/IVACG 1988
X, xerophthalmia

eventually occur; superimposed infection is a frequent accompaniment and both lead to blindness.

In PEM, retinol-binding protein along with other proteins is reduced. This suggests vitamin A deficiency, although body stores are not necessarily reduced.

Vitamin A in malnourished children

Vitamin A supplementation (single oral dose of 60 mg retinol palmitate) appears to improve morbidity and mortality from measles. It has also been suggested that a similar supplementation reduces morbidity and/or mortality from diarrhoeal diseases and respiratory infections and improves growth. Despite low circulating concentrations of vitamin A in HIV-infected individuals, supplementation of HIV-infected pregnant women does not appear to reduce the risk of mother-to-child transmission of HIV.

Diagnosis

In parts of the world where deficiency is common, diagnosis is made on the basis of the clinical features, and deficiency should always be suspected if any degree of malnutrition is present. Blood levels of vitamin A will usually be low, but the best guide to the diagnosis is a response to replacement therapy.

Treatment

Urgent treatment with retinol palmitate 30 mg orally should be given on 2 successive days. In the presence of vomiting and diarrhoea, 30 mg of vitamin A is given intramuscularly. Associated malnutrition must be treated, and superadded bacterial infection should be treated with antibiotics. Referral for specialist ophthalmic treatment is necessary in severe cases.

Prevention

Most western diets contain enough dairy products and green vegetables, but vitamin A is added to foodstuffs (e.g. margarine) in some countries. Vitamin A is not destroyed by cooking.

In some developing countries vitamin A supplements are given at the time the child attends for measles vaccination. Food fortification programmes are another approach. Education of the population is necessary and people should be encouraged to grow their own vegetables. In particular, pregnant women and children should be encouraged to eat green vegetables.

Other effects of vitamin A

The effect of β-carotene in cardiovascular and other diseases is discussed below in the section entitled 'Dietary antioxidants' (p. 223). Retinoic acid and some synthetic retinoids are used in dermatology (p. 1247).

Possible adverse effects

- *High intakes of vitamin A*. Chronic ingestion of retinol can cause liver and bone damage, hair loss, double vision, vomiting, headaches and other abnormalities. Single doses of 300 mg in adults or 100 mg in children can be harmful.
- *Retinol is teratogenic*. The incidence of birth defects in infants is high with vitamin A intakes of more than 3 mg a day during pregnancy. In pregnancy, extra vitamin A or consumption of liver is not recommended in the UK. However, β-carotene is not toxic.

Vitamin D

See page 559.

Vitamin K

Vitamin K is found as phylloquinone (vitamin K_1) in green leafy vegetables, dairy products, rape seed and soya bean oils. Intestinal bacteria can synthesize the other major form of vitamin K, menaquinone (vitamin K_2), in the terminal ileum and colon. Vitamin K is absorbed in a similar manner to other fat-soluble substances in the upper small gut. Some menaquinones must also be absorbed as this is the major form found in the human liver.

Function

Vitamin K is a cofactor necessary for the production not only of blood clotting factors (p. 432), but also for proteins necessary in the formation of bone.

Vitamin K is a cofactor for the post-translational carboxylation of specific protein-bound glutamate residues in γ-carboxyglutamate (Gla). Gla residues bind calcium ions to phospholipid templates, and this action on factors II, VII, IX and X, and on proteins C and S, is necessary for coagulation to take place.

Bone osteoblasts contain three vitamin K-dependent proteins, osteocalcin, matrix Gla protein and protein S, which have a role in bone matrix formation. Osteocalcin contains three Gla residues which bind tightly to the hydroxyapatite matrix depending on the degree of carboxylation; this leads to bone mineralization. There is, however, no convincing evidence that vitamin K deficiency or antagonism affects bone other than rapidly growing bone.

Vitamin K deficiency

Vitamin K deficiency results in inadequate synthesis of clotting factors (p. 440), which leads to an increase in the prothrombin time and haemorrhage. Deficiency occurs in the following circumstances:

The newborn

Deficiency occurs in the newborn owing to:

- poor placental transfer of vitamin K
- little vitamin K in breast milk

- no hepatic stores of menaquinone (no intestinal bacteria in the neonate).

Deficiency leads to a haemorrhagic disease of the newborn, which can be prevented by prophylactic vitamin K. Vitamin K (phytomenadione 1 mg, i.m.) is given to all neonates after risks have been discussed with parents and consent obtained.

Cholestatic jaundice

When bile flow into the intestine is interrupted, malabsorption of vitamin K occurs as no bile salts are available to facilitate absorption and the prothrombin time increases. This can be corrected by giving 10 mg of phytomenadione intramuscularly. (Note that an increased prothrombin time because of liver disease does not respond to vitamin K injection, there being no shortage of vitamin K, just bad liver function.) In patients with chronic cholestasis (e.g. primary biliary cirrhosis) oral therapy using a water-soluble preparation, menadiol sodium phosphate 10 mg daily, is used.

Concomitant vitamin K antagonists

Oral anticoagulants antagonize vitamin K (p. 445). Antibacterial drugs also interfere with the bacterial synthesis of vitamin K.

Vitamin E

Vitamin E includes eight naturally occurring compounds divided into tocopherols and tocotrienoles. The most active compound and the most widely available in food is the natural isomer d- (or RRR) α-tocopherol, which accounts for 90% of vitamin E in the human body. Vegetables and seed oils, including soya bean, saffron, sunflower, cereals and nuts, are the main sources. Animal products are poor sources of the vitamin.

Vitamin E is absorbed with fat, transported in the blood largely in low-density lipoproteins (LDL).

An individual's vitamin E requirement depends on the intake of polyunsaturated fatty acids (PUFAs). Since this varies widely, no daily requirement is given in the UK. The requirement stated in the USA is approximately 7–10 mg per day, but average diets contain much more than this. If PUFAs are taken in large amounts, more vitamin E is required.

Function

The biological activity of vitamin E results principally from its antioxidant properties. In biological membranes it contributes to membrane stability. It protects cellular structures against damage from a number of highly reactive oxygen species, including hydrogen peroxide, superoxide and other oxygen radicals. Vitamin E may also affect cell proliferation and growth.

Vitamin E deficiency

The first deficiency to be demonstrated was a haemolytic anaemia described in premature infants. Infant formulations now contain vitamin E.

Deficiency is seen only in children with abetalipoproteinaemia (p. 284) and in patients on long-term parenteral nutrition. The severe neurological deficit (gross ataxia) can be prevented by vitamin E injections.

Plasma or serum levels of α-tocopherol can be measured and should be corrected for the level of plasma lipids by expressing the value as milligrams per milligram of plasma lipid.

Epidemiological data and clinical trials

Animals fed an atherogenic diet supplemented with α-tocopherol develop many fewer new atheromatous lesions than those fed an atherogenic diet alone; there may be regression of existing lesions.

There is also evidence for vitamin E intake and blood α-tocopherol levels as an independent risk factor for the development of ischaemic heart disease (IHD) in healthy, well-nourished individuals eating a western diet. This has been shown in comparisons of different communities in the WHO 'MONICA' observational study.

Randomized trials involving vitamin E supplementation have produced conflicting results, possibly due to factors such as short duration of treatment, use of suboptimal doses or without the concurrent administration of vitamin C, although there may be benefit in certain subgroups.

There are very few trials to assess the role of vitamin E in prevention of peripheral vascular disease and for cancer prevention, although there may be some benefit in patients with Alzheimer's disease.

WATER-SOLUBLE VITAMINS

Water-soluble vitamins are non-toxic and relatively cheap and can therefore be given in large amounts if a deficiency is possible. The daily requirements of water-soluble vitamins are given in Table 5.8.

Thiamin (vitamin B_1)

Thiamin consists of pyrimidine and thiazole rings. The alcohol side-chain is esterified with one, two or three phosphates (Fig. 5.7).

Function

Thiamin diphosphate, often called thiamin pyrophosphate (TPP), is an essential cofactor, particularly in carbohydrate metabolism.

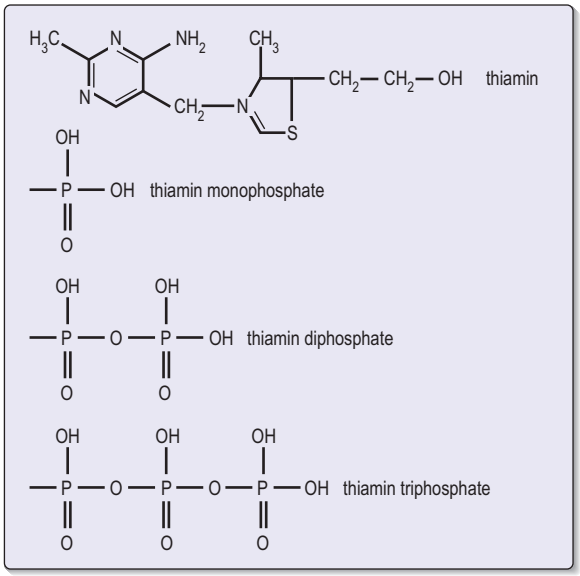


Fig. 5.7 **Thiamin.**

TPP is involved in the oxidative decarboxylation of acetyl CoA in mitochondria. In formation of acetyl CoA (from pyruvate) and in the Krebs cycle, TPP is the key enzyme for the decarboxylation of α-ketoglutarate to succinyl CoA. TPP is also the cofactor for transketolase, a key enzyme in the hexose monophosphate shunt.

Thiamin is found in many foodstuffs, including cereals, grains, beans, nuts, as well as pork and duck. It is often added to food (e.g. in cereals) in developed countries. The dietary requirement (see Table 5.8) depends on energy intake, more being required if the diet is high in carbohydrates.

Following absorption, thiamin is found in all body tissues, the majority being in the liver. Body stores are small and signs of deficiency quickly develop with inadequate intake.

There is no evidence that a high oral intake is dangerous, but ataxia has been reported after high parenteral therapy.

Thiamin deficiency

Thiamin deficiency is seen:

- as *beriberi*, where the only food consumed is polished rice
- in *chronic alcohol-dependent patients* who are consuming virtually no food at all
- in *starved patients* (e.g. with carcinoma of the stomach), and in severe prolonged hyperemesis gravidarum, anorexia nervosa and prolonged total starvation in healthy subjects (e.g. fasts for political reasons). It can also occur in patients given parenteral nutrition with little or no thiamine as large doses of glucose increase requirements of thiamine and can precipitate deficiency, e.g. during refeeding.

Beriberi

This is now confined to the poorest areas of South East Asia. It can be prevented by eating undermilled or par-boiled rice, or by fortification of rice with thiamine. The prevention of beriberi needs a general increase in overall food consumption so that the staple diet is varied and contains legumes and pulses, which contain a large amount of thiamin. There are two main clinical types of beriberi which, surprisingly, only rarely occur together.

Dry beriberi usually presents insidiously with a symmetrical polyneuropathy. The initial symptoms are heaviness and stiffness of the legs, followed by weakness, numbness, and pins and needles. The ankle jerk reflexes are lost and eventually all the signs of polyneuropathy that may involve the trunk and arms are found (p. 1172). Cerebral involvement occurs, producing the picture of the Wernicke–Korsakoff syndrome (p. 1174). In endemic areas, mild symptoms and signs may be present for years without unduly affecting the patient.

Wet beriberi causes oedema. Initially this is of the legs, but it can extend to involve the whole body, with ascites and pleural effusions. The peripheral oedema may mask the accompanying features of dry beriberi.

Thiamin deficiency impairs pyruvate dehydrogenase with accumulation of lactate and pyruvate, producing peripheral vasodilatation and eventually oedema. The heart muscle is also affected and heart failure occurs, causing a further increase in the oedema. Initially there are warm extremities, a full, fast, bounding pulse and a raised venous pressure ('high-output state'), but eventually heart failure advances and a poor cardiac output ensues. The electrocardiogram may show conduction defects.

Infantile beriberi occurs, usually acutely, in breast-fed babies at approximately 3 months of age. The mothers show no signs of thiamin deficiency but presumably their body stores must be virtually nil. The infant becomes anorexic, develops oedema and has some degree of aphonia. Tachycardia and tachypnoea develop and, unless treatment is instituted, death occurs quickly.

Diagnosis

In endemic areas the diagnosis of beriberi should always be suspected and if in doubt treatment with thiamine should be instituted. A rapid disappearance of oedema after thiamine (50 mg i.m.) is diagnostic. Other causes of oedema must be considered (e.g. renal or liver disease), and the polyneuropathy is indistinguishable from that due to other causes. The diagnosis is confirmed by measurement of the circulating thiamin concentration or transketolase activity in red cells using fresh heparinized blood. This enzyme is dependent on TPP. The assay is performed with and without added TPP; an increase in activity of 25% with TPP indicates deficiency.

Treatment

Thiamine 50 mg i.m. is given for 3 days, followed by 50 mg of thiamine daily by mouth. The response in wet beriberi occurs in hours, giving dramatic improvement, but in dry beriberi improvement is often slow to occur. In most cases all the B vitamins are given because of multiple deficiency. Infantile beriberi is treated by giving thiamine to the mother, which is then passed on to the infant via the breast milk.

Thiamin deficiency in patients with alcohol dependence or acute illness

In the western world alcohol-dependent patients and those with severe acute illness receiving high carbohydrate infusions without vitamins are the only major groups to suffer from thiamin deficiency. Rarely they develop wet beriberi, which must be distinguished from alcoholic cardiomyopathy. More usually, however, thiamin deficiency presents with polyneuropathy or with the Wernicke–Korsakoff syndrome.

This syndrome, which consists of dementia, ataxia, varying ophthalmoplegia and nystagmus (see p. 1174), presents acutely and should be suspected in all heavy drinkers. If treated promptly it is reversible; if left it becomes irreversible. It is a major cause of dementia in the USA.

Urgent treatment with thiamine 250 mg i.m. or i.v. infusion once daily is given for 3 days, often combined with other B-complex vitamins. Anaphylaxis can occur. Thiamine must always be given before any intravenous glucose infusion.

Riboflavin

Riboflavin is widely distributed throughout all plant and animal cells. Good sources are dairy products, offal and leafy vegetables. Riboflavin is not destroyed appreciably by cooking, but is destroyed by sunlight. Riboflavin is a flavoprotein that is a cofactor for many oxidative reactions in the cell.

There is no definite deficiency, although many communities have low dietary intakes. Studies in volunteers taking a low riboflavin diet have produced:

- angular stomatitis or cheilosis (fissuring at the corners of the mouth)
- a red, inflamed tongue
- seborrhoeic dermatitis, particularly involving the face (around the nose) and the scrotum or vulva.

Conjunctivitis with vascularization of the cornea and opacity of the lens has also been described. It is probable, however, that many of the above features are due to multiple deficiencies rather than the riboflavin itself.

Riboflavin 5 mg daily can be tried for the above conditions, usually given as the vitamin B complex.

Niacin

This is the generic name for the two chemical forms, nicotinic acid and nicotinamide, the latter being found in the two pyridine nucleotides, nicotinamide adenine dinucleotide (NAD) and nicotinamide adenine dinucleotide phosphate (NADP). Both act as hydrogen acceptors in many oxidative reactions, and in their reduced forms (NADH and NADPH) act as hydrogen donors in reductive reactions. Many oxidative steps in the production of energy require NAD, and NADP is equally necessary in the hexose monophosphate shunt (p. 413) for the generation of NADPH, which is necessary for fatty-acid synthesis.

Niacin is found in many foodstuffs, including plants, meat (particularly offal) and fish. Niacin is lost by removing bran from cereals but is added to processed cereals and white bread in many countries.

Niacin can be synthesized in humans from tryptophan, 60 mg of tryptophan being converted to 1 mg of niacin. The amount of niacin in food is given as the 'niacin equivalent', which is equal to the amount of niacin plus one-sixtieth of the tryptophan content. Eggs and cheese contain tryptophan.

Kynureninase and kynurenine hydroxylase, key enzymes in the conversion of tryptophan to nicotinic acid, are both B_6 and riboflavin dependent, and deficiency of these B vitamins can also produce pellagra.

Pellagra

This is rare and is found in people who eat virtually only maize, for example in parts of Africa. Maize contains niacin in the form of niacytin, which is biologically unavailable, and has a low content of tryptophan. In central America, pellagra has always been rare because maize (for the cooking of tortillas) is soaked overnight in calcium hydroxide, which releases niacin. Many of the features of pellagra can be explained purely by niacin deficiency; but some are probably due to multiple deficiencies, including deficiencies of proteins and of other vitamins.

Clinical features

The classical features are of dermatitis (Fig. 5.8), diarrhoea and dementia. Although this is an easily remembered triad, not all features are always present and the mental changes are not a true dementia.

- *Dermatitis*. In the areas of skin exposed to sunlight, initially there is redness followed by cracks with occasional ulceration. Chronic thickening, dryness and pigmentation develop. The lesions are always symmetrical and often affect the dorsal surfaces of the hands. The perianal skin and vulva are frequently involved. Casal's necklace or collar is the term given to the skin lesion around the neck, which is confined to this area by the clothes worn.
- *Diarrhoea*. This is often a feature but constipation is occasionally seen. Other gastrointestinal manifestations include a painful, red, raw tongue, glossitis and angular stomatitis. Recurring mouth infections occur.
- *Dementia*. This occurs in chronic disease. In milder cases there are symptoms of depression, apathy and sometimes thought disorders. Tremor and an encephalopathy frequently occur. Hallucinations and acute psychosis are also seen with more severe cases.

Fig. 5.8 **Pellagra.** A child with pellagra showing chronic thickening and pigmentation of the skin, particularly on the backs of hands. Courtesy of Professor David Warrell.

Pellagra may also occur in the following circumstances:

- Isoniazid therapy can lead to a deficiency of vitamin B_6, which is needed for the synthesis of nicotinamide from tryptophan. Vitamin B_6 is now given concomitantly with isoniazid.
- In Hartnup's disease, a rare inborn error, in which basic amino acids including tryptophan are not absorbed by the gut. There is also loss of this amino acid in the urine.
- In generalized malabsorption (rare).
- In alcohol-dependent patients who eat little.
- Very low protein diets given for renal disease or taken as a food fad.
- In the carcinoid syndrome and phaeochromocytomas, tryptophan metabolism is diverted away from the formation of nicotinamide to form amines.

Diagnosis and treatment

In endemic areas this is based on the clinical features, remembering that other vitamin deficiencies can produce similar changes (e.g. angular stomatitis). Nicotinamide (approximately 300 mg daily by mouth) with a maintenance dose of 50 mg daily is given with dramatic improvement in the skin and diarrhoea. Mostly, however, vitamin B complex is given, as other deficiencies are often present.

An increase in the protein content of the diet and treatment of malnutrition and other vitamin deficiencies is essential.

Vitamin B_6

Vitamin B_6 exists as pyridoxine, pyridoxal and pyridoxamine, and is found widely in plant and animal foodstuffs. Pyridoxal phosphate is a cofactor in the metabolism of many amino acids. Dietary deficiency is extremely rare. Some drugs (e.g. isoniazid, hydralazine and penicillamine) interact with pyridoxal phosphate, producing B_6 deficiency. The polyneuropathy occurring after isoniazid usually responds to vitamin B_6.

Sideroblastic anaemia may respond to vitamin B_6 (see p. 397).

A polyneuropathy has occurred after high doses (>200 mg) given over many months. Vitamin B_6 is used for premenstrual tension: a daily dose of 10 mg should not be exceeded.

Biotin and pantothenic acid

Biotin is involved in a number of carboxylase reactions. It occurs in many foodstuffs and the dietary requirement is small. Deficiency is extremely rare and is confined to a few people who consume raw eggs, which contain an antagonist (avidin) to biotin. It has also been reported in patients receiving long-term parenteral nutrition without adequate amounts of biotin. It causes a dermatitis that responds to biotin replacements.

Pantothenic acid is widely distributed in all foods and deficiency in humans has not been described.

Vitamin C

Ascorbic acid is a powerful reducing agent controlling the redox potential within cells. It is involved in the hydroxylation of proline to hydroxyproline, which is necessary for the formation of collagen. The failure of this biochemical pathway in vitamin C deficiency accounts for virtually all of the clinical effects seen.

Humans, along with a few other animals (e.g. primates and the guinea-pig), are unusual in not being able to synthesize ascorbic acid from glucose.

Vitamin C is present in all fresh fruit and vegetables. Unfortunately, ascorbic acid is easily leached out of vegetables when they are placed in water and it is also oxidized to dehydro-ascorbic acid during cooking or exposure to copper or alkalis. Potatoes are a good source as many people eat a lot of them, but vitamin C is lost during storage.

It has been suggested that ascorbic acid in high dosage (1–2 g daily) will prevent the common cold. While there is some scientific support for this, clinical trials have shown no significant effect. Vitamin C supplements have also been advocated to prevent atherosclerosis and cancer, but again a clear benefit has not been demonstrated.

Vitamin C deficiency is seen mainly in infants fed boiled milk and in the elderly and single people who do not eat vegetables. In the UK it is also seen in Asians eating only rice and chapattis, and in food faddists.

Scurvy

In adults the early symptoms of vitamin C deficiency may be non-specific, with weakness and muscle pain. Other features are shown in Table 5.11. In infantile scurvy there is irritability, painful legs, anaemia and characteristic subperiosteal haemorrhages, particularly into the ends of long bones.

Table 5.11 Clinical features of vitamin C deficiency (scurvy)

Keratosis of hair follicles with 'corkscrew' hair
Perifollicular haemorrhages
Swollen, spongy gums with bleeding and superadded infection, loosening of teeth
Spontaneous bruising
Spontaneous haemorrhage
Anaemia
Failure of wound healing

Diagnosis

The anaemia is usually hypochromic but occasionally a normochromic or megaloblastic anaemia is seen. The type of anaemia depends on whether iron deficiency (owing to decreased absorption or loss due to haemorrhage) or folate deficiency (folate being largely found in green vegetables) is present.

Plasma ascorbic acid is very low in obvious deficiency and a vitamin C level of less than 11 μmol/L (0.2 mg per 100 mL) indicates vitamin C deficiency. The leucocyte–platelet layer (buffy coat) of centrifuged blood corresponds to vitamin C concentrations in other tissues. The normal level of leucocyte ascorbate is 1.1–2.8 pmol per 10^6 cells.

Treatment

Initially the patient is given 250 mg of ascorbic acid daily and encouraged to eat fresh fruit and vegetables. Subsequently, 40 mg daily will maintain a normal exchangeable body pool of about 900 mg (5.1 mmol).

Prevention

Orange juice should be given to bottle-fed infants. The intake of breast-fed infants depends on the mother's diet. In the elderly, eating adequate fruit and vegetables is the best way to avoid scurvy. Careful surveillance of the elderly, particularly those who live alone, is necessary. Ascorbic acid supplements should only be necessary occasionally.

Vitamin B_{12} and folate

These are dealt with on page 398 and daily requirements are shown in Table 5.8.

Folate. In many developed countries, up to 15% of the population have a partial deficiency of 5,10-methylene tetrahydrofolate reductase, a key folate-metabolizing enzyme. This is due to a point mutation and is associated with an increase in neural tube defects and hyperhomocysteinaemia, which may lead to cardiovascular damage. Autoantibodies against folate receptors have been found in serum from women who have had a pregnancy complicated by neural tube defects. However, the role of this in the pathogenesis is unclear.

In the USA and some other countries, enriched cereals are fortified with 1.4 mg/kg grain of folic acid to increase daily requirements.

Dietary antioxidants

Free radicals are generated during inflammatory processes, radiotherapy, smoking, and during the course of a wide range of diseases. They may cause uncontrolled damage of

multiple cellular components, the most sensitive of which are unsaturated lipids, proteins and DNA, and they also disrupt the normal replication process. They have been implicated as a cause of a wide range of diseases, including malignant, acute inflammatory and traumatic diseases, cardiovascular disease, neurodegenerative conditions such as Alzheimer's disease, senile macular degeneration, and cataract. The defence against uncontrolled damage by free radicals is provided by antioxidant enzymes (e.g. catalase, superoxide dismutase) and antioxidants, which may be endogenous (e.g. glutathione) or exogenous (e.g. vitamins C and E, carotenoids). A possible causal link between lack of antioxidants and cardiovascular disease has emerged from epidemiological studies although several RCTs have not confirmed this.

Epidemiological studies

Dietary intake

- A high intake of fruits and vegetables has been linked to reduced risk of heart disease, cerebrovascular disease and total cardiovascular morbidity and mortality.
- A high intake of nuts (rich in vitamin E) and dietary components, e.g. red wine, onions, apples (rich in flavonoids), which are strong scavengers of free radicals, has also been linked to reduced risk of cardiovascular disease.
- The seasonal variation in cardiovascular disease, which is higher in winter, has been related to decreased intake of fresh fruit and vegetables in winter.
- The decline in cardiovascular disease in the USA since the 1950s has been associated with a simultaneous increase in the intake of fresh fruit and vegetables.

Status of antioxidant nutrients

The level of antioxidant nutrients in the circulation has been reported to be inversely related to cardiovascular morbidity and mortality, extent of atherosclerosis assessed by intra-arterial ultrasound, and clinical signs of ischaemic heart disease. The tissue content of lycopene, a marker of vegetable intake, has been reported to be low in patients with myocardial infarction.

Antioxidants, especially vitamin E, have been shown to prevent the initiation and progression of atherosclerotic disease in animals. They also reduce the oxidation of low-density lipoprotein (LDL) in the arterial wall in vitro. Oxidation of LDL is an initial event in the atherosclerotic process (p. 744). However these epidemiological studies show an association rather than a causal link and RCTs comparing the antioxidant against a control group are necessary.

Randomized controlled trials (RCT) (see also p. 929)

The results of such trials have been formally evaluated through a series of systematic reviews and meta-analyses.

- *For primary or secondary prevention* of cardiovascular disease, intervention with β-carotene, α-tocopherol (vitamin E) and ascorbic acid (vitamin C) has demonstrated no significant benefit.
- Vitamin E or β-carotene given in, e.g. stroke and fatal and non-fatal myocardial infarction, has also not yielded benefits.
- There is a report of increased risk of intracerebral and subarachnoid haemorrhage in healthy individuals receiving carotene and α-tocopherol.
- A meta-analysis has shown a small but significant overall increased risk of cardiovascular death and all-cause mortality in individuals treated with β-carotene (compared to the control group).
- An increased risk of developing lung cancer by administering large doses of β-carotene to subjects with a history of heavy smoking.
- Although administration of antioxidant nutrients has been proposed in a wide range of acute (e.g. critical illness, pancreatitis) and chronic diseases, the evidence base from randomized RCTs is generally not strong.
- In some cases improvement in indices of free radical damage had been demonstrated (e.g. in acute inflammatory conditions), but with little evidence of clinical benefit.

Epidemiological studies are also confounded by other associated variables, e.g. eating a low-fat diet or undertaking more exercise. The latter may be more valuable in the causal pathway than the intake of antioxidants. Diets rich in fresh fruit and vegetables also contain a range of antioxidants that were not tested in the clinical trials. Therefore, the results of large-scale RCTs using various combinations and doses of antioxidant nutrients are currently awaited. In the meantime, the policy of encouraging 'healthy' behaviour, which includes increased physical activity and a varied diet rich in fresh fruit and vegetables, and nuts, is still generally recommended both for the population as a whole and for those at risk of cardiovascular disease.

Homocysteine, cardiovascular disease and B vitamins

The circulating concentration of the amino acid homocysteine is an independent risk factor for cardiovascular disease (p. 747). A high concentration is related to ischaemic heart disease, stroke, thrombosis, pulmonary embolism, coronary artery stenosis, and heart failure. The strength of the association is similar to smoking or hyperlipidaemia.

Proposed mechanisms, based on experimental evidence, by which homocysteine detrimentally affects vascular function, include:

- the direct damaging effects of homocysteine on endothelial cells of blood vessels
- an increase in blood vessel stiffness
- an increase in blood coagulation.

Homocysteine is not found in food, but results from metabolism within the body which depends on folic acid, vitamin B_{12} and pyridoxine (vitamin B_6) (Fig. 5.9). Deficiency of one or more of these vitamins is common in the elderly, which would increase the concentration of homocysteine. If an elevated homocysteine concentration was causally linked to cardiovascular disease then it should be possible to lower the risk by administering one or more of these vitamins to lower the homocysteine concentration. Some recent studies, however, suggest that lowering homocysteine concentrations in this way does not reduce the risk of cardiovascular disease.

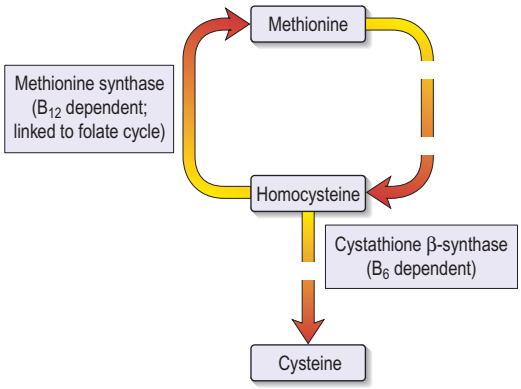


Fig. 5.9 **Homocysteine metabolism.**

FURTHER READING

Institute of Medicine. *Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B6, Folate, Vitamin B12, Pantothenic Acid, Biotin, and Choline.* Washington DC: The National Academies Press, 2000.

Institute of Medicine. *Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Protein, and Amino Acids.* Washington, DC: The National Academies Press, 2005.

Loscalzo J. Homocysteine trials – clear outcomes for complex reasons. *New England Journal of Medicine* 2006; **354**: 1629–1632.

Vivekananthan DP, Penn MS, Sapp SK et al. Use of antioxidant vitamins for the prevention of cardiovascular disease: meta-analysis of randomised trials. *Lancet* 2003; **361**: 2017–2023.

MINERALS

A number of minerals have been shown to be essential in animals, and an increasing number of deficiency syndromes are becoming recognized in humans. Long-term total parenteral nutrition allowed trace element deficiency to be studied in controlled conditions; now trace elements are always added to long-term parenteral nutrition regimens. It is highly probable (but difficult to study because of multiple deficiencies) that trace-element deficiency is also a frequent accompaniment of all PEM states. Sodium (RNI 70 mmol/day), potassium (RNI 90 mmol/day), magnesium (RNI 12.3 mmol/day for men and 10.9 for women) and chloride are discussed in Chapter 12.

Iron (see also p. 393)

The daily RNI for men is 160 μmol (8.7 mg) and for women 260 μmol (14.8 mg). Iron deficiency is common world-wide, affecting both developing and developed countries. It is particularly prevalent in women of reproductive age. Dietary iron overload is seen in South African men who cook and brew in iron pots.

Copper

The daily RNI of copper is 1.2 mg (19 μmol). Shellfish, legumes, cereals and nuts are good dietary sources.

Deficiency

Menkes' kinky hair syndrome is a rare condition caused by malabsorption of copper. The Menkes' disease gene (*ATP7A*) encodes a copper-transporting ATPase and has a homology to the gene in Wilson's disease. Infants with this sex-linked recessive abnormality develop growth failure, mental retardation, bone lesions and brittle hair. Anaemia and neutropenia also occur. This condition, which serves as a model for copper deficiency, supports the idea that some of the clinical features seen in PEM are due to copper deficiency. Breast and cow's milk are low in copper, and supplementation is occasionally necessary when first treating PEM.

Copper toxicity

This occurs in Wilson's disease; see page 357.

Zinc

The daily RNI of zinc is 9.5 mg (145 μmol) for men, 7 mg (110 μmol) for women and it is widely available in food. Zinc is involved in many metabolic pathways, often acting as a coenzyme; it is essential for the synthesis of RNA and DNA.

Deficiency

Acrodermatitis enteropathica is an inherited disorder caused by malabsorption of zinc. Infants develop growth retardation, severe diarrhoea, hair loss and associated *Candida* and bacterial infections. This condition provides a model for zinc deficiency. Zinc supplementation results in a complete cure. Deficiency probably also plays a role in PEM and in many diseases in children in the developing world. Zinc supplementation has been shown to be of some benefit in, for example, the prevention of diarrhoeal diseases and acute respiratory infections; it also improves growth.

Zinc levels have also been shown to be low in some patients with malabsorption or skin disease, and in patients with AIDS, but the exact role of zinc in these situations is disputed. Zinc has low toxicity, but high zinc levels from water stored in galvanized containers interfere with iron and copper metabolism. Wound healing is impaired with moderate zinc deficiency and is improved by zinc supplements. Impaired taste and smell, hair loss and night blindness are also features of severe zinc deficiency.

Iodine

The daily RNI of iodine is 140 μg (1.1 μmol) for men and women, and it is found in milk, meat and seafoods.

It exists in foodstuffs as inorganic iodides which are efficiently absorbed. Iodine is a constituent of the thyroid hormones (p. 983).

Deficiency

Many areas throughout the world lack iodine in the soil, and so iodine deficiency, which impairs brain development, is a WHO priority. Two billion people (one-third children) worldwide have insufficient iodine intake. Endemic goitre occurs in remote areas where the daily intake is below 70 μg, and in those parts 1–5% of babies are born with cretinism. In these areas, iodized oil should be given intramuscularly to all reproductive women every 3–5 years. Salt iodization is now used in many contries and is a simple, cost-effective way to prevent deficiency.

www.studentconsult.com
More online

Fluoride

In areas where the level of fluoride in drinking water is less than 1 p.p.m. (0.7–1.2 mg/L), dental caries is relatively more prevalent. Fluoridation of the water provides 1–2 mg daily, resulting in a reduction of about 50% of tooth decay in children. There is little fluoride in food except for seafish and tea, the latter providing 70% of the daily intake. Fluoride-containing toothpaste may add up to 2 mg a day.

Excessive fluoride intake in areas where the water fluoride level is above 3 mg/L can result in fluorosis, in which there is infiltration into the enamel of the teeth, producing pitting and discoloration.

Selenium

Clinical deficiency of selenium is rare except in areas of China where Keshan disease, a selenium-responsive cardiomyopathy, occurs. Selenium deficiency may also cause a myopathy. Toxicity has been described with very high intakes.

Calcium (see also p. 558)

In the UK, the daily RNI of calcium is 700 mg (17.5 mmol), but substantially higher values are recommended in the USA. It is found in many foodstuffs, with two-thirds of the intake coming from milk and milk products, and only 5% from vegetables. In the UK most flour is fortified. Calcium absorption from the gastrointestinal tract is vitamin D-dependent. Ninety-nine per cent of body calcium is in the skeleton.

Increased calcium is required in pregnancy and lactation, when dietary intake must be increased. Calcium deficiency is usually due to vitamin D deficiency.

Phosphate (see also p. 559)

The daily RNI of phosphate is the same as that of calcium, i.e. 17.5 mmol. Phosphates are present in all natural foods, and dietary deficiency has not been described. Patients taking large amounts of aluminium hydroxide can, however, develop phosphate deficiency owing to binding in the gut lumen. It can also be seen in total parenteral nutrition. Symptoms include anorexia, weakness and osteoporosis.

Other trace elements

The possible significance of cadmium, chromium, cobalt, manganese, molybdenum, nickel and vanadium is shown in Table 5.12.

Table 5.12 Other trace elements (see text)

Element	Deficiency
Cadmium	?
Chromium	Glucose intolerance
Cobalt	Anaemia (vitamin B_{12})
Manganese	Skin rash, ? osteoporosis,? mood
Molybdenum	? (case study involving parenteral nutrition)
Nickel	? Animals only
Vanadium	?

FURTHER READING

Institute of Medicine. *Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Silicon, Vanadium, and Zinc*. Washington, DC: IOM, 2001.

Prentice A. What are the dietary requirements for calcium and vitamin D? *Calcified Tissue International* 2002; **70**: 83–88.

NUTRITION AND AGEING

Many animal studies have shown that life expectancy can be extended by restricting food intake. It is, however, not known whether the ageing process in humans can be altered by nutrition.

www.studentconsult.com
Ask the authors

THE AGEING PROCESS

The process of ageing is not well understood. While wear and tear may play a role, it is an insufficient explanation for the causation of ageing. The 'programmed' theories depend on inbuilt biological clocks that regulate lifespan, and involve genes that are responsible for controlling signals that influence various body systems. The 'error' theories involve environmental stressors that induce damage (e.g. mitochondrial DNA damage or cross-linking).

The search for a single cause of ageing, e.g. a single gene defect, has been replaced by the view that ageing is a complex multifactorial process that involves an interaction between genetic, environmental, and stochastic (random damage to essential molecules) causes. The following theories have been suggested:

Molecular theories

- *Gene regulation* – ageing, for example, results from changes in expression of genes that regulate both development and ageing. An insulin-like signalling pathway has been linked to the lifespan of worms, flies and mice (activation of a transcription factor in response to reduced insulin-like signalling prolongs lifespan).
- *Codon restriction* – inadequate mRNA translation resulting from inadequate decoding of codons in mRNA.
- *Error catastrophe* – errors in gene expression result in abnormal proteins.
- *Somatic mutation* – cumulative molecular damage mainly to genetic material.
- *Dysdifferentiation* – cumulative random molecular damage detrimentally affects gene expression.

Mutations in genes encoding Lamin A are found in fibroblasts of elderly people and in progeria syndromes.

Cellular theories

- *Cellular senescence-telomere* – an increase in senescent cells occurs from:
  (a) *Loss of telomeres*, which is known to occur with ageing (with each cell division a small amount of DNA is necessarily lost at each end of the chromosome). Activation of the telomerase enzyme regenerates telomeres, prevents senescence (replicative senescence) and immortalizes cell cultures. Cancer cells are known to activate telomerase. Accelerated telomerase shortening occurs in progerias, such as Werner's syndrome. A recent study has also found that telomere length in leucocytes predicts coronary artery disease in middle-aged men at high risk and these individuals might benefit from statin treatment.

(b) *Damage* due to a variety of other factors, including DNA damage (stress-induced senescence).

- *Free radical* – production of free radicals during oxidative metabolism which damages fat, protein and DNA.
- *Wear and tear* – cumulative damage from normal injury/stress which is unable to repair itself.
- *Apoptosis theory* – programmed cell death due to genetic events.

System theories

These theories involve loss in the function of neuroendocrine or immune systems with consequent age-related physiological changes and an increase in autoimmunity.

Whole body metabolism and energy expenditure theory proposes that there is a fixed limit to the cumulative energy expenditure and metabolism during a lifetime, so that if this limit is reached quickly the lifespan is short. Energy restriction in rodents reduces energy expenditure and prolongs lifespan, but there is a lack of studies in primates or humans.

Evolutionary theories

- *Cumulative mutation* – mutations that accumulate during a lifetime act in older age rather than during the active reproductive period (for which there is evolutionary selection), producing pathology and senescence. The theory was initially based on the observation that Huntington's disease, a dominant lethal mutation which typically manifests itself between 35 and 55 years, allows affected individuals to reproduce.
- *Disposable soma* – the somatic body is maintained to ensure reproductive success, after which it is disposable. Factors that may enhance reproductive success may have detrimental effects on ageing – a possible example being androgen secretion, which may be beneficial to reproduction but potentially detrimental with development of prostatic cancer and cardiovascular disease in later life.

Several of these theories have strong nutritional components. *Disability and dependency* in older humans are at least partly due to poor nutrition, and correction of deficiencies or nutrient imbalances can prevent the decline in function from falling below the disability threshold (Fig. 5.10). In this way some loss of function may be prevented or reversed, especially if other measures, such as physical activity, which increases muscle mass and strength, are undertaken.

Early origins of health and disease in older adults

A low birth weight (and/or length) is associated with reduced height, as well as reduced mass and fat-free mass in adult life. These relationships can occur independently of genetic factors, since the smaller of identical twins becomes a shorter and lighter adult. Relationships have also been reported between growth of the fetus and a variety of diseases and risk factors for disease in adults and older people. These include cardiovascular disease, especially ischaemic heart disease, hypertension, and diabetes, and even obesity and fat distribution. However, the strength of association for some of these conditions is in some cases weak. Animal studies involving dietary modifications (e.g. protein and zinc, even within the normal range) during pregnancy or in early postnatal life have clearly demonstrated effects, such as hypertension. This can persist, not only through the lifetime of the offspring, but also through to their offspring. The extent to which these apply to humans is uncertain, and the mechanisms are poorly understood. Since relationships with cardiovascular disease in old age have also been related to growth in the first few years of life as well as starvation during puberty, it is likely that cumulative environmental stresses, including nutritional, from the time of implantation of the fertilized egg, to fetal and postnatal growth and development, and adult life, summate to produce an overall disease risk (Fig. 5.10).

Nutritional requirements in the elderly

These are qualitatively similar to the requirements of younger adults, but as energy expenditure is less, there is a lower energy requirement. However, maintaining physical activity is required for the overall health of the elderly.

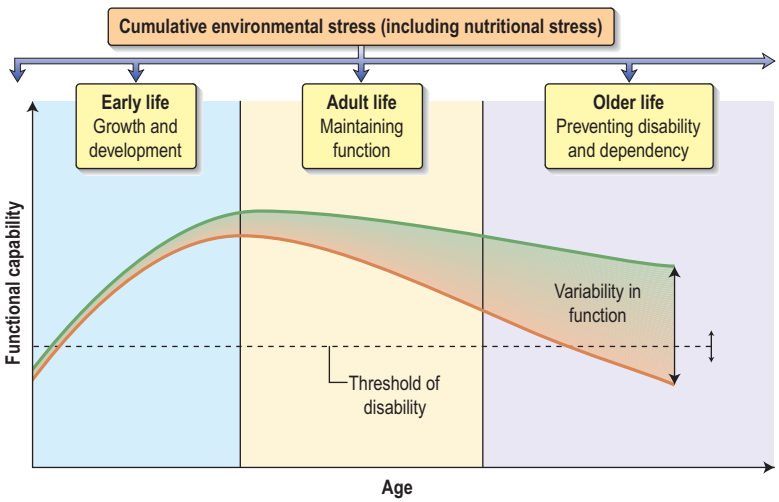


Fig. 5.10 **Nutrition and ageing.** Nutrition is a contributory cause of the variability in function during the lifespan. Appropriate nutrition may improve function, or delay deterioration below the threshold of disability and dependence.

The daily energy requirement of 'elderly people' (aged 60 and above, irrespective of age) has been set to be approximately 1.5 × BMR. The BMR is reduced, owing to a fall in the fat-free mass, from an average of 60 kg to 50 kg in men and from 40 kg to 35 kg in women. In disease, physical activity is usually decreased. The diet should contain approximately the same proportions of nutrients, and essential nutrients are still required. Although disabilities and diseases are common in older people, the RNIs are intended for healthy people without disease. The nutrient requirements for those with disease are less well defined.

Nutritional deficits in the elderly are common and may be due to many factors, such as dental problems, lack of cooking skills (particularly in widowers), depression and lack of motivation. Significant malnourishment in developed countries is usually secondary to social problems or disease. In the elderly who are in institutions, multiple nutrient deficiencies are common. Vitamin D supplements may be required because often these people do not go into the sunlight. Owing to the high prevalence of osteoporosis in elderly people, daily calcium intake of 1–1.5 g/day is often recommended.

FURTHER READING

Artandi SE. Telomeres, telomerase, and human disease. *New England Journal of Medicine* 2006; **355**: 1195–1197.

Cox LS, Faragher RG. From old organisms to new molecules: integrative biology and therapeutic targets in accelerated human ageing. *Cellular and Molecular Life Sciences* 2007; **64**: 2620–2641.

Gluckman PD, Hanson MA, Cooper C. Effect of in utero and early life conditions on adult health and disease. *New England Journal of Medicine* 2008; **359**: 61–73.

Rattan SI. Theories of biological aging: genes, proteins, and free radicals. *Free Radical Research* 2006; **40**: 1230–1238.

Vijg J, Suh Y. Functional genomics of ageing. *Mechanisms of Ageing and Development* 2003; **124**: 3–8.

Vina J, Borras C, Miquel J. Theories of ageing. *IUBMB Life* 2007; **59**: 249–254.

Weinert BT, Timiras PS. Physiology of aging. Invited review: theories of aging. *Journal of Applied Physiology* 2003; **95**: 1706–1716.

OBESITY

Obesity is almost invariable in developed countries and almost all people accumulate some fat as they get older. The World Health Organization acknowledges that obesity (body mass index >30 kg/m^2) is a world-wide problem which also affects many developing countries. Obesity implies an excess storage of fat, and this can most easily be detected by looking at the undressed patient.

Most patients suffer from simple obesity, but in certain conditions obesity is an associated feature (Table 5.13). Even in the latter situation, the intake of calories must have exceeded energy expenditure over a prolonged period of time. Hormonal imbalance is often incriminated in women (e.g. postmenopause or when taking contraceptive pills), but most weight gain in such cases is usually small and due to water retention.

Not all obese people eat more than the average person, but all obviously eat more than they need.

Table 5.13 Conditions in which obesity is an associated feature

Genetic syndromes associated with hypogonadism (e.g. Prader–Willi syndrome, Laurence–Moon–Biedl syndrome)
Hypothyroidism
Cushing's syndrome
Stein–Leventhal syndrome
Drug-induced (e.g. corticosteroids)
Hypothalamic damage (e.g. due to trauma, tumour)

Suggested mechanisms

Genetic and environmental factors

These have always been difficult to separate when studying obesity. However, feeding experiments in both monozygotic and dizygotic twins, reared together or apart, suggest that genetic influences account for 70% of the difference in body mass index (BMI) later in life, and that the childhood environment has little or no influence.

These feeding experiments also showed that weight gain did not occur in all pairs of twins, suggesting that environmental factors operate. Genetic factors have led to the discovery of a putative gene, firstly in the obese (*ob ob*) mouse and now in humans. The *ob* gene was shown to be expressed solely in both white and brown adipose tissue. The *ob* gene is found on chromosome 7 and produces a 16 kDa protein called leptin. In the *ob ob* mouse a mutation in the *ob* gene leads to production of a non-functioning protein. Administration of normal leptin to these obese mice reduces food intake and corrects the obesity. A similar situation has been described in a very rare genetic condition causing obesity in humans, in which leptin is not expressed.

In massively obese subjects, leptin mRNA in subcutaneous adipose tissue is 80% higher than in controls. Plasma levels of leptin are also very high, correlating with the BMI. Weight loss due to food restriction decreases plasma levels of leptin. However, in contrast to the *ob ob* mouse, the leptin structure is normal, and abnormalities in leptin are not the prime cause of human obesity.

Leptin secreted from fat cells was thought to act as a feedback mechanism between the adipose tissue and the brain, acting as a 'lipostat' (adipostat), controlling fat stores by regulating hunger and satiety (see below). However, many other signals are involved. It is interesting that obesity is largely restricted to humans, and animals that are domesticated or in zoos.

Food intake

Many factors related to the home environment, such as finance and the availability of sweets and snacks, will affect food intake. Some individuals eat more during periods of heavy exercise or during pregnancy and are unable to get back to their former eating habits. The increase in obesity in social class 5 can usually be related to the type of food consumed (i.e. food containing sugar and fat). Psychological factors and how food is presented may override complex biochemical interactions.

It has been shown that obese patients eat more than they admit to eating, and over the years a very small daily excess of intake over expenditure can lead to a large accumulation of fat. For example, a 44 kJ (10.5 kcal) daily excess would lead to a 10 kg weight gain over 20 years.

Control of appetite

Appetite is the desire to eat and this usually initiates food intake. Following a meal, satiation occurs. This depends on gastric and duodenal distension and the release of many substances peripherally and centrally.

Following a meal, cholecystokinin (CCK), bombesin, glucagon-like peptide 1 (GLP1), enterostatin, and somatostatin are released from the small intestine, and glucagon and insulin from the pancreas. All of these hormones have been implicated in the control of satiety. Centrally, the hypothalamus – particularly the lateral hypothalamic area, and paraventricular and arcuate nuclei – plays a key role in integrating signals involved in appetite and bodyweight regulation (Fig. 5.11). There are two main pathways in the arcuate nucleus (see Fig. 5.11):

- *The central appetite-stimulating (orexogenic) pathway* in the ventromedial part of the arcuate nucleus, which expresses NPY (neuropeptide Y) and AgRP (agouti-related protein). Animal studies suggest that this pathway also decreases energy expenditure.
- *The central appetite-suppressing (anorexogenic pathway, or leptin–melanocortin pathway)* in the dorsolateral part of the arcuate nucleus, which expresses POMC/CART (pro-opiomelanocortin/cocaine-and-amfetamine-regulated transcript). In this pathway, α-MSH (α-melanocyte-stimulating hormone), formed by cleavage of POMC by PC1 (prohormone convertase), exerts its appetite-suppressing effect via the Mc4R (melanocortin-4 receptors) in areas of the brain that regulate food intake and autonomic activity. Animal studies suggest that this pathway also increases energy expenditure.

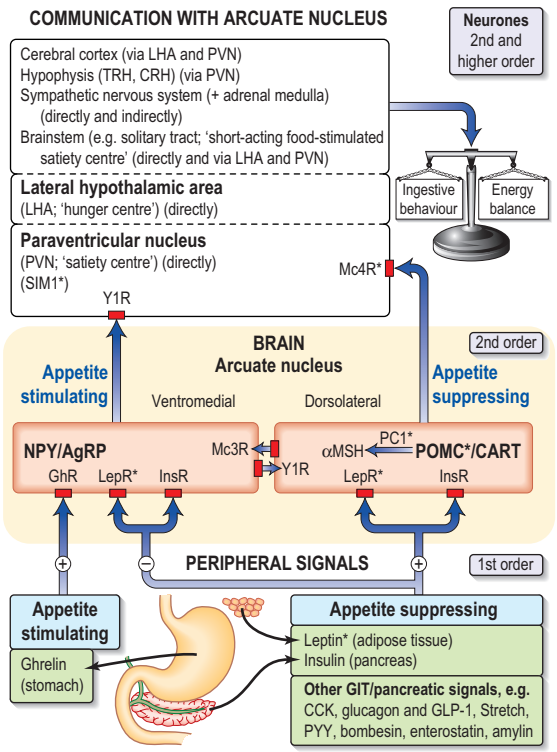


Fig. 5.11 **Peripheral signals and central pathways (central nervous system) involved in the control of food intake. (i)** the stimulatory and suppressive peripheral signals and central pathways are identified (NPY/AgRP (neuropeptide Y/agouti-related protein) involved in the stimulatory pathway) and POMC/CART (pro-opiomelanocortin/cocaine-and-amfetamine-regulated transcript) involved in the suppressive pathway). POMC is converted to melanocortins including α-MSH (α-melanocyte-stimulating hormone) through the action of prohormone convertase (PC1).
(ii) The solid red areas represent receptors for a variety of signals (GhR, ghrelin receptor; LepR, leptin receptor; InsR, insulin receptor; Mc3R, melanocortin 3 receptor; Mc4R, melanocortin 4 receptor; Y1R, Y1 subtype of neuropeptide Y (NPY) receptor).
(iii) Asterisks (*) indicate mutations that have resulted in human obesity. CCK, cholecystokinin; CRH, corticotrophin-releasing hormone; GLP-1, glucagon-like peptide; PVN, paraventricular nucleus; PYY, peptide YY; TRH, thyrotrophin-releasing hormone.

These pathways interact with each other and feed into the lateral hypothalamus, which communicates with other parts of the brain, and influence the autonomic nervous system and ingestive behaviour. These central pathways are in turn influenced by a variety of peripheral signals which can also be classified as appetite stimulating or appetite suppressing.

- **Peripheral appetite-suppressing**: *Leptin and insulin* act centrally to activate the appetite-suppressing pathway (whilst also inhibiting the appetite-stimulating pathway). Since these hormones circulate in proportion to adipose tissue mass, they can be regarded as long-term signals, although they probably also modulate short-term signals (insulin also responds acutely to meal ingestion). Peptide YY (PYY) is produced by the L cells of the large bowel and distal small bowel in proportion to the energy ingested. The release of this rapidly responsive (short-acting) signal begins shortly after food intake, suggesting that the initial response involves neural pathways, before ingested nutrients reach the site of PYY production. PYY is thought to reduce appetite, at least partly through inhibition of the appetite-stimulating pathway (NPY/AgRP-expressing neurones).
- **Peripheral appetite-stimulating**: *Ghrelin* is a 28-amino-acetylated peptide produced by the oxyntic cells of the fundus of the stomach. It is the first known gastrointestinal tract peptide that stimulates appetite by activating the central appetite-stimulating pathway. The circulatory concentration is high before a meal and is reduced rapidly by ingestion of a meal or glucose (cf. peptide YY, which increases after a meal). It may also act as a long-term signal, as its circulating concentration in weight-stable individuals is inversely related to body mass index over a wide range (cf. insulin and leptin which are positively related to BMI (see below)). It is also increased in several situations in which there is a negative energy balance, e.g. long-term exercise, very low-calorie diets, anorexia nervosa and both cancer and cardiac cachexia (an exception is vertical banded gastric bypass surgery, where its concentration is low rather than high).

  The single gene mutations affecting this pathway in humans, e.g. leptin, leptin receptor, *POMC*, *Mc4R*, *PC1*, and *SIM1*, are rare and recessive, with the exception of the Mc4R, which is common and dominant with incomplete penetrance. It appears that the *Mc4R mutation* accounts for 2–6% of human obesity. Affected individuals are obese without disturbances in pituitary

function or resting energy expenditure, although children tend to be tall. However, these mutations are of little significance as obesity is predominantly polygenic in origin (the human obesity gene map has already identified several hundreds of candidate genes).

- Another system, the **endocannabinoid system**, is involved in both central and peripheral regulation of food intake and control of energy balance. There are two receptors: endocannabinoid in the brain and CB_2 in the periphery. CB_1 receptors are located in the cerebral cortex, cerebellum and hippocampus.

The present obesity epidemic is mainly due to behaviour and lifestyle changes (although it may be that individuals with certain genes are affected more than others) with a trebling in the prevalence of obesity in the UK over the last 25 years as well as a vast increase in developing countries.

The control of appetite is extremely complex. For example, if one considers only one signal, i.e. leptin, there can be leptin resistance where obese individuals have high circulating leptin but with no reduction of appetite. In contrast, in acute starvation, leptin concentrations decrease to lower levels than expected from the prevailing adipose tissue mass. It is known that cytokines, such as TNF and IL-2, which are elevated in a wide range of inflammatory and traumatic conditions, also suppress appetite, although the exact pathways involved are not entirely clear. Finally, there are a range of transmitters in the central nervous system, some of which appear to inhibit appetite (dopamine, serotonin, γ-aminobutyric acid) and others which stimulate appetite (e.g. opioids).

Energy expenditure

Basal metabolic rate (BMR). BMR in obese subjects is higher than in lean subjects, which is not surprising since obesity is associated with an increase in lean body mass.

Physical activity. Obese patients tend to expend more energy during physical activity as they have a larger mass to move. On the other hand, many obese patients decrease their amount of physical activity. The energy expended on walking at 3 miles per hour is only 15.5 kJ/ min (3.7 kcal/min) and therefore increasing exercise plays only a small part in losing weight. Nevertheless, because increased body fat develops insidiously over many years, any change in energy balance is helpful.

Thermogenesis

About 10% of ingested energy is dissipated as heat and is unconnected with physical activity. This dietary induced thermogenesis has been reported to be lower in obese and post-obese subjects than in lean subjects.

This would tend to favour energy deposition in obesity and those predisposed to obesity. However, other workers have found no difference in dietary induced thermogenesis between lean and obese subjects.

Brown adipose tissue in animals, when stimulated by cold or food, dissipates the energy derived from ingested food into heat. This can be a major component of overall energy balance in small mammals. However, the importance of brown adipose tissue thermogenesis in adult humans is likely to be very small, and of doubtful clinical significance. β_3-adrenergic receptors are the principal receptors mediating catecholamine-stimulated lipolysis in brown adipose tissue and to a lesser extent at other sites. Drugs with β_3-adrenergic activities have been developed, but side-effects have limited their use.

Table 5.14 Potential benefits that may result from the loss of 10 kg in patients who are initially 100 kg and suffer from co-morbidities

Mortality	20–25% fall in total mortality 30–40% fall in diabetes-related deaths 40–50% fall in obesity-related cancer deaths
Blood pressure	Fall of about 10 mmHg (systolic and diastolic)
Diabetes	Reduces risk of developing diabetes by >50% Fall of 30–50% in fasting blood glucose Fall of 15% in HbA_{1c}
Serum lipids	Fall of 10% in total cholesterol Fall of 15% in LDL cholesterol Fall of 30% in triglycerides Increase of 8% in HDL cholesterol

Morbidity and mortality

Obese patients are at risk of early death, mainly from diabetes, coronary heart disease and cerebrovascular disease. The greater the obesity, the higher the morbidity and mortality rates. For example, men who are 10% overweight have a 13% increased risk of death, while the increase in mortality for those 20% overweight is 25%. The rise is less in women, and in men over 65 obesity is not an independent risk factor. Weight reduction reduces this mortality and therefore should be strongly encouraged. The benefits are probably greater in more obese subjects (Table 5.14).

Clinical features

Most patients recognize their own problems, although often they are unaware of the main foods that cause obesity. Many symptoms are related to psychological problems or social pressures, such as the woman who cannot find fashionable clothes to wear.

The degree of obesity can be assessed by comparison with tables of ideal weight for height, from the BMI (Box 5.4), and by measuring skinfold thickness. The latter should be measured over the middle of the triceps muscle; normal values are 20 mm in a man and 30 mm in a woman. A central distribution of body fat (a waist/hip circumference ratio of >1.0 in men and >0.9 in women) is associated with a higher risk of morbidity and mortality than is a more peripheral distribution of body fat (waist/hip ratio <0.85 in men and <0.75 in women). This is because fat located centrally, especially inside the abdomen, is more sensitive to lipolytic stimuli, with the result that the abnormalities in circulating lipids are more severe.

Table 5.15 shows the conditions and complications that are associated with obesity.

The relationship between cardiovascular disease (hypertension or ischaemic heart disease), hyperlipidaemia, smoking, physical exercise and obesity is complex. Difficulties arise in interpreting mortality figures because of the number of factors involved. Many studies do not differentiate

Box 5.4 Ranges of body mass index (BMI) used to classify degrees of overweight and associated risk of co-morbidities

WHO classification	BMI (kg/m^2)	Risk of co-morbidities
Overweight	25–30	Mildly increased
Obese	>30	
Class I	30–35	Moderate
Class II	35–40	Severe
Class III	>40	Very severe

Table 5.15 Conditions and complications associated with obesity

Psychological
Osteoarthritis of knees and hips
Varicose veins
Hiatus hernia
Gallstones
Postoperative problems
Back strain
Accident proneness
Obstructive sleep apnoea
Hypertension
Breathlessness
Ischaemic heart disease
Stroke
Diabetes mellitus (type 2)
Hyperlipidaemia
Menstrual abnormalities
Increased morbidity and mortality
Increased cancer risk
Heart failure

between the types of physical exercise taken or take into account the cuff-size artefact in the measurement of blood pressure (an artefact will occur if a large cuff is not used in patients with a large arm). Nevertheless, obesity almost certainly plays a part in all of these diseases and should be treated. An exception is that stopping smoking, even if accompanied by weight gain, is more beneficial than any of the other factors. Physical fitness is also helpful, and there is some evidence to suggest that a fit obese person may have similar or even lower cardiovascular risk than a leaner unfit person.

Metabolic syndrome

There are two classification systems which are shown in Table 5.16. The differences are:

- A large waist is an absolute requirement for the International Diabetes Federation (IDF), but not in the ATP III NCEP.
- The IDF criteria use lower cut-off values for waist circumference (close to values of people with a BMI of 25 kg/m^2) and lower fasting blood glucose concentrations.

Table 5.16 Classification systems for metabolic syndrome: ATP III of the National Cholesterol Education Program (NCEP) and International Diabetes Federation (IDF)

Risk factor	ATP III NCEP (any 3 of the 5 features)	International Diabetes Federation (large waist + any other 2 features)
Waist circumference		
Men	>102 cm (40 in)	>94 cm (37 in)
Women	>88 cm (35 in)	>80 cm (35 in)
Triglycerides	>150 mg/dL (1.7 mmol/L)	>150 mg/dL (1.7 mmol/L)
HDL cholesterol		
Men	<40 mg/dL (1.03 mmol/L)	≤40 mg/dL (1.03 mmol/L)
Women	<50 mg/dL (1.29 mmol/L)	<50 mg/dL (1.29 mmol/L)
Blood pressure	≥130/≥85 mmHg	≥130/≥85 mmHg
Fasting glucose	≥110 mg/dL (6.1 mmol/L)	≥100 mg/dL (5.6 mmol/L)

ATP III = Adult Treatment Panel 3

This means that the prevalence of metabolic syndrome will be higher using the IDF criteria. These criteria will identify at-risk patients at an earlier stage. This could lead to further investigations from this initial screening, and institution of preventative as well as therapeutic measures.

Overweight/obesity and insulin resistance, which causes glucose and lipid disturbances, seem to form the basis of many features of the metabolic syndrome. Other classification systems also exist, e.g. using BMI instead of waist circumference.

The metabolic syndrome is a combination of risk factors (Table 5.16). Its clinical role in the prediction of the risk of cardiovascular disease is in doubt as the sum of the combined risk factors in the syndrome does not offer more than the individual factors added together.

Treatment

Dietary control

This largely depends on a reduction in calorie intake.

The most common diets allow a daily intake of approximately 4200 kJ (1000 kcal), although this may need to be nearer 6300 kJ (1500 kcal) for someone engaged in physical work. Very low calorie diets are also advocated by some, usually over shorter periods of time, but unless they are accompanied by changes in lifestyle, weight regain is likely. Patients must realize that prolonged dieting is necessary for large amounts of fat to be lost. Furthermore, a permanent change in eating habits is required to maintain the new low weight. It is relatively easy for most people to lose the first few kilograms, but long-term success in moderate obesity is poor, with an overall success rate of no more than 10%.

Many dietary regimens aim to produce a weight loss of approximately 1 kg per week. Weight loss will be greater initially owing to accompanying protein and glycogen breakdown and consequent water loss. After 3–4 weeks, further

weight loss may be very small because only adipose tissue is broken down and there is less accompanying water loss.

Patients must understand the principles of energy intake and expenditure, and the best results are obtained in educated, well-motivated patients. Constant supervision by healthcare professionals, by close relatives or through membership of a slimming club helps to encourage compliance. It is essential to establish realistic aims. A 10% weight loss, which is regarded by some as a 'success' (see Table 5.14) is a realistic initial aim.

An increase in exercise will increase energy expenditure and should be encouraged – provided there is no contraindication – since weight control is usually not achieved without exercise. The effects of exercise are complex and not entirely understood. However, exercise alone will usually produce little long-term benefit. On the other hand there is evidence to suggest that in combination with dietary therapy, it can prevent weight being regained. In addition, regular exercise (30 min daily) will improve general health.

The diet should contain adequate amounts of protein, vitamins and trace elements. A diet of 4200 kJ (1000 kcal) per day should be made up of more than 50 g protein, approximately 100 g of carbohydrate, and 40 g of fat. The carbohydrate should be in the form of complex carbohydrates such as vegetables and fruit rather than simple sugars. Alcohol contains 29 kJ/g (7 kcal/g) and should normally be discouraged. It can be substituted for other foods in the diet, but it often reduces the willpower. With a varied diet, vitamins and minerals will be adequate and supplements are not necessary.

A balanced diet, attractively presented, is of much greater value and safer than any of the slimming regimens often advertised in magazines.

Most obese people oscillate in weight; they often regain the lost weight, but many manage to lose weight again. This 'cycling' in bodyweight may play a role in the development of coronary artery disease.

A wide range of diets are available, including low-fat or low-carbohydrate diets, and some suit certain individuals better than others. The following general statements can be made about them.

- All low-calorie diets produce loss of bodyweight and fat, irrespective of dietary composition. Short-term weight loss is faster on low-carbohydrate diets, as a result of greater loss of body water, which is regained after the end of dietary therapy.
- Very low-fat diets are often low in vitamins E, B_{12}, and zinc. Very low-carbohydrate diets may be nutritionally inadequate, and may lead to deficiencies.
- Low-fat diets decrease LDL triglycerides and increase HDL, whereas low-carbohydrate diets produce a greater decrease in HDL and triglyceride, with no change in LDL.
- There are some potential long-term concerns with low-carbohydrate diets (high in fat and protein), including increased risk of osteoporosis, renal stones and atheroma (due to high saturated fat, high *trans* fat and cholesterol and the lack of fruits, vegetables and whole grains), but long-term studies are lacking.
- Low-energy-density diets, often bulky and rich in fibre and complex carbohydrates, may be more satiating but they are often less palatable than high-energy-dense diets which may affect long-term compliance.
- Liquids, e.g. soft drinks, appear to be less satiating than solid foods.
- A recent study has shown that Mediterranean and low-carbohydrate diets are as effective as a low-fat diet for weight loss.

Behavioural modification

The aim of behavioural modification is to encourage the patient to take personal responsibility for changing lifestyle, which will determine dietary habits and physical activity. Family therapy may also be useful, especially when it involves obese children. It can be time-consuming and expensive. Cognitive behavioural therapy is even more time-consuming and expensive.

Drug therapy

Drugs can be used in the short term (up to 3 months) as an adjunct to the dietary regimen, but they do not substitute for strict dieting.

Centrally acting drugs:

- Drugs acting on both serotoninergic and noradrenergic pathways, e.g. sibutramine, tesofensine.
- Cannabinoid-1 receptor blockers, e.g. rimonabant (now withdrawn due to depression/suicide risk), acting on the endocannabinoid system.
- Drugs acting on the noradrenergic pathways do suppress appetite but all have been withdrawn in the UK because of cardiovascular side-effects.

Peripherally acting drugs:

- *Orlistat* is an inhibitor of pancreatic and gastric lipases. It reduces dietary fat absorption and aids weight loss. Weight regain occurs after the drug is stopped. It has been used continuously in a large-scale trial for up to 2 years. The patients complain of diarrhoea during treatment and to avoid this take a low-fat diet resulting in weight loss.
- *Incretins.* Glucagon-like peptide 1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP) are being used in type 2 diabetes mellitus (p. 1039). They suppress appetite and are being used in obesity (Fig. 5.11)

A systematic review of long-term pharmacotherapy concluded that there was a paucity of long-term studies with antiobesity agents, and that in weight loss trials of 1-year duration, orlistat and sibutramine appear to be only modestly effective in promoting weight loss (2.7 and 4.3 kg greater weight loss respectively than the control group). *Rimonabant* produces a greater weight loss (6.7 kg at 1 year compared to a control group) in addition to hypocaloric dieting but it can produce anxiety and severe depression; it does not have FDA approval. Other randomized trials show that a combination of lifestyle modification and pharmacotherapy (sibutramine) produces greater weight loss than either treatment alone.

Surgical treatment

Surgery is used in some cases of morbid obesity (BMI >40 kg/m^2) or patients with a BMI >35 kg/m^2 and obesity-

related complications, after conventional medical treatments have failed. A variety of gastrointestinal surgical procedures have been used. They fall into two main groups:

- restrictive procedures, which restrict the ability to eat
- malabsorptive procedures, which reduce the ability to absorb nutrients.

A systematic analysis of bariatric surgical procedures concluded that, in comparison to non-surgical treatments, they produced significantly more weight loss (23–37 kg), which was maintained to 8 years and associated with improvement in quality of life and co-morbidities.

- *Roux-en-Y gastric bypass*. This procedure incorporates both restrictive and malabsorptive elements (gastrojejunostomy). This procedure, like other malabsorptive procedures, may result in nutrient deficiencies requiring careful long-term follow-up.
- *Bilio-pancreatic diversion (including the duodenal switch variation)*. This is another malabsorptive procedure that requires long-term evaluation.
- *Laparoscopic adjustable gastric banding*. This is a restrictive procedure, in which a band is placed around the upper stomach to produce a small proximal pouch and a large distal remnant. It has a perioperative mortality of < 0.5%, and results in loss of about 60% (43–78%) of the excess weight after 3 years, although longer-term follow-up studies are required. Laparoscopic adjustable gastric banding has been reported to produce greater weight loss and fewer side-effects (e.g. vomiting) and operative revisions than vertical banded gastroplasty.
- *Liposuction*. The removal of large amounts of fat by suction (liposuction) does not deal with the underlying problem and weight regain frequently occurs. There is no reduction in cardiovascular risk factors.

Prevention

Preventing obesity must always be the goal because most obese people find it difficult to maintain any weight loss they have managed to achieve. All health professionals must be aware of the dangers of obesity and encourage children, young as well as older adults, from gaining too much weight. A small gain each year over a long period produces an obese individual for whom treatment is difficult. Public health policies should consider creation of public places to encourage physical activity and fitness, education about the benefits of losing weight or not gaining it through healthy eating and physical activity, and changes in food composition (alternatives to high-fat, high-energy-dense foods).

The present obesity epidemic has resulted from lifestyle changes, and so it is appropriate to promote lifestyle changes, not only as the first line therapy for most overweight and obese individuals, but also in the prevention of overweight and obesity. Lifestyle modification would involve changes in the amount of time watching television and using computers, use of bicycle paths, dietary changes, and educational activities of patients and public, parents and children. All the specific therapies discussed above, including pharmacotherapy, should be part of a package that involves lifestyle modification.

FURTHER READING

Baggio LL, Drucker DJ. Biology of incretins: GLP-1 and GIP. *Gastroenterology* 2007; **132**: 2131–2157.

Byers T. Overweight and mortality among baby boomers – now we're getting personal. *New England Journal of Medicine* 2006; **355**: 758–760.

Eckel RH. Nonsurgical management of obesity in adults. *New England Journal of Medicine* 2008; **358**: 1941–1950.

Franzosi MG. Should we continue to use BMI as a cardiovascular risk factor? *Lancet* 2006; **368**: 624–625.

Iltz JL, Baker DE, Setter SM et al. Exenatide: an incretin mimetic for the treatment of type 2 diabetes mellitus. *Clinical Therapeutics* 2006; **28**: 652–665.

Kahn R. Metabolic syndrome – what is the clinical significance? *Lancet* 2008; **371**: 1892–1893.

McMahon MM, Sarr MG, Clark MM et al. Clinical management after bariatric surgery: value of a multidisciplinary approach. *Mayo Clinic Proceedings* 2006; **81**: S34–45.

National Institute for Health and Clinical Excellence. Obesity: the prevention, identification, assessment and management of overweight and obesity in adults and children. Clinical Guideline 43, 2006. www.nice.org.uk

Shai I, SchwarzFuchs D, Henkin Y et al. Weight loss with a low-carbohydrate, Mediterranean, or low-fat diet. *New England Journal of Medicine* 2008; **359**: 229–241.

Wadden TA, Butryn ML, Wilson C. Lifestyle modification for the management of obesity. *Gastroenterology* 2007; **132**: 2226–2238.

NUTRITIONAL SUPPORT IN THE HOSPITAL PATIENT

Nutritional support is recognized as being necessary in many hospitalized patients. The pathophysiology and hallmarks of malnutrition have been described earlier (p. 213); here the forms of nutritional support that are available are discussed, along with special nutritional requirements in some diseases.

Principles

Some form of nutritional supplementation is required in those patients who cannot eat, should not eat, will not eat or cannot eat enough. All patients should be screened for malnutrition on admission and the findings linked to a care plan, preferably under the supervision of a trained multidisciplinary team. The Council of Europe has produced 10 key characteristics of good nutritional care in hospital (bapen.org.uk). Plans are discussed with patients and consent is taken for any invasive procedure (e.g. nasogastric tube, parenteral nutrition). If the patient is unable to give consent, the healthcare team should act in the patient's best interest, taking into account previously expressed wishes of the patient and views of the family. It is usually necessary to provide nutritional support for:

- all severely malnourished patients on admission to hospital
- moderately malnourished patients who, because of their physical illness, are not expected to eat for more than 3–5 days
- normally nourished patients expected not to eat for more than 5 days or to eat less than half their intake for more than 8–10 days.

Enteral rather than parenteral nutrition should be used if the gastrointestinal tract is functioning normally.

In the *refeeding syndrome*, the shifts of water and electrolytes that occur after parenteral and enteral nutrition can be life threatening. Carbohydrate intake stimulates insulin release which leads to cellular uptake of phosphate, potassium and magnesium. Complications include hypophosphataemia, hypokalaemia, hypomagnesaemia and fluid overload because of sodium retention (decreased renal excretion of sodium and water). Patients who have eaten little or nothing for more than 5 days should initially receive no more than 50% of their energy requirements (NICE guidelines).

Nutritional requirements for adults

- *Water*. Typical requirements are ~2–3 L/day. Increased requirements occur in patients with large-output fistulae, nasogastric aspirates and diarrhoea. Reduced requirements occur in patients with oedema, hepatic failure, renal failure (oliguric and not dialysed) and brain oedema.
- *Energy*. Typical requirements are ~7.5–10.0 MJ/day (1800–2400 kcal/day). Disease increases resting energy expenditure but decreases physical activity. Extra energy is given for repletion and reduced energy for obesity.
- *Protein*. Typically 10–15 g N/day (62–95 g protein/day) or 0.15–0.25 g N/kg/day (0.94–1.56 g protein/kg/day). Extra protein may be needed in severely catabolic conditions, such as extensive burns.
- *Major minerals*. Typical requirements for sodium and potassium are 70–100 mmol/day. Increased requirements occur in patients with gastrointestinal effluents. The excretion of these minerals in various effluents can provide an indication of the additional requirements (Table 12.10). Low requirements may be necessary in those with fluid overload (or patients with hypernatraemia and hyperkalaemia). The requirements of calcium and magnesium are higher for enteral than for parenteral nutrition because only a proportion of these minerals is absorbed by the gut.
- *Trace elements*. For trace elements such as iodide, fluoride and selenium that are well absorbed, the requirements for enteral and parenteral nutrition are similar. For other trace elements, such as iron, zinc, manganese and chromium, the requirements for parenteral nutrition are substantially lower than for enteral nutrition (Fig. 5.12).
- *Vitamins*. Many vitamins are given in greater quantities in patients receiving parenteral nutrition than in those receiving enteral nutrition (Fig. 5.13). This is because patients on parenteral nutrition may have increased requirements, partly because of severe disease, partly because they may already have depleted pools of vitamins, and partly because some vitamins degrade during storage. Vitamin K is usually absent from parenteral nutrition regimens and therefore it may need to be administered separately.

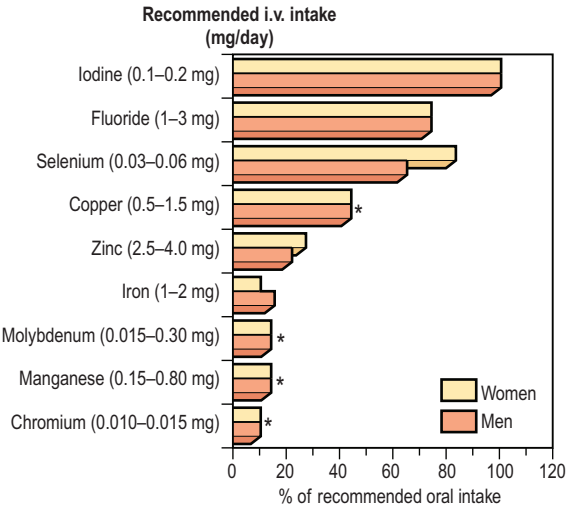


Fig. 5.12 **Trace elements. Recommended intravenous intake in absolute values and as a percentage of recommended oral intake**. Trace elements marked with an asterisk are those for which there was too little information to establish a recommended value for dietary oral intake, therefore the midpoint of estimated safe and adequate oral intake is used for comparison.

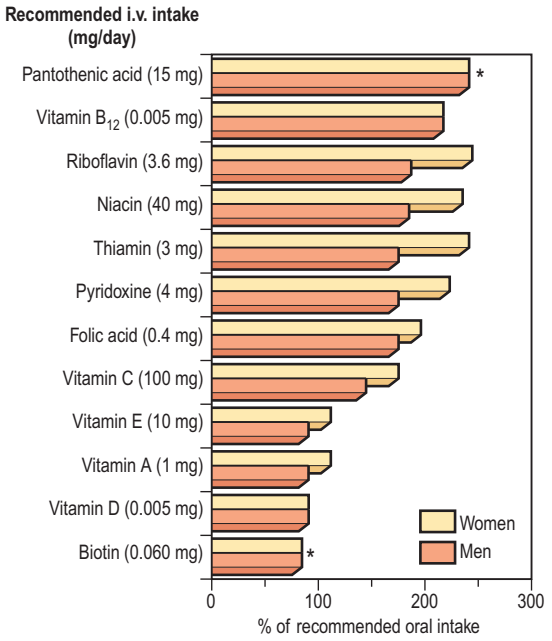


Fig. 5.13 **Vitamins. Recommended intravenous intake in absolute values and as a percentage of recommended oral intake**. Vitamins marked with an asterisk are those for which there was too little information to establish recommended dietary oral intake, therefore the midpoint of estimated safe and adequate oral intake is used for comparison.

Enteral nutrition (EN)

Feeds can be given by various routes:

- By mouth (food can be supplemented with solid or liquid supplements with multiple benefits).
- By fine-bore nasogastric tube (Practical Box 5.1)
- Percutaneous endoscopic gastrostomy (PEG) is useful for patients who need enteral nutrition for a prolonged

Practical Box 5.1 Enteral feeding via nasogastric tube

The procedure should be explained to the patient and consent taken.

Procedure

- Insert fine-bore tube intranasally with wire stylet.
- Confirm position of tube in stomach by aspiration of gastric contents and auscultation of the epigastrium.
- Check by X-ray if aspiration or auscultation is unsuccessful.

Problems

No satisfactory way of keeping nasogastric tubes in place (up to 60% come out).

Main complications

- Regurgitation and aspiration into bronchus.
- Blockage of the nasogastric tube.
- Gastrointestinal side-effects, the most common being diarrhoea.
- Metabolic complications including hyperglycaemia and hypokalaemia, as well as low levels of magnesium, calcium and phosphate.

period (e.g. more than 30 days), such as those with swallowing problems following a head injury or in elderly people after a stroke. A catheter is placed percutaneously into the stomach under endoscopic control.

- With needle catheter jejunostomy, a fine catheter is inserted into the jejunum at laparotomy and brought out through the abdominal wall.

Diet formulation (see Table 5.17)

A polymeric diet with whole protein and fat can be used, except in patients with severely impaired gastrointestinal function who may require a predigested (i.e. elemental) diet. In these patients, the nitrogen source is purified low-molecular-weight peptides or amino acid mixtures, with sometimes the fat being given partly as medium-chain triglycerides.

Management

Daily amounts of diet vary between 1.5 and 2.5 L, but small amounts are started in patients with suspected poor gastric emptying and severe malnutrition (to avoid the refeeding syndrome).

Hypercatabolic patients require a high supply of nitrogen (15 g daily) and often will not achieve positive nitrogen balance until the primary injury is resolved.

The success of enteral feeding depends on careful supervision of the patient, with monitoring of weight, biochemistry and diet charts.

Parenteral nutrition

Peripheral parenteral nutrition

Specially formulated mixtures for peripheral use are available, with a low osmolality and containing lipid emulsions. Heparin and corticosteroids can be added to the infusion and local application of glyceryl trinitrate patches reduces the occurrence of thrombophlebitis and prolongs catheter life.

- Peripheral cannulas can be inserted into a mid arm vein (20 cm) and can last up to 5 days.

Table 5.17 Standard enteric diet, providing 8.4 MJ per day (= 2000 kcal)

Energy

Carbohydrate as glucose polymers (49–53% of total energy)
Fat as triglycerides (30–35% of total energy)

Nitrogen

Whole protein (10–14 g of nitrogen/day)
Additional electrolytes, vitamins and trace elements

Features

Ratio of energy to nitrogen kJ:g = 620:1 (kcal:g = 150:1)
Osmolality = 285–300 mOsm/kg

Practical Box 5.2 Central catheter placement for parenteral nutrition

This should be performed only by experienced clinicians under aseptic conditions in an operating theatre.

Give an explanation and obtain consent from the patient.

- The patient is placed supine with 5° of head-down tilt to avoid air embolism.
- The skin below the midpoint of the right clavicle is infiltrated with 1–2% lidocaine and a 1 cm skin incision is made.
- A 20-gauge needle on a syringe is inserted beneath the clavicle and first rib and angled towards the tip of a finger held in the suprasternal notch.
- When blood is aspirated freely, the needle is used as a guide to insert the cannula through the skin incision and into the subclavian vein.
- The catheter is advanced so that its tip lies in the distal part of the superior vena cava.
- A skin tunnel is created under local anaesthetic using an introducer inserted through a point about 10 cm below and medial to the incision and passed upwards to the incision.
- The proximal end of the catheter (with hub removed) is passed backwards through the introducer to emerge 10 cm below the clavicle, where it is sutured to the chest wall.
- The original infraclavicular entry incision is now sutured.

- A longer (60 cm) peripherally inserted central catheter (PICC) inserted into an antecubital fossa vein has its distal end lying in a central vein; here there is less risk of thrombophlebitis and hyperosmolar solutions can be given.

With careful management, these catheters can last for up to a month. Peripheral parenteral nutrition is often preferred initially, allowing time to consider the necessity for having to insert a central venous catheter.

Parenteral nutrition via a central venous catheter (PN) (see Practical Box 5.2)

A silicone catheter is placed into a central vein, usually using the infraclavicular approach to the subclavian vein. The skin-entry site should be dressed carefully and not disturbed unless there is a suggestion of catheter-related sepsis.

Complications of catheter placement include central vein thrombosis, pneumothorax and embolism, but the major problem is catheter-related sepsis. Organisms, mainly staphylococci, enter along the side of the catheter, leading to septicaemia. Sepsis can be prevented by careful and sterile placement of the catheter, by not removing the dressing over the catheter entry site, and by not giving other substances (e.g. blood products, antibiotics) via the central vein catheter.

Sepsis should be suspected if the patient develops fever and leucocytosis. In two-thirds of cases, organisms can be grown from the catheter tip after removal. Treatment involves removal of the catheter and appropriate systemic antibiotics.

Nutrients

With PN it is possible to provide sufficient nitrogen for protein synthesis and calories to meet energy requirements. Electrolytes, vitamins and trace elements are also necessary. All of these substances are infused simultaneously.

Nitrogen source

Most patients receive at least 11–15 g N per day, in the form of synthetic L-amino acids.

Energy source

This is provided by glucose, with additional calories provided by a fat emulsion. Fat infusions provide a greater number of calories in a smaller volume than can be provided by carbohydrate. Fat infusions are not hypertonic and they also prevent essential fatty acid deficiency.

Essential fatty acid deficiency has been reported in long-term parenteral nutritional regimens without fat emulsions. It causes a scaly skin, hair loss and a delay in healing.

Electrolytes, vitamins and trace elements

(See Figs 5.12 and 5.13.) Initially, the electrolyte status should be monitored on a daily basis and electrolyte solutions given as appropriate. Fat- and water-soluble vitamins and minerals including trace elements should be given routinely.

Administration and monitoring

Peripheral parenteral nutrition. Administered via 3-L bags over 24 hours, with the constituents being premixed under sterile conditions by the pharmacy. Table 5.18 shows the composition which provides 9 g of nitrogen and 1700 calories in 24 hours.

Central venous PN regimen. Most hospitals now use premixed 3-L bags. A standard parenteral nutrition regimen which provides 14 g of nitrogen and 2250 calories over 24 hours is also given in Table 5.18.

Monitoring includes:

- *Blood tests*. Daily plasma electrolytes and glucose. Twice weekly FBC, liver biochemistry and function, calcium, phosphate, and magnesium, zinc and glycerides weekly.
- *Nutritional status*. Weekly weight and skin fold thickness if appropriate callipers are available.
- *Nitrogen balance* (p. 210) assessment, but this requires complete collections of urine.

Complications

- Catheter-related (see above)
- Metabolic (e.g. hyperglycaemia – insulin therapy is usually necessary)
- Fluid and electrolyte disturbances
- Hypercalcaemia
- Liver dysfunction
- Nutrient deficiencies (if inadequately provided).

Table 5.18 Examples of parenteral nutrition regimens

Peripheral: all mixed in 3-L bags and infused over 24 hours		
Nitrogen	L-amino acids 9 g/L	1 L
Energy	Glucose 20%	1 L
	Lipid 20%	0.5 L
	+ Trace elements, electrolytes, and water-soluble and fat-soluble vitamins, heparin 1000 UL and hydrocortisone 100 mg; insulin is added if required. Nitrogen 9 g, non-protein calories 7206 kJ (1700 kcal)	
Central: all mixed in 3-L bags and infused over 24 hours		
Nitrogen	L-amino acids 14 g/L	1 L
Energy	Glucose 50%	0.5 L
	Glucose 20%	0.5 L
	+ Lipid 10% as either Intralipid or Lipofundin	0.5 L Fractionated soya oil 100 g/L Soya oil 50 g, medium-chain triglycerides 50 g/L
	+ Electrolytes, water-soluble vitamins, fat-soluble vitamins, trace elements, heparin and insulin may be added if required. Nitrogen 14 g, non-protein calories 9305 kJ (2250 kcal)	

NUTRITIONAL SUPPORT IN THE HOME PATIENT

In both high- and low-income countries there is considerably more undernutrition in the community than in hospital. However, the principles of care are very similar: detection of malnutrition and the underlying risk factors; treatment of underlying disease processes and disabilities; correction of specific nutrient deficiencies and provision of appropriate nutritional support. This typically begins with dietary advice, and may involve the provision of 'meals on wheels' by social services. A systematic review of the use of nutritional supplements in the community came to the following conclusions:

- Supplements are generally of more value in *patients with a BMI <20 kg/m^2 and children with growth failure* (weight for height <85% of ideal) than in those with better anthropometric indices. They are likely to be of little or no value in patients with little weight loss and a BMI >20 kg/m^2. The supplemental energy intake in such subjects largely replaces oral food intake.
- Supplements may be of value in *weight-losing patients* (e.g. >10% weight loss compared to pre-illness) with a BMI >20 kg/m^2, and in *children with deteriorating growth performance* without chronic protein–energy undernutrition.
- The functional benefits vary according to the patient group. In patients with *chronic obstructive airways disease* the observed functional benefits were increased respiratory muscle strength, increase in handgrip strength, and an increase in walking distance/duration of exercise. In the *elderly* the benefits were reduced number of falls, or increase in activities of daily living, and reduced pressure sore surface area. In patients with *HIV/AIDS* there were changes in immunological function and improved cognition. Patients with *liver disease* experienced a lower incidence of severe infections and had a lower frequency of hospitalization.

- *Acceptability and compliance* are likely to be better when a choice of supplements (of type, flavour, consistency) and schedule is decided in conjunction with the patient and/or carer. Changes in these may be necessary when there is a change in patterns of daily activities, disease status, and 'taste fatigue' with prolonged use of the same supplement.
- Nutritional *counselling and monitoring* are recommended before and after the start of supplements (see also below).

Some patients receive enteral tube feeding and parenteral nutrition at home. At any one point in time in developed countries enteral tube feeding occurs more frequently at home than in hospital. In adults the commonest reason for starting home tube feeding is for swallowing difficulties. This involves patients with neurological disorders, such as motor neurone disease, multiple sclerosis and Parkinson's disease, but the commonest single diagnosis is cerebrovascular disease. Approximately 2% of patients who have had a stroke in the UK receive home enteral tube feeding (HETF). However, in a British Nutrition survey (1996–2000) of patients with these disorders (apart from Parkinson's), only 15% in total were able to return to oral feeding after one year. The swallowing capabilities of patients should be assessed regularly in order to avoid unnecessary tube feeding. The patients and/or carers should have adequate training, contacts with appropriate health professions, and a reliable delivery service for feeds and ancillary equipment. They should also be clear about how to manage simple problems associated with the feeding tube, which is usually a gastrostomy tube rather than a nasogastric tube.

Home parenteral nutrition is practised much less frequently, usually under the supervision of specialist centres. The potential value of intestinal transplantation in patients with long-term intestinal failure is still being assessed.

FURTHER READING

Elia M, Stroud M. Nutrition in acute care. *Clinical Medicine* 2004; **4**: 405–407.

Mehanna HM, Moledina J, Travis J. Refeeding syndrome. *British Medical Journal* 2008; **336**: 1495–1498.

National Institute for Health and Clinical Excellence. Nutrition support in adults. Clinical Guideline 32, 2006. www.nice.org.uk

Smith T, Elia M. Artificial nutrition support in hospital: indications and complications. *Clinical Medicine* 2006; **6**: 457–460.

Stratton RJ, Green CJ, Elia M. *Disease-related Malnutrition. An Evidence-based Approach to Treatment*. Oxford: CABI Publishing (CAB International), 2003.

Stroud M, Duncan H, Nightingale J. Guidelines for enteral feeding in adult hospital patients. *Gut* 2003; **52** (Suppl VII): vii1–vii12.

Zaloga GP. Parenteral nutrition in adult inpatients with functioning gastrointestinal tracts: assessment of outcomes. *Lancet* 2006; **367**: 1101–1111.

FOOD ALLERGY AND FOOD INTOLERANCE

Many people ascribe their various symptoms to food, and many such sufferers are seen and started on exclusion diets. The scientific evidence that food does harm in most instances is weak, although adverse reactions to food certainly exist. These can be divided into those that involve immune mechanisms (food allergy) and those that do not (food intolerance).

Food allergy

Food allergy, which is estimated to affect up to about 5% of young children and about 1–2% of adults, may be IgE mediated or non-IgE mediated (T-cell mediated). The IgE-mediated reactions tend to occur early after a food challenge (within minutes to an hour). Adults tend to be allergic to fish, shellfish and peanuts, while children tend to be allergic to cow's milk, egg white, wheat, and soy. Peanuts are very allergenic and peanut allergy persists throughout life. The following conditions can result from food allergy:

- *Acute hypersensitivity*. An example is urticaria, vomiting or diarrhoea after eating nuts, strawberries or shellfish. These IgE-mediated reactions do not usually produce clinical problems as the patients have already learned to avoid the suspected food. Inadvertent ingestion of the incriminating food can sometimes occur, leading to angioneurotic oedema (p. 1244).
- *Eczema and asthma*. These tend to affect young children and are often due to egg and are IgE mediated.
- *Rhinitis and asthma*. These have been produced by foods such as milk and chocolate, mainly in atopic subjects.
- *Chronic urticaria*. This has been treated successfully by an exclusion diet.
- *Food-sensitive enteropathy*. This may manifest itself as coeliac disease (gluten (wheat) sensitive enteropathy), and cow's milk enteropathy (in infants) and is T-cell mediated.

Food intolerance

- *Migraine*. This sometimes follows the intake of foods such as chocolate, cheese and alcohol, which are rich in certain amines, such as tyramine. Patients on monoamine oxidase inhibitors, which are involved in the metabolism of these amines, are particularly vulnerable.
- *Irritable bowel syndrome*. In some patients this seems to be related to ingestion of certain food items, such as wheat, but the mechanisms are not clearly defined.
- *Chinese restaurant syndrome*. Monosodium glutamate, a flavour enhancer used in cooking Chinese food, may produce dizziness, faintness, nausea, sweating and chest pains.
- *Lactose intolerance*. Patients develop abdominal bloating and diarrhoea following ingestion of lactose, which is present in milk (p. 277). This is probably the commonest form of food intolerance world-wide, and may be genetic in origin.
- *Phenylketonuria*. This can also be classified as a form of food intolerance, and is due to lack of phenylalanine hydroxylase, which is necessary for the metabolism of phenylalanine present in dietary protein.

A number of other inborn errors of metabolism can also be regarded as forms of food intolerance.

Food intolerance may be due to a constituent of food (e.g. the histamine in mackerel or canned food, or the tyramine in

cheeses), chemical mediators released by food (e.g. histamine may be released by tomatoes or strawberries), or toxic chemicals found in food (e.g. the food additive tartrazine).

Many other additives and compounds with certain E numbers have been implicated as causing reactions, but the evidence for this is poor.

There is little or no evidence to suggest that diseases such as arthritis, behaviour and affective disorders and Crohn's disease are due to ingestion of a particular food.

Multiple vague symptoms such as tiredness or malaise are also not due to food allergy. Most of the patients in this group are suffering from a psychiatric disorder (p. 1192).

Management

- A careful history may help to delineate the causative agent, particularly when the effects are immediate.
- Skin-prick testing with allergen and measurement in the serum of antigen or antibodies have not correlated with symptoms and are usually misleading. 'Fringe' techniques such as hair analysis, although widely advertised, are of no value.
- Diagnostic exclusion diets are sometimes used, but they are time-consuming. They can occasionally be of value in identifying a particular food causing problems.
- Dietary challenge consists of the food and the test being given sublingually or by inhalation in an attempt to reproduce the symptoms. Again this may be helpful in a few cases.
- Most people who have acute reactions to food realize it and stop the food, and do not require medical attention. In the remainder of patients, a small minority seem to be helped by modifying their diet, but there is no good scientific evidence to support these exclusion diets.

FURTHER READING

American Gastroenterology Association. American Gastroenterological Association medical position statement: guidelines for the evaluation of food allergy. *Gastroenterology* 2001; **120**: 1023–1025.

American Gastroenterology Association. Review on the evaluation of food allergy in GI disorder, official recommendation of American Gastroenterology Association. *Gastroenterology* 2001; **120**: 1026–1040.

Kimber I, Dearman RJ. Factors affecting the development of food allergy. *Proceedings of the Nutrition Society* 2002; **6**: 435–439.

Lack G, Fox D, Northstone K et al. Factors associated with the development of peanut allergy in childhood. *New England Journal of Medicine* 2003; **348**: 977–985.

ALCOHOL

Alcohol is a popular 'nutrient' consumed in large quantities all over the world. In many countries, alcohol consumption is becoming a major problem (see p. 1211).

Ethanol (ethyl alcohol) is oxidized, in the steps shown in Box 5.5, to acetaldehyde. Acetaldehyde is then converted to acetate, 90% in the liver mitochondria. Acetate is released into the blood and oxidized by peripheral tissues to carbon dioxide and water.

Alcohol dehydrogenases are found in many tissues and it has been suggested that enzymes present in the gastric mucosa may contribute substantially to ethanol metabolism.

Box 5.5 The main pathways of ethanol oxidization

- Alcohol dehydrogenase:

$$\underset{\text{(ethanol)}}{CH_3CH_2OH} + NAD^+ \xrightarrow{ADH} \underset{\text{(acetaldehyde)}}{CH_3CHO} + NADH + H^+ \quad [1]$$

- The liver microsomal enzyme oxidizing system (MEOS) including the specific P450 enzyme, CYP2EI, which is induced by ethanol:

$$CH_3CH_2OH + NADPH + H^+ + O_2 \xrightarrow{MEOS} CH_3CHO + NADP + 2H_2O \quad [2]$$

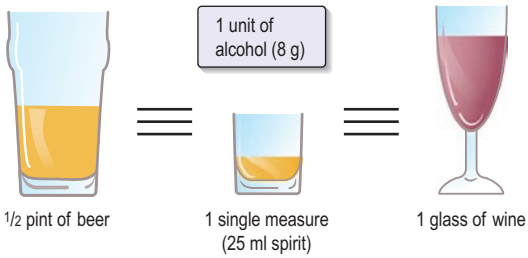


Fig. 5.14 **Measures of 1 unit of alcohol.**

Ethanol itself produces 29.3 kJ/g (8 kcal/g), but many alcoholic drinks also contain sugar, which increases their calorific value. For example, one pint of beer provides 1045 kJ (250 kcal), so the heavy drinker will be unable to lose weight if he or she continues to drink.

Effects of excess alcohol consumption

Excess consumption of alcohol leads to two major problems, both of which can be present in the same patient:

- alcohol dependence syndrome (p. 1212)
- physical damage to various tissues.

Each unit of alcohol (defined as one half pint of normal beer, one single measure of spirit, or one small glass of wine) contains 8 g of ethanol (Fig. 5.14). All the long-term effects of excess alcohol consumption are due to excess ethanol, irrespective of the type of alcoholic beverage; i.e. beer and spirits are no different in their long-term effects. Short-term effects, such as hangovers, depend on additional substances, particularly other alcohols such as isoamyl alcohol, which are known as congeners. Brandy and bourbon contain the highest percentage of congeners.

The amount of alcohol that produces damage varies and not everyone who drinks heavily will suffer physical damage. For example, only 20% of people who drink heavily develop cirrhosis of the liver. The effect of alcohol on different organs of the body is not the same; in some patients the liver is affected, in others the brain or muscle. The differences may be genetically determined.

Thiamin deficiency contributes to both neurological (confusion, Wernicke–Korsakoff syndrome; see p. 1174) and some of the non-neurological manifestations (cardiomyopathy). Susceptibility to damage of different organs is variable and the figures given in Box 5.6 are given only as a guide to

Box 5.6 Guide to sensible drinking of alcohol

Daily maximum

3 units for men
2 units for women

To help achieve this:

Use a standard measure.
Do not drink during the daytime.
Have alcohol-free days each week.

Remember

Health can be damaged without being 'drunk'.
Regular heavy intake is more harmful than occasional binges.
Do not drink to 'drown your problems'.
In the UK the drink-before-driving limit of alcohol in the blood is 800 mg/L (80 mg%).
One unit of alcohol is eliminated per hour, therefore spread drinking time.
Food decreases absorption and therefore results in a lower blood alcohol level.
4–5 units are sufficient to put the blood alcohol level over the legal driving limit in a 70 kg man (less in a lighter person).

Table 5.19 Physical effects of excess alcohol consumption

Central nervous system
Epilepsy
Wernicke–Korsakoff syndrome
Polyneuropathy

Muscles
Acute or chronic myopathy

Cardiovascular system
Cardiomyopathy
Beriberi heart disease
Cardiac arrhythmias
Hypertension

Metabolism
Hyperuricaemia (gout)
Hyperlipidaemia
Hypoglycaemia
Obesity

Endocrine system
Pseudo-Cushing's syndrome

Respiratory system
Chest infections

Gastrointestinal system
Acute gastritis
Carcinoma of the oesophagus or large bowel
Pancreatic disease
Liver disease

Haemopoiesis
Macrocytosis (due to direct toxic effect on bone marrow or folate deficiency)
Thrombocytopenia
Leucopenia

Bone
Osteoporosis
Osteomalacia

sensible drinking. Heavy persistent drinkers for many years are at greater risk than heavy sporadic drinkers.

Liver disease

In general the effects of a given intake of alcohol seem to be worse in women. The following figures are for men and should be reduced by 50% for women:

- 160 g ethanol per day (20 single drinks) carries a high risk
- 80 g ethanol per day (10 single drinks) carries a medium risk
- 40 g ethanol per day (five single drinks) carries little risk.

Alcohol consumption in pregnancy

Women are advised not to drink alcohol at all during pregnancy because even small amounts of alcohol consumed can lead to 'small babies'. The fetal alcohol syndrome is characterized by mental retardation, dysmorphic features and growth impairment; it occurs in fetuses of alcohol-dependent women.

Summary

A summary of the physical effects of alcohol is given in Table 5.19. Details of these diseases are discussed in the relevant chapters. The effects of alcohol withdrawal are discussed on page 1213.

FURTHER READING

Lieber CS. Medical disorders of alcoholism. *New England Journal of Medicine* 1995; **333**: 1058–1063.

CHAPTER BIBLIOGRAPHY

Ahima RS. Nutrition, obesity and metabolism. *Gastroenterology (Special Issue)* 2007; **132**: 2085–2275.

Gibney MJ, Elia M, Ljungqvist O et al. *Clinical Nutrition.* Oxford: Blackwell Science, 2006.

National Institute for Health and Clinical Excellence. Nutrition support in adults. Clinical Guideline 32, 2006. www.nice.org.uk

National Institute for Health and Clinical Excellence. Obesity: the prevention, identification, assessment and management of overweight and obesity in adults and children. Clinical Guideline 43, 2006. www.nice.org.uk

SIGNIFICANT WEBSITES

http://www.who.int/nutgrowthdb/
World Health Organization site, provides information on world-wide nutritional issues, resources and research

http://www.fao.org/
Food and Agriculture Organization (FAO) – autonomous body within the United Nations, aims to improve health through nutrition and agricultural productivity, especially in rural populations

http://www.ific.org/
International Food Information Council (IFIC) – non-profit organization providing access to health and nutrition resources to improve communication of health and nutrition information to consumers

http://www.ag.uiuc.edu/~food-lab/nat/
Free analysis of the nutrient content of food available to anyone (at University of Illinois, USA)

http://www.ama-assn.org/ama/pub/category/10931.html
American Medical Association: Assessment and management of adult obesity

http://www.nhlbi.nih.gov/health/public/heart/obesity/lose_wt/profmats.htm
National Heart, Lung and Blood Institute: Aim for a healthy weight

http://www.hda-online.org.uk/downloads/pdfs/obesity_evidence_briefing.pdf
Health Development Agency: Management of obesity and overweight

Selected nutrition journals (for more extensive website addresses see *Journal of Nutrition* 1997; **127**: 1527–1532):

http://www.faseb.org/ajcn
American Journal of Clinical Nutrition

http://www.nutrition.org/
The Journal of Nutrition

http://www.nature.com/ijo
International Journal of Obesity

http://www.cabi-publishing.org/JOURNALS/BJN/Index.asp
The British Journal of Nutrition

http://www.ilsi.org/publications/reviews.html
Nutrition Reviews

http://clinnutr.org/
Journal of Parenteral and Enteral Nutrition

http://www.naturesj.com/ejcn/
European Journal of Clinical Nutrition

6 Gastrointestinal disease

Gastrointestinal symptoms and signs 241
Clinical examination 243
Investigations 244
The mouth and salivary glands 246
The salivary glands 248
The pharynx and oesophagus 248
Gastro-oesophageal reflux disease (GORD) 250
Motility disorders 253
Other oesophageal disorders 254
Oesophageal tumours 256
The stomach and duodenum 258
Helicobacter pylori infection 258
Gastropathy and gastritis 262
Management of dyspepsia in the community 263
Gastric tumours 263
Acute and chronic gastrointestinal bleeding 267
The small intestine 270
Malabsorption 277
Miscellaneous intestinal diseases 282
Tumours of the small intestine 284
Inflammatory bowel disease (IBD) 285
The colon and rectum 294
Constipation 295
Miscellaneous colonic conditions 297
Diverticular disease 298
Anorectal disorders 299
Colonic tumours 300
Diarrhoea 305
Functional gastrointestinal disorders 309
Functional oesophageal disorders 310
Functional gastroduodenal disorders 311
Functional bowel disorders 311
The acute abdomen 314
The peritoneum 317

In *developed countries* gastrointestinal symptoms are a common reason for attendance to primary care clinics and to hospital outpatients. Approximately 75% of these consultations are for non-organic symptoms. The clinician's main task is therefore to recognize when organic disease must be sought or excluded, remembering that 20% of all cancers occur in the gastrointestinal tract (Fig. 6.1).

In *developing countries*, malnutrition and poor hygiene make infection a more probable diagnosis. The clinician needs to recognize and treat these infections promptly and also help with prevention by encouraging good hygiene.

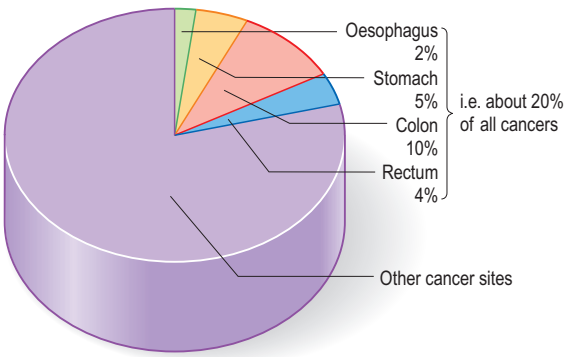


Fig. 6.1 **Incidence (approximate) of cancers at various sites of the gastrointestinal tract.**

GASTROINTESTINAL SYMPTOMS AND SIGNS

Dyspepsia and indigestion

Dyspepsia is an inexact term used to describe upper abdominal symptoms such as heartburn, acidity, pain or discomfort, nausea, wind, fullness or belching. Patients refer to all of these as 'indigestion', but may also include other symptoms such as constipation and the presence of undigested vegetable material in the stool. 'Indigestion' is common; 80% of the population will say they have had indigestion at some time.

'Alarm' features are suggestive of serious diseases such as cancer. They are:

- dysphagia
- weight loss
- protracted vomiting
- anorexia
- haematemesis or melaena.

However among patients investigated for dysphagia, anorexia, weight loss and vomiting, only a minority have significant gastrointestinal pathology. Even among patients with a history of vomiting blood or melaena, organic disease is by no means invariable. Nonetheless, concern about missing treatable cancer is so great that most patients with alarm symptoms will be recommended to undergo endoscopy.

Ask the authors
www.studentconsult.com

FURTHER READING

Vakil N, Moayyedi P, Fennerty MB et al. Limited value of alarm features in the diagnosis of upper gastrointestinal malignancy: systematic review and meta-analysis. *Gastroenterology* 2006; **131**(2): 390–401.

Dysphagia and odynophagia

These symptoms are described on page 249.

Vomiting

Vomiting centres are located in the lateral reticular formation of the medulla and are stimulated by the chemoreceptor trigger zones (CTZs) in the floor of the fourth ventricle, and also by vagal afferents from the gut. These zones are directly stimulated by drugs, motion sickness and metabolic causes.

There are usually three phases:

- nausea – a feeling of wanting to vomit, often associated with autonomic effects including salivation, pallor and sweating
- retching – a strong involuntary unproductive effort to vomit associated with abdominal muscle contraction
- vomiting – the expulsion of gastric contents through the mouth.

Many gastrointestinal (and non-gastrointestinal) conditions are associated with vomiting (Table 6.1).

Large volumes of vomit and sometimes projectile vomiting suggest gastric outlet or upper intestinal obstruction. *Faeculent vomit* suggests low intestinal obstruction or the presence of a gastrocolic fistula.

Nausea and frequent regurgitation of small amounts of gastric contents without other abdominal symptoms are frequently psychogenic.

Haematemesis is vomiting fresh or altered blood ('coffee-grounds') (see p. 267).

Early-morning nausea and vomiting is seen in pregnancy, alcohol dependence and some metabolic disorders (e.g. uraemia).

Flatulence

This term describes excessive wind. It is used to indicate belching, abdominal distension, gurgling and the passage of flatus per rectum. Swallowing air (aerophagia) is described on page 311. Some of the swallowed air passes into the intestine where most of it is absorbed, but some remains to be passed rectally. Intestinal bacterial breakdown of food also produces a small amount of gas. Rectal flatus thus consists of nitrogen, carbon dioxide, hydrogen and methane. It is normal to pass rectal flatus up to 20 times per day.

Table 6.1 Causes of vomiting – some examples

Any gastrointestinal disease

Infections
Viruses (influenza, norovirus)
Bacterial (pertussis, urinary infection)

Central nervous system disease
Raised intracranial pressure
Vestibular disturbance
Migraine

Metabolic
Uraemia
Hypercalcaemia

Diabetic ketoacidosis

Drugs
Antibiotics
Chemotherapy
Digoxin
Immunotherapy
Incretins
Levodopa
Opiates

Reflex
Myocardial infarction
Biliary colic

Psychogenic

Pregnancy

Alcohol excess

Diarrhoea and constipation

These are common complaints and in the community are not usually due to serious disease. They are described in detail on pages 305 and 295 respectively. Some simple rules are useful; for example, a single episode of diarrhoea is commonly due to dietary indiscretion or anxiety; watery stools of large volume are always due to organic disease; bloody diarrhoea usually implies colonic disease; acute diarrhoea lasting 2–5 days is most often due to an infective cause. Stool cultures can be useful but not always possible.

Patients often consider themselves constipated if their bowels are not open on most days, though normal stool frequency is very variable, from 3 times daily to 3 times a week. The difficult passage of hard stool is also regarded as constipation, irrespective of stool frequency. True constipation with hard stools is rarely due to organic colonic disease.

Abdominal pain

Pain is stimulated mainly by the stretching of smooth muscle or organ capsules. Severe acute abdominal pain can be due to a large number of gastrointestinal conditions, and normally presents as an emergency (see p. 314). An apparent 'acute abdomen' can occasionally be due to referred pain from the chest, as in pneumonia, or to metabolic causes, such as diabetic ketoacidosis or porphyria.

Clinical enquiry in patients with abdominal pain should include:

- site, intensity, character, duration and frequency of the pain
- aggravating and relieving factors
- associated symptoms, including non-gastrointestinal symptoms.

Upper abdominal pain

Epigastric pain is very common; it is often a dull ache, but sometimes sharp and severe and may be related to food intake. Although functional dyspepsia is the commonest diagnosis, the symptoms of peptic ulcer disease can be identical. *Heartburn* – a burning pain behind the sternum – is a common symptom of reflux.

Right hypochondrial pain may originate from the gall bladder or biliary tract. Biliary pain can also be epigastric. *Biliary pain* is typically intermittent, lasts a few hours and remits spontaneously to recur weeks or months later. Hepatic congestion (e.g. in hepatitis or cardiac failure) and sometimes peptic ulcer can present with pain in the right hypochondrium. Chronic, persistent or constant pain in the right (or left) hypochondrium in a well-looking patient is a frequent functional symptom; this chronic pain is not due to gall bladder disease (see p. 368).

Lower abdominal pain

Acute pain in the left iliac fossa is usually colonic in origin (e.g. acute diverticulitis). Chronic pain is most commonly associated with functional bowel disorders.

Lower abdominal pain in women occurs in a number of gynaecological disorders and the differentiation from GI disease may be difficult.

Persistent pain in the right iliac fossa over a long period is not due to chronic appendicitis and is most commonly functional.

Proctalgia fugax is a severe pain deep in the rectum that comes on suddenly but lasts only for a short time. It is not due to organic disease.

Abdominal wall pain

Persistent abdominal pain with localized tenderness which is not relieved by tensing the abdominal muscles appears to arise from the abdominal wall itself. Causes are thought to include nerve entrapment, external hernias and entrapment of internal viscera (commonly omentum) within traumatic ruptures of abdominal wall musculature.

Anorexia and weight loss

Anorexia describes reduced appetite. It is common in systemic disease and may be seen in psychiatric disorders, particularly anorexia nervosa. Anorexia often accompanies cancer but is usually a late symptom and not of diagnostic help. Weight loss is almost always due to reduced food intake and is a frequent accompaniment of gastrointestinal diseases. Weight loss in malabsorption is primarily due to anorexia. Weight loss with a normal or increased dietary intake occurs with hyperthyroidism and other catabolic states. Weight loss should always be assessed objectively as patients' impressions are unreliable. Appetite and satiety are described on page 229.

CLINICAL EXAMINATION

A systematic general examination is performed, starting with the hands, skin and nutritional status and noting the presence of anaemia or jaundice. Detailed examination of the gastrointestinal tract starts with the mouth and tongue, before examining the abdomen with the patient lying flat.

Examination of the abdomen

Inspection

Abdominal distension, whether due to flatus, fat, fetus, fluid or faeces, must be looked for. Intermittent distension is most commonly a feature of functional bowel disorders.

Palpation

Some abdominal organs may be felt even when normal, usually in thin people, and such organs are usually only just palpable (Fig. 6.2). Reidel's lobe is an anatomical variant consisting of a palpable enlargement of the lateral portion of the right lobe of the liver. Any palpable mass is carefully felt to evaluate its size, shape and consistency and whether it moves with respiration, to decide which organ is involved.

Figure 6.3a (p. 245) shows a normal CT scan at the level of T12. The hernial orifices should be examined if intestinal obstruction is suspected.

A succussion splash suggests gastric outlet obstruction if the patient has not drunk for 2–3 hours; the splash of fluid in the stomach can be heard with a stethoscope laid on the abdomen when the patient is moved.

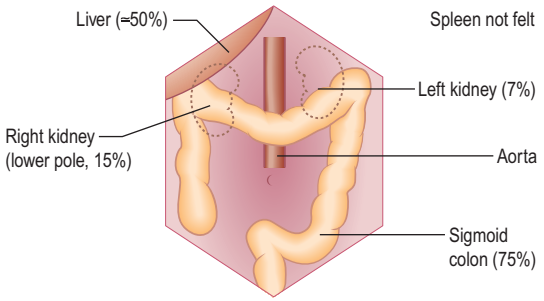


Fig. 6.2 **The organs sometimes palpable in thin subjects.**

Percussion

Abdominal percussion detects the areas of dullness caused by the liver and spleen, ascites or over masses. It can also detect a full bladder. *Ascites* is a term for excess fluid in the peritoneal cavity. It is detected clinically by central abdominal resonance due to gas within small bowel loops with dullness in the flanks which shifts when the patient lies on their side. This 'shifting dullness' is a reliable physical sign, but 1–2 L of fluid must be present. A large ovarian cyst can sometimes produce an enlarged abdomen, but the dullness is more centrally placed than in ascites.

Auscultation

Auscultation is not of great value in abdominal disease, except in the acute abdomen (see p. 314). Abdominal bruits are often present in normal subjects and are rarely clinically significant.

Examination of the rectum and sigmoid colon

A digital examination of the rectum should be performed in all patients with a change in bowel habit, rectal bleeding and prior to proctoscopy or sigmoidoscopy.

- *Proctoscopy* (Practical Box 6.1) is performed in all patients with a history of bright red rectal bleeding to look for anorectal pathology such as haemorrhoids; a rigid sigmoidoscope is too narrow and long to enable adequate examination of the anal canal.
- *Sigmoidoscopy* is part of the routine in-hospital examination in all cases of diarrhoea and in patients with lower abdominal symptoms such as a change in bowel habit or rectal bleeding. The rigid sigmoidoscope allows inspection of a maximum of 20–25 cm of distal colon.
- *Flexible sigmoidoscopy (FS)* (60 cm) allows a more extensive examination, usually up to the splenic flexure, and often with less discomfort. It can be performed in the outpatient department after evacuation of the distal colon using an enema or suppository. FS is useful in patients with increased stool frequency or looseness or rectal bleeding. Most rectal bleeding is due to benign ano-rectal disease (haemorrhoids or fissure-in-ano). Up to 60% of colonic neoplasms occur within the range of FS (see Fig. 6.47, p. 304). It can also be used to biopsy or remove lesions in the sigmoid area seen radiologically, and for the follow-up of patients with distal colitis. If the presenting complaint is constipation with hard stools FS has the same rate of abnormality as in a control population.

Practical Box 6.1 Sigmoidoscopy and proctoscopy

Sigmoidoscopy

- The technique using a 25-cm rigid sigmoidoscope is easy to learn, provides valuable information and is safe in competent hands
- No bowel preparation is required
- Explain to the patient the nature of the procedure, obtain consent
- The technique is relatively painless. In the irritable bowel syndrome, the patient's pain is often reproduced by air insufflation

1. Rectal examination is initially performed
2. The sigmoidoscope is passed into the anus, pointing towards the symphysis pubis. The obturator is removed, and the instrument passed under direct vision to the rectosigmoid junction and beyond if possible (using air insufflation)
3. The mucosa of the anus and rectum is inspected. The normal mucosa is shiny with superficial vessels and no contact bleeding
4. Biopsies can be taken of any lesions that are seen or from apparently normal-looking mucosa, which occasionally shows histological evidence of inflammation

Proctoscopy

1. The proctoscope is passed into the anus and the obturator is removed
2. The patient strains down as the proctoscope is removed
3. Haemorrhoids are seen as purplish veins in the left lateral, right posterior or right anterior positions
4. Fissures may also be seen, but pain often prevents the procedure from being performed

Stool examination

It is occasionally useful to confirm a patient's account (e.g. passing of blood or steatorrhoea). The shape and size may be helpful (e.g. 'rabbit dropping' or ribbon-like stools in the irritable bowel syndrome). Stool charts for recording weight and frequency of defecation are useful in inpatients to follow the progress of diarrhoea.

INVESTIGATIONS

Routine haematology and biochemistry, followed by radiology and endoscopy, are the principal investigations. The investigation of small bowel disease is discussed in more detail on page 276. Manometry is mainly used in oesophageal disease (see p. 250) and anorectal disorders (see p. 295).

Imaging

Full clinical information must be provided before the examination, and ideally the images should be reviewed with the radiologist.

Plain X-rays of the chest and abdomen are chiefly used in the investigation of an acute abdomen. Interpretation depends on analysis of gas shadows inside and outside the bowel. Plain films are particularly useful where obstruction or perforation is suspected, in assessment of the severity and extent of acute colitis and to assess faecal loading in constipation. Calcification may be seen in gall bladder stones and in chronic pancreatitis, though CT is more sensitive for both.

Ultrasound involves no radiation and is the first-line investigation for abdominal distension, e.g. ascites, mass or suspected inflammatory conditions. It can show dilated fluid-filled loops of bowel in obstruction, and thickening of the bowel wall. It can be used to guide biopsies or percutaneous drainage. In the acute abdomen, ultrasound can diagnose cholecystitis, appendicitis, enlarged mesenteric glands and other inflammatory conditions.

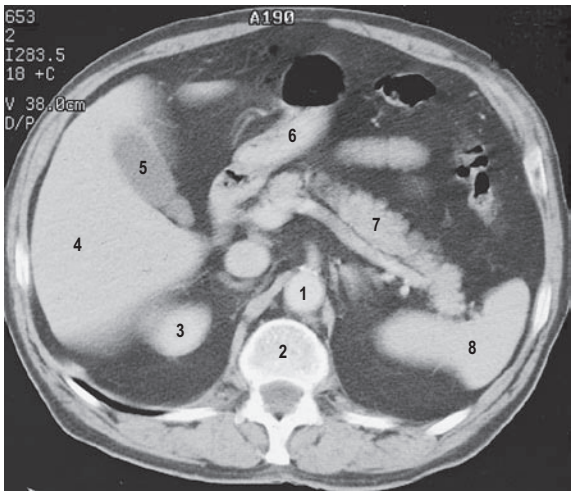
- *Endoscopic ultrasound (EUS)* is performed with a gastroscope incorporating an ultrasound probe at the tip. It is used diagnostically for lesions in the oesophageal or gastric wall including the detailed TNM staging of oesophageal/gastric cancer and for the detection and biopsy of pancreatic tumours and cysts.
- *Endoanal and endorectal ultrasonography* are performed to define the anatomy of the anal sphincters (see p. 298), to detect perianal disease and to stage superficial rectal tumours.

Computed tomography involves high-dose radiation. Modern multi-slice fast scanners and techniques involving intraluminal and intravenous contrast provide accurate diagnoses in many cases. Intraluminal contrast may be used and may be positive (Gastrografin or Omnipaque) or negative (usually water). The bowel wall and mesentery are well seen after intravenous contrast especially with negative intraluminal contrast. Clinically unsuspected diseases of other abdominal organs are quite often also revealed (Fig. 6.3a).

- *CT* is widely used as a first-line investigation for the acute abdomen. CT is sensitive for small volumes of gas from a perforated viscus as well as leakage of contrast from the gut lumen.
- Inflammatory conditions such as abscesses, appendicitis, diverticulitis, Crohn's disease and its complications are well demonstrated. In high-grade bowel obstruction CT is usually diagnostic of both the presence and the cause of the obstruction.
- CT is widely used in cancer staging and as guidance for biopsy of tumour or lymph nodes.
- *CT pneumocolon/CT colonography* (virtual colonoscopy) after CO_2 insufflation into a previously cleansed colon provides an alternative to colonoscopy for diagnosis of colon mass lesions (Fig. 6.3b). It is being evaluated as a screening test for colon pathology with sensitivities of over 90% for >10 mm polyps.
- *Unprepared CT* is a good test for colon cancer in the frail (often elderly) patient who would have problems with bowel preparation.

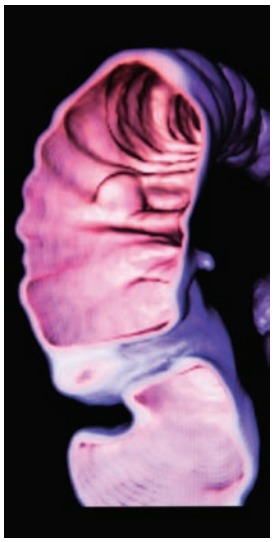
Magnetic resonance imaging. MRI uses no radiation and is particularly useful in the evaluation of rectal cancers and abscesses and fistulae in the perianal region. It is also useful in small bowel disease and in hepatobiliary and pancreatic disease.

Positron emission tomography (PET) relies on detection of the metabolism of fluoro-deoxyglucose. It is used for staging oesophageal, gastric and colorectal cancer and in the detection of metastatic and recurrent disease. PET/CT adds additional anatomical information.



(a)

(b)



(c)

Fig. 6.3 **(a) CT scan of the normal abdomen at the level of T12.** 1, aorta; 2, spine; 3, top of right kidney; 4, liver; 5, gall bladder; 6, stomach (containing air); 7, pancreas; 8, spleen. **(b) CT cross-sectional (2-dimensional) image of a colonic polyp on a long stalk.** The colon has been emptied as for visual colonoscopy. The pedunculated polyp and its stalk show enhancement after intravenous contrast. **(c) 3D reconstruction of part of the colon** (false colour) generated by a computer programme from multiple axial CT images. A sessile polypoid lesion is shown. (b and c Courtesy of Dr Paul Jenkins.)

Contrast studies

- *Barium swallow* examines the oesophagus and proximal stomach. Its main use is for dysphagia. Barium is swallowed in both the upright and prone positions. Anatomical lesions and motility disorders can be investigated. Reflux, as demonstrated by the retrograde flow of barium from the stomach into the oesophagus, is best observed with the patient tipped head down. It can be provoked by asking the patient to drink water (water siphon test) and by distension of the stomach with gas generated from swallowed effervescent granules. Observing the passage of bread or a marshmallow coated in barium can be useful in evaluating patients with dysphagia.
- *Double-contrast barium meal* examines the oesophagus, stomach and duodenum. Barium is given to produce mucosal coating. Effervescent granules producing carbon dioxide in the stomach create a double contrast between gas and barium. This test has a high accuracy for the detection of significant pathology – ulcers and cancer – but requires good technique. Gastroscopy is a more sensitive test for small superficial mucosal lesions and for bleeding, and enables biopsy of suspicious areas.
- *Small bowel meal or follow-through* specifically examines the small bowel. Ingested barium passes through the small bowel into the right colon, usually taking about an hour. The fold pattern and calibre of the small bowel are assessed. Specific views of the terminal ileum are usually obtained using compression to separate the loops of bowel and particularly to identify early changes in patients with suspected Crohn's disease.
- *Small bowel enema (enteroclysis)* is an alternative specific technique for small bowel examination. A tube is passed into the duodenum and a large volume of dilute barium is introduced. It is particularly used to demonstrate strictures or adhesions when there is suspicion of intermittent obstruction.
- *Barium enema* examines the colon and is used for altered bowel habit. Colonoscopy has largely replaced this examination for rectal bleeding, polyps and inflammatory bowel disease. The colon is first thoroughly cleansed with oral laxative preparations. Barium and air or CO_2 are insufflated via a rectal catheter and double-contrast views obtained of the entire colon. Rectal examination and sigmoidoscopy should precede this examination. Some patients find the examination rather difficult to tolerate, particularly the elderly, frail or immobile in whom other tests such as CT should be used.
- *Absorbable water soluble (Gastrografin or Omnipaque) contrast agents* should be used in preference to barium when perforation is suspected anywhere in the gut.

Radioisotopes

Radionuclides are used to a varying degree depending on local enthusiasm, availability and expertise. Some more common indications and techniques are to:

- detect urease activity of *Helicobacter pylori* – ^{13}C urea breath test (see p. 260)
- assess oesophageal reflux – gamma camera scan after oral [^{99m}Tc]technetium–sulphur colloid
- measure rate of gastric emptying – sequential gamma camera scans after oral [^{99m}Tc]technetium–sulphur colloid or ^{111}In-DTPA (indium-labelled diethylene triamine penta-acetic acid)
- demonstrate a Meckel's diverticulum – gamma camera scan after i.v. [^{99m}Tc]pertechnetate, which has affinity for gastric mucosa
- assess extent of inflammation and presence of inflammatory collections in inflammatory bowel disease – gamma camera scan after i.v. ^{99m}Tc-HMPAO (hexamethylpropylene amine oxime) labelled white cells

- evaluate neuroendocrine tumours and their metastases – gamma camera scan after i.v. radiolabelled octreotide or MIBG (meta-iodobenzylguanidine)
- assess obscure gastrointestinal bleeding:
  - measure faecal loss of i.v. ^{51}Cr labelled red cells
  - gamma camera abdominal scan after i.v. injection of red cells labelled with ^{99m}Tc (only useful if the bleeding is at more than 2 mL/minute)
- measure albumin loss in the stools (in *protein-losing enteropathy*) – following albumin labelled in vivo with i.v. $^{51}CrCl_3$ (this test has been replaced by the measurement of the intestinal clearance of α_1 antitrypsin)
- assess bile salt malabsorption (in patients with unexplained diarrhoea) – gamma camera scan to measure both isotope retention and faecal loss of orally administered ^{75}Selenium-homocholic acid taurine (SeHCAT) (see p. 308)
- detect bacterial overgrowth in the small bowel – measure $^{14}CO_2$ in breath following oral ^{14}C glycocholic acid.

Endoscopy

Video endoscopes have largely replaced fibreoptic instruments. They have a video chip and lens assembly mounted in the tip of the instrument to relay colour images to a television monitor. The tip of the endoscope can be angulated in all directions. Channels in the instrument are used for air insufflation, water injection, suction and for the passage of accessories such as biopsy forceps or brushes for obtaining tissue, snares for polypectomy and needles for injection therapies. Permanent photographic or video records of the procedure can be obtained.

- *Oesophagogastroduodenoscopy (OGD*, 'gastroscopy') is often used as the investigation of choice for upper GI disorders by gastroenterologists because of easy access and the possibility of therapy and mucosal biopsy. Findings include reflux oesophagitis, gastritis, ulcers and cancer. Relative contraindications include severe chronic obstructive pulmonary disease, a recent myocardial infarction, or severe instability of the atlanto-axial joints. The mortality for diagnostic endoscopy is 0.001% with significant complications in 1:1000, usually when performed as an emergency (e.g. GI haemorrhage).
- *Colonoscopy* allows good visualization of the whole colon and terminal ileum. Biopsies can be obtained and polyps removed. The success rate for reaching the terminal ileum should be at least 90% after training. Cancer, polyps and diverticular disease are the commonest significant findings. Perforation occurs in 1:1000 examinations and in 2% after polypectomy. The mortality is 0.02% for diagnostic colonoscopy (Practical Box 6.2).
- *Balloon enteroscopy* can examine the small bowel from the duodenum to the ileum but specialized enteroscopes are needed. The indications are limited (mainly gastrointestinal blood loss), the technique is time-consuming and the instruments are expensive so it is available only in a few centres.
- *Video (wireless) capsule endoscopy* is now used for the evaluation of obscure GI bleeding (after negative gastroscopy and colonoscopy) and for the detection of small bowel tumours and occult inflammatory bowel disease.

Practical Box 6.2 Gastroscopy and colonoscopy

- Explain the procedure to the patient, including benefits and risks
- Discuss need for sedation
- Obtain written informed consent

Gastroscopy

1. Patient should be fasted for at least 4 hours
2. Give oxygen and monitor oxygen saturation with an oximeter
3. Give lidocaine throat spray or sedation (midazolam ± opiate if required)
4. Pass the gastroscope to the duodenum under direct vision
5. Examine during insertion and withdrawal
6. Gastroscopy takes 5–15 minutes, depending on the indication and findings
7. Withhold fluid and food until LA/sedation wears off
8. Complications are rare: beware of over-sedation, perforation and aspiration
9. Patient must be accompanied home if sedation given

Colonoscopy

1. Stop oral iron a week before the procedure
2. Restrict diet to low residue foods for 48 hours; clear fluids only for 24 hours
3. Use local bowel cleansing regime, usually starting 24 hours beforehand (e.g. two sachets of Picolax and 2–4 bisacodyl tablets, or Moviprep 2 litres, or KleanPrep 4 litres, or local alternative; more if constipated)
4. Give oxygen and monitor O_2 levels
5. Give sedation (midazolam ± opiate) if required by patient
6. Pass the colonoscope to the caecum or ileum under direct vision
7. Detailed examination during withdrawal
8. Colonoscopy takes 15–30 minutes, depending on the colon anatomy, indication and findings
9. Withhold fluid and food until sedation wears off
10. Observe patient for at least an hour after sedation given
11. Complications are rare: beware of over-sedation, perforation and aspiration
12. Patient must be accompanied home if sedation given

THE MOUTH AND SALIVARY GLANDS

The oral cavity extends from the lips to the pharynx and contains the tongue, teeth and gums. Its primary functions are mastication, swallowing and speech. Problems in the mouth are extremely common and, although they may be trivial, they can produce severe symptoms. Poor dental hygiene is often a factor.

Stomatitis is inflammation in the mouth from any cause, such as ill-fitting dentures. Angular stomatitis is inflammation of the corners of the mouth.

The '*burning mouth syndrome*' consists of a burning sensation with a clinically normal oral mucosa. It occurs more commonly in middle-aged and elderly females. It is probably psychogenic in origin. *Halitosis* is a common symptom and is due to poor oral hygiene, anxiety (often when halitosis is more apparent to the patient than real) and rarer causes, e.g. oesophageal stricture and pulmonary sepsis.

The mouth

Recurrent aphthous ulceration

Idiopathic aphthous ulceration is common and affects up to 25% of the population. Recurrent painful round or ovoid

mouth ulcers are seen with inflammatory halos. They are commoner in females and non-smokers, usually appear first in childhood and tend to reduce in number and frequency before age 40. Other family members may be affected. There is no sign of systemic disease. *Minor aphthous ulcers* are the most common, are less than 10 mm diameter, have a grey/white centre with a thin erythematous halo and heal within 14 days without scarring. They rarely affect the dorsum of the tongue or hard palate. *Major aphthous ulcers* (more than 10 mm diameter) often persist for weeks or months and heal with scarring.

The cause is not known. Deficiencies of iron, folic acid or vitamin B_{12} (with or without gastrointestinal disorders) are sometimes found but are not causally linked. Secondary causes, e.g. Crohn's disease, should be excluded.

There are no specific effective therapies. Sufferers should avoid oral trauma and acidic foods or drinks which cause pain. Topical (1% triamcinolone) or systemic corticosteroids may lessen the duration and severity of the attacks. Chlorhexidine gluconate or tetracycline mouthwash, dapsone, colchicine, thalidomide and azathioprine have all been used with variable effect.

Other causes

See Table 6.2.

Neoplasia (squamous cell carcinoma)

Malignant tumours of the mouth account for 1% of all malignant tumours in the UK. The majority develop on the floor of the mouth or lateral borders of the tongue. Early lesions may be painless, but advanced tumours are easily recognizable as hard indurated ulcers with raised and rolled edges. Aetiological agents include tobacco, heavy alcohol consumption and the areca nut. Premalignant lesions include leucoplakia (single adherent white patch), lichen planus, submucous fibrosis and erythroplakia (a red patch). The previous male predominance has declined. *Treatment* is by surgical excision but may require extensive neck dissection to remove involved lymph nodes and/or radiotherapy.

Oral white patches

Transient white patches are either due to *Candida* infection or are very occasionally seen in systemic lupus erythematosus. Oral candidiasis in adults is seen in seriously ill or immunocompromised patients, diabetics and following therapy with broad-spectrum antibiotics or inhaled steroids. Local causes include mechanical, irritative or chemical trauma from drugs (e.g. aspirin).

Persistent white patches can be due to *leucoplakia*, which is associated with alcohol and (particularly) smoking, and is premalignant. A biopsy should always be taken; histology shows alteration in the keratinization and dysplasia of the epithelium. Treatment is unsatisfactory. Isotretinoin possibly reduces disease progression. *Oral lichen planus* presents as white striae, which can rarely extend into the oesophagus.

Table 6.2 Causes of mouth ulcers

Idiopathic aphthous ulceration (commonest)

Gastrointestinal disease
Inflammatory bowel disease
Coeliac disease

Infection
Viral – HSV, HIV, Coxsackie
Fungal – candidiasis
Bacterial – syphilis, tuberculosis

Systemic disease
Reactive arthritis (see p. 535)
Behçet's syndrome
Systemic lupus erythematosus

Trauma
e.g. Dental caries

Neoplasia
e.g. Squamous cell carcinoma

Drugs, e.g.
In erythema multiforme major
Chemotherapy, antimalarials

Skin disease
Pemphigoid
Pemphigus
Lichen planus

Oral pigmented lesions

Non-neoplastic lesions

Racial pigmentation is scattered and symmetrically distributed. Amalgam tattoo is the most common form of localized oral pigmentation and consists of blue-black macules involving the gingivae and results from dental amalgam sequestering into the tissues. Diseases causing pigmentation include Peutz–Jeghers syndrome and Addison's disease. Heavy metals, such as lead, bismuth and mercury, and drugs (e.g. phenothiazines and antimalarials) all cause pigmentation of the gums.

Neoplastic lesions

These include melanotic naevi on the hard palate and buccal mucosa. These are rarer in the mouth than on the skin. Malignant melanomas are rare, more common in males, and occur mainly on the upper jaw. The 5-year survival is only 5%.

The tongue

The tongue may be involved in stomatitis with similar lesions to those described above.

Glossitis is a red, smooth, sore tongue associated with B_{12}, folate or iron deficiency. It is also seen in infections due to *Candida* and in riboflavin and nicotinic acid deficiency.

A black hairy tongue is due to a proliferation of chromogenic microorganisms causing brown staining of elongated filiform papillae. The causes are unknown, but heavy smoking and the use of antiseptic mouthwashes have been implicated.

A geographic tongue is an idiopathic condition occurring in 1–2% of the population and may be familial. There are erythematous areas surrounded by well-defined, slightly raised irregular margins. The lesions are usually painless and the patient should be reassured.

The gums

The gums (gingivae) are the mucous membranes covering the alveolar processes of the mandible and the maxilla.

Chronic gingivitis follows the accumulation of bacterial plaque. It resolves when the plaque is removed. It is the most common cause of bleeding gums.

Acute (necrotizing) ulcerative gingivitis ('Vincent's angina') is characterized by the proliferation of spirochaete and fusiform bacteria associated with poor oral hygiene and smoking. Treatment is with oral metronidazole 200 mg three times daily for 3 days, improved oral hygiene and chlorhexidine gluconate mouthwash.

Desquamative gingivitis is a clinical description of smooth, red atrophic gingivae caused by lichen planus or mucous membrane pemphigoid. The diagnosis is confirmed by biopsy.

Gingival swelling may be due to inflammation or fibrous hyperplasia. The latter may be hereditary (gingival fibromatosis) or associated with drugs (e.g. phenytoin, ciclosporin, nifedipine). Inflammatory swellings are seen in pregnancy, gingivitis and scurvy. Swelling due to infiltration is seen in acute leukaemia and Wegener's granulomatosis.

The teeth

Dental caries causes tooth decay and 'cavities'. The main cause in man is *Streptococcus mutans* which is cariogenic only in the presence of dietary sugar. Dental caries can progress to pulpitis and pulp necrosis, and spreading infection can cause dentoalveolar abscesses. If there is soft tissue swelling, antibiotics (e.g. amoxicillin or metronidazole) should be prescribed prior to dental intervention.

Erosion of the teeth can result from exposure to acid (e.g. in bulimia nervosa) or, very occasionally, in patients with gastro-oesophageal reflux disease.

Oral manifestations of HIV infection

In the UK, 60% of HIV-infected patients have characteristic oral lesions. Lesions strongly associated with HIV infection include candidiasis (with erythema and/or white exudates), erythematous candidiasis, oral hairy leucoplakia, Kaposi's sarcoma, non-Hodgkin's lymphoma, necrotizing ulcerative gingivitis and necrotizing ulcerative periodontitis and are described elsewhere.

All oral lesions are much less common since the introduction of HAART (see p. 192).

FURTHER READING

Lehner T. The mouth and salivary glands. In: Cox TM, Firth JD, Warrell DA (eds) *Oxford Textbook of Medicine,* 4th edn. Oxford: Oxford University Press, 2003.

Scully C. Aphthous ulceration. *New England Journal of Medicine* 2006; **355**: 165–172.

Selwitz RH, Ismail AI, Pitts NB. Dental caries. *Lancet* 2007; **369**: 51–59.

THE SALIVARY GLANDS

Excessive salivation (*ptyalism*) occurs prior to vomiting, but may be secondary to other intraoral pathology. It can be psychogenic.

Dry mouth (*xerostomia*) can result from a variety of causes:

- Sjögren's syndrome
- drugs (e.g. antimuscarinic, antiparkinsonian, antihistamines, lithium, monoamine oxidase inhibitors, tricyclic and related antidepressants, and clonidine)
- radiotherapy
- psychogenic causes
- dehydration, shock and renal failure.

The principles of management are to preserve what flow remains, stimulate flow and replace saliva (glycerine and lemon mouthwash and artificial saliva).

Sialadenitis

Acute sialadenitis is viral (mumps) or bacterial. Bacterial sialadenitis is a painful ascending infection with *Staphylococcus aureus*, *Streptococcus pyogenes* and *Strep. pneumoniae*, usually secondary to secretory failure. Pus can be expressed from the affected duct.

Kikuchi disease is a rare, self-limiting condition which presents with cervical lymphadenopathy, fever and sometimes parotid involvement. The cause is unknown.

Salivary duct obstruction due to calculus

Obstruction to salivary flow is usually due to a calculus. There is a painful swelling of the submandibular gland after eating and stones can sometimes be felt in the floor of the mouth. Plain X-ray films and sialography will show the calculus; removal of the obstruction by sialendoscopy gives complete relief.

Sarcoidosis (see also p. 868)

Sarcoidosis can involve the major salivary glands as part of *Heerfordt's syndrome* (uveo-parotid fever). When combined with lacrimal gland enlargement it is known as the *Mikulicz syndrome*.

Neoplasms

Salivary gland neoplasms account for 3% of all tumours world-wide. The majority occur in the parotid gland. The pleomorphic adenoma is the most common and 15% of these undergo malignant transformation. Malignant tumours classically result in lower motor neurone 7th cranial nerve signs. Recurrence following surgical excision is common.

THE PHARYNX AND OESOPHAGUS

Structure and function

The oesophagus is a muscular tube approximately 20 cm long that connects the pharynx to the stomach just below the diaphragm. In the upper portion, both the outer longitudinal layer and inner circular muscle layers are striated. In the lower two-thirds of the oesophagus, including the thoracic and abdominal parts containing the lower oesophageal sphincter, both layers are composed of smooth muscle. Diseases of the striated muscle of the upper oesophagus include polymyositis and myasthenia gravis, whereas scleroderma and achalasia involve the distal smooth muscle portion.

The oesophagus is lined by stratified squamous epithelium which extends distally to the squamocolumnar junction where the oesophagus joins the stomach, recognized endoscopically by a zig-zag ('Z') line.

The oesophagus is separated from the pharynx by the *upper oesophageal sphincter*, which is normally closed due to tonic activity of the nerves supplying the cricopharyngeus. The *lower oesophageal sphincter* (LOS) consists of a 2–4 cm zone in the distal end of the oesophagus that has a high resting tone and, assisted by the diaphragmatic sphincter, is largely responsible for the prevention of gastric reflux.

During swallowing, the bolus of food is voluntarily moved from the mouth to the pharynx. Contraction of the pharynx initiates reflex relaxation of the upper sphincter and food enters the oesophagus. Starting in the pharynx, a primary peristaltic wave progresses down the whole oesophagus. The LOS relaxes when swallowing is initiated, before the arrival of the food bolus carried by the peristaltic wave.

The smooth muscle of the thoracic oesophagus and lower oesophageal sphincter is supplied by vagal autonomic motor nerves consisting of extrinsic preganglionic fibres and

intramural postganglionic neurones in the myenteric plexus (Fig. 6.4). There are parallel excitatory and inhibitory pathways. The excitatory pathway consisting of preganglionic and postganglionic neurones acts via acetylcholine and also substance P at the postganglionic synapse. The inhibitory pathway comprises preganglionic cholinergic and postganglionic nitrergic neurones (NANC fibres) that release vasoactive intestinal peptide, nitric oxide, ATP and substance P.

Swallowing

Between swallows the muscles of the oesophagus are relaxed except for those of the sphincters. The upper oesophageal sphincter remains closed because of tonic activity of its innervation. The lower oesophageal sphincter remains closed due to the unique property of its muscle, illustrated by the tonic contraction of the denervated LOS in achalasia. **T**ransient **l**ower **o**esophageal **s**phincter **r**elaxations (TLESRs) occur in normal people, and more in reflux sufferers.

Swallowing is mediated by a complex reflex involving a swallowing centre in the dorsal motor nucleus of the vagus in the brainstem. Once activated, the swallowing centre neurones send pre-programmed discharges of inhibition followed by excitation to the motor nuclei of the cranial nerves. This results in initial relaxation, followed by distally progressive activation of neurones to the oesophageal smooth muscle and LOS. Pharyngeal and oesophageal peristalsis mediated by this swallowing reflex causes *primary peristalsis*. *Secondary peristalsis* arises as a result of stimulation by a food bolus in the lumen, mediated by a local intra-oesophageal reflex. *Tertiary contractions* indicate pathological non-propulsive contractions resulting from aberrant activation of local reflexes within the myenteric plexus.

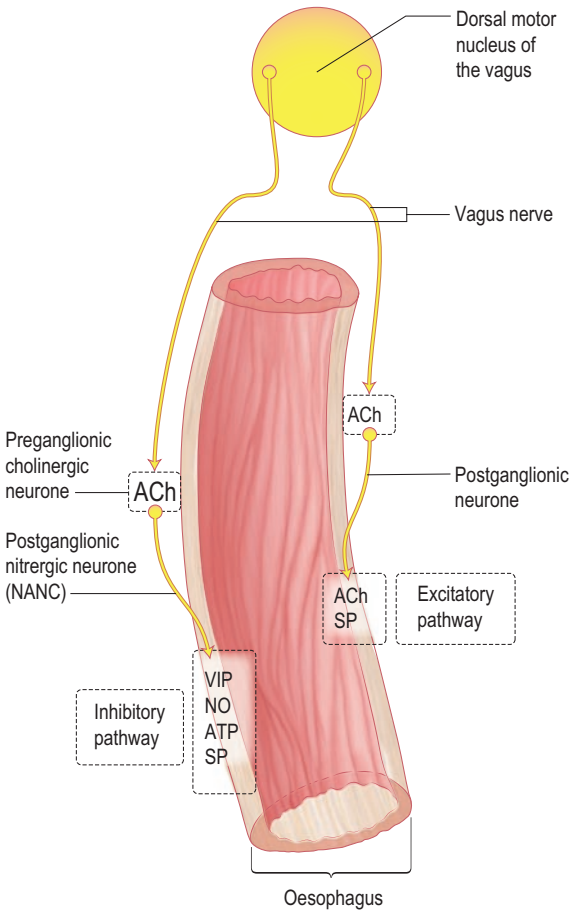


Fig. 6.4 **Innervation of the oesophagus.** The excitatory pathway consists of vagal preganglionic neurones releasing acetylcholine (ACh), connecting to postganglionic neurones that release ACh and substance P. The inhibitory pathway consists of vagal preganglionic neurones releasing ACh, connecting to postganglionic neurones that release nitric oxide (NO), vasoactive intestinal peptide (VIP), adenosine triphosphate (ATP) and substance P (SP).

Symptoms of oesophageal disorders

Major oesophageal symptoms are:

- ***Dysphagia***. This term means a sensation of obstruction during the passage of liquid or solid through the pharynx or oesophagus, i.e. within 15 seconds of food leaving the mouth. The characteristics of the progression of dysphagia to solids can be helpful, e.g. intermittent slow progression with a history of heartburn suggests a benign peptic stricture; relentless progression over a few weeks suggests a malignant stricture. The slow onset of dysphagia for solids and liquids at the same time suggests a motility disorder, e.g. achalasia (see p. 253). The causes are shown in Table 6.3.
- ***Odynophagia*** is pain during the act of swallowing and is suggestive of oesophagitis. Causes include reflux, infection, chemical oesophagitis due to drugs such as bisphosphonates or slow-release potassium, or associated with oesophageal stenosis.
- ***Substernal discomfort, heartburn***. This is a common symptom of reflux of gastric contents into the oesophagus. It is usually a retrosternal burning pain that can spread to the neck, across the chest, and when severe can be difficult to distinguish from the pain of ischaemic heart disease. It is often worst lying down at night when gravity promotes reflux, or on bending or stooping.
- ***Regurgitation*** is the effortless reflux of oesophageal contents into the mouth and pharynx. Uncommon in normal subjects, it occurs frequently in patients with gastro-oesophageal reflux disease or organic stenosis.

Table 6.3 Causes of dysphagia

Disease of mouth and tongue (e.g. tonsillitis)

Neuromuscular disorders
Pharyngeal disorders
Bulbar palsy (e.g. motor neurone disease)
Myasthenia gravis

Oesophageal motility disorders
Achalasia
Scleroderma
Diffuse oesophageal spasm
Presbyoesophagus
Diabetes mellitus
Chagas' disease

Extrinsic pressure
Mediastinal glands
Goitre
Enlarged left atrium

Intrinsic lesion
Foreign body
Stricture:
benign – peptic, corrosive
malignant – carcinoma
Lower oesophageal ring
Oesophageal web
Pharyngeal pouch

Signs of oesophageal disorders

The main sign of oesophageal disease is weight loss due to reduced food intake. Cervical lymphadenopathy with cancer is uncommon. Rarely a pharyngeal pouch may be seen to swell the neck during drinking.

Investigation of oesophageal disorders

- Barium swallow and meal.
- Oesophagoscopy.
- Manometry (Fig. 6.5) is performed by passing a catheter through the nose into the oesophagus and measuring the pressures generated within the region of the lower oesophageal sphincter (LOS) and body of oesophagus. It is the gold standard for the assessment of oesophageal motor activity. It is not a primary investigation and should be performed only when the diagnosis has not been achieved by history, barium radiology or endoscopy. Recordings are usually made over a short time period, or much more rarely for up to 24 hours.
- pH monitoring – 24-hour ambulatory monitoring uses a pH-sensitive probe positioned in the lower oesophagus and is used to identify acid reflux episodes (pH < 4). Catheter and implantable sensors are available; both are insensitive to alkali. Although only 5–10% of recorded acid reflux episodes are perceived by the patient, pH is a valuable means of relating episodes of acid reflux with patient's symptoms.
- Impedance uses a catheter to measure the resistance to flow of 'alternating current' in the contents of the oesophagus. Combined with pH it allows assessment of both acid and alkaline reflux which is useful if pain does not respond to a proton-pump inhibitor (PPI).
- Radiolabelled technetium (see p. 245) is used in some centres.

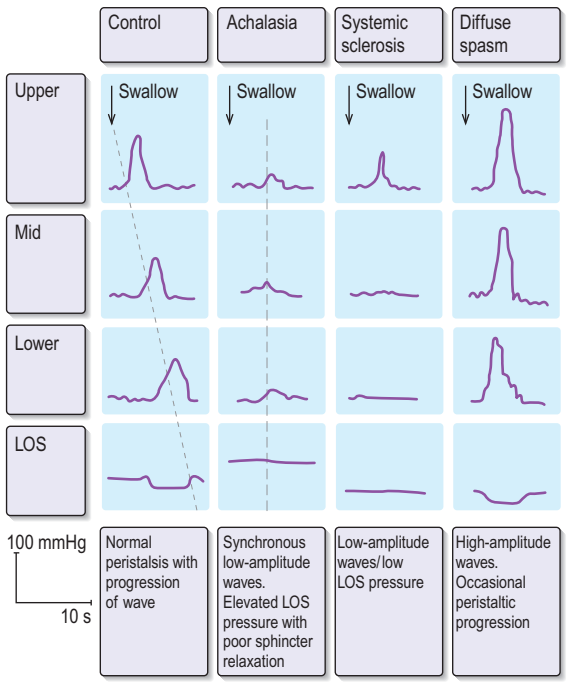


Fig. 6.5 **Oesophageal manometric patterns in normals and diseased states.** LOS, lower oesophageal sphincter. Dotted line shows peristaltic wave.

Acid and alkali infusion into the oesophagus (Bernstein test) to reproduce the pain is obsolete.

The 'Proton Pump Inhibitor Test' – an empirical trial of a treatment dose of PPI to observe the effect on heartburn or 'non-cardiac chest pain' – is practical and widely used. It has an 80% sensitivity and 74% specificity. It should not be used in patients with dysphagia until cancer has been excluded.

GASTRO-OESOPHAGEAL REFLUX DISEASE (GORD)

Pathophysiology (Fig. 6.6)

Small amounts of gastro-oesophageal reflux are normal. The lower oesophageal sphincter (LOS) in the distal oesophagus is in a state of tonic contraction and relaxes transiently to allow the passage of a food bolus (see p. 248). Sphincter pressure also increases in response to rises in intra-abdominal and intragastric pressures.

Other antireflux mechanisms involve the intra-abdominal segment of the oesophagus which acts as a flap valve. In addition, the mucosal rosette formed by folds of the gastric mucosa and the contraction of the crural diaphragm at the LOS acting like a pinchcock, prevent acid reflux. A large hiatus hernia can impair this mechanism. The oesophagus is also normally rapidly 'cleared' of refluxate by secondary peristalsis, gravity and salivary bicarbonate.

The clinical features of reflux occur when the antireflux mechanisms fail, allowing acidic gastric contents to make prolonged contact with the lower oesophageal mucosa. The sphincter relaxes transiently independently of a swallow (Transient Lower oEsophageal Sphincter Relaxation, TLESR) after meals and this is the cause of almost all reflux in normals and about two-thirds in GORD patients.

Oesophageal mucosal defence mechanisms

- *Surface.* Mucus and the unstirred water layer trap bicarbonate. This mechanism is a weak buffering mechanism compared to that in the stomach and duodenum.
- *Epithelium.* The apical cell membranes and the junctional complexes between cells act to limit diffusion of H^+ into the cells. In oesophagitis, the junctional complexes are damaged leading to increased H^+ diffusion and cellular damage.
- *Postepithelium.* Bicarbonate normally buffers acid in the cells and intracellular spaces. Hydrogen ions impair the growth and replication of damaged cells.

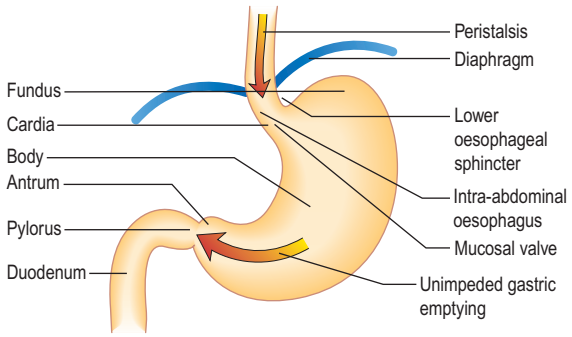


Fig. 6.6 **The main antireflux mechanisms – shown on the right.**

- *Sensory mechanisms.* Acid stimulates primary sensory neurones in the oesophagus by activating the vanilloid receptor-1 (VR1). This can initiate inflammation and release of proinflammatory substances from the tissue and produce pain. Pain can also be due to contraction of longitudinal oesophageal muscle.

Clinical features

Heartburn is the major feature. Factors associated with GORD are shown in Table 6.4.

The burning is aggravated by bending, stooping or lying down which promote acid exposure, and may be relieved by oral antacids. The patient complains of pain on drinking hot liquids or alcohol.

The correlation between heartburn and oesophagitis is poor. Some patients have mild oesophagitis but severe heartburn; others have severe oesophagitis without symptoms, and may present with a haematemesis or iron deficiency anaemia from chronic blood loss. Psychosocial factors are often determinants of symptom severity. Many patients erroneously ascribe their symptoms to their hiatus hernia (Box 6.1) but the symptoms are due to reflux.

Differentiation of cardiac and oesophageal pain can be difficult; 20% of cases admitted to a coronary care unit have GORD (Box 6.2). In addition to the clinical features, a trial of a PPI is always worthwhile and if symptoms persist, ambulatory pH and impedence monitoring should be performed.

Regurgitation of food and acid into the mouth occurs, particularly on bending or lying flat. Aspiration pneumonia is unusual without an accompanying stricture, but cough and asthma can occur and respond slowly (1–4 months) to PPI.

Table 6.4 Factors associated with gastro-oesophageal reflux

Pregnancy or obesity
Fat, chocolate, coffee or alcohol ingestion
Large meals
Cigarette smoking
Drugs – antimuscarinic, calcium-channel blockers, nitrates
Systemic sclerosis
After treatment for achalasia
Hiatus hernia

Box 6.1 Hiatus hernia

Sliding hiatus hernia

The oesophageal-gastro junction and part of the stomach 'slides' through the hiatus so that it lies above the diaphragm.

- Present in 30% of people over 50 years
- Produces no symptoms – any symptoms are due to reflux

Rolling or para-oesophageal hernia

Part of the fundus of the stomach prolapses through the hiatus alongside the oesophagus.

- The lower oesophageal sphincter remains below the diaphragm and remains competent
- Occasionally severe pain occurs due to volvulus or strangulation

Diagnosis and investigations

The clinical diagnosis can usually be made without investigation. Unless there are alarm signs, especially dysphagia (see p. 241), patients under the age of 45 years can safely be treated initially without investigations. If investigation is required, there are two aims:

- *Assess oesophagitis and hiatal hernia by endoscopy*. If there is oesophagitis (Fig. 6.7) or Barrett's oesophagus (see p. 252), reflux is confirmed.
- *Document reflux by intraluminal monitoring* (Fig. 6.8). 24-hour intraluminal pH monitoring or impedance combined with manometry is helpful if the PPI test is negative and should always be performed to confirm reflux before surgery. Excessive reflux is defined as a

Box 6.2 Features of the pain of gastro-oesophageal reflux and cardiac ischaemia

Reflux pain: burning, worse on bending, stooping or lying down

Seldom radiates to the arms
Worse with hot drinks or alcohol
Relieved by antacids

Cardiac ischaemic pain:

Gripping or crushing
Radiates to neck or left arm
Worse with exercise
Accompanied by dyspnoea

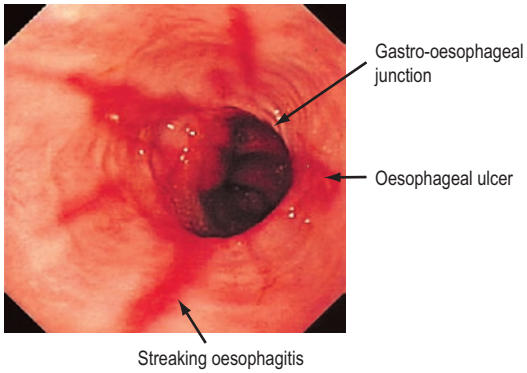


Fig. 6.7 **Oesophagitis seen on endoscopy.**

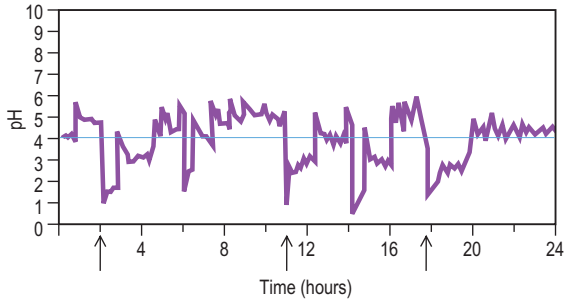


Fig. 6.8 **24-hour intraluminal pH monitoring.** Five reflux episodes (pH < 4) occurred, but only three gave symptoms (arrows).

pH < 4 for more than 4% of the time. There should also be a good correlation between reflux (pH < 4.0) and symptoms. It is also helpful to assess oesophageal dysmotility as a potential cause of the symptoms.

Treatment

Approximately half of patients with reflux symptoms in primary care can be treated successfully with simple antacids, loss of weight and raising the head of the bed at night. Precipitating factors should be avoided, with dietary measures, reduction in alcohol consumption and cessation of smoking. These measures are simple to say and difficult to carry out, but useful in mild disease in compliant patients. *Helicobacter pylori* eradication in GORD has little effect on the symptoms but is usually advised (see p. 261).

Drugs

Alginate-containing antacids (10 mL three times daily) are the most frequently used agents for GORD. They form a gel or 'foam raft' with gastric contents to reduce reflux. Available over the counter, they are often used by the patient before consultation. Common proprietary antacids contain sodium which may exacerbate fluid retention; aluminium hydroxide has less sodium than magnesium trisilicate. Magnesium-containing antacids tend to cause diarrhoea whilst aluminium-containing compounds cause constipation.

The dopamine antagonist prokinetic agents metoclopramide and domperidone are occasionally helpful as they enhance peristalsis and speed gastric emptying.

H_2-receptor antagonists (e.g. cimetidine, ranitidine, famotidine and nizatidine) are frequently used for acid suppression if antacids fail as they can often be obtained over the counter.

Proton pump inhibitors (PPIs: omeprazole, rabeprazole, lansoprazole, pantoprazole, esomeprazole) inhibit gastric hydrogen/potassium-ATPase. PPIs reduce gastric acid secretion by up to 90% and are the drugs of choice for all but mild cases. Most patients with GORD will respond well but this is only 20–30% of patients presenting with heartburn. Patients with severe symptoms may need twice-daily PPIs and prolonged treatment, often for years. Once oesophageal sensitivity has normalized, a lower dose, e.g. omeprazole 10 mg, may be sufficient for maintenance. The patients who do not respond to a PPI are sometimes described as having non-erosive reflux disease (NERD) (Fig. 6.9), when the endoscopy is normal. These patients are usually female and often the symptoms are functional, although a small group have a hypersensitive oesophagus, giving discomfort with only slight changes in pH.

Surgery

Surgery should never be performed for a hiatus hernia alone. The best predictors of a good surgical result are typical reflux symptoms with documented acid reflux which correlates with symptoms and responds to a PPI. With such highly selected cases in experienced hands, the laparoscopic Nissen fundoplication has over a 90% satisfaction rate at 5 years, and recently available 10-year data show satisfaction rates remain high at 88%. Current surgical techniques return the oesophagogastric junction to the abdominal cavity, mobilize the gastric fundus, close the diaphragmatic crura snugly and involve a short tension-free fundoplication. Oesophageal stricture and erosive esophagitis can be treated successfully.

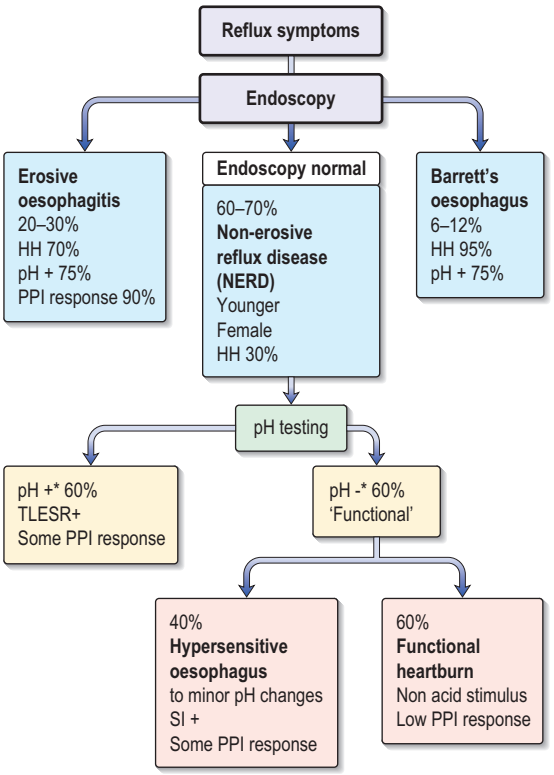


Fig. 6.9 **Outcome of patients with reflux symptoms.** *pH+, meets standard criteria for excessive acid reflux on intraluminal testing; pH–, does not meet standard criteria for excessive acid reflux; SI+, symptom index; TLESR, transient lower (o)esophageal sphincter relaxation; HH, hiatus hernia; PPI, proton pump inhibitor.

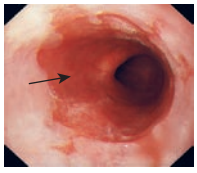


Fig. 6.10 **Barrett's oesophagus.** Endoscopic view showing a segment of columnar-lined oesophagus (arrow).

Indications for operation are not clear cut but include intolerance to medication, the desire for freedom from medication, the expense of therapy and the concern of long-term side-effects. The most common cause of mechanical fundoplication failure is recurrent hiatus hernia.

Patients with oesophageal dysmotility unrelated to acid reflux, patients who do not respond to PPIs and those with underlying functional bowel disease should not have surgery.

Complications

Peptic stricture

Peptic stricture usually occurs in patients over the age of 60. The symptoms are those of intermittent dysphagia for solids which worsen gradually over a long period. Mild cases may respond to PPI alone. More severe cases need endoscopic dilatation and long-term PPI therapy. Surgery is required if medical treatment fails.

Barrett's oesophagus (Fig. 6.10)

Barrett's oesophagus is a condition in which part of the normal esophageal squamous epithelium is replaced by metaplastic columnar mucosa to form a segment of 'columnar-lined oesophagus' (CLO). It is a complication of severe gastro-oesophageal reflux disease due to severe LOS hypotension and there is almost always a hiatus hernia. The diagnosis is made by endoscopy showing proximal displacement

of the squamocolumnar mucosal junction and biopsies demonstrating intestinal metaplasia. It may be seen as a continual sheet, or finger-like projections extending upwards from the squamocolumnar junction or as islands of columnar mucosa interspersed in areas of residual squamous mucosa. Recent guidelines suggest recording the length of circumferential CLO (C measurement) as well as the maximum length (M measurement) to aid assessment of progression or regression.

Central obesity increases the risk of Barrett's by 4.3 times. Long segment (>3 cm) and short segment (<3 cm) Barrett's is found respectively in 5% and 15% of patients, undergoing endoscopy for reflux symptoms. It is also often found incidentally in endoscoped patients without reflux symptoms. Barrett's is commonest in middle-aged men. The major concern is that 0.5–1.0% of Barrett's patients develop oesophageal adenocarcinoma per year. Barrett's increases the chance of developing adenocarcinoma 30- to 50-fold. Progression of Barrett's to cancer occurs through increasingly worse grades of dysplasia. The dysplasia is patchy, and biopsies from all four quadrants (every 2 cm) of the Barrett's segment are recommended, as well as biopsies from macroscopically abnormal areas. High-grade dysplasia (HGD) can be hard to differentiate from early adenocarcinoma and there is a high risk of progression. Indeed, 30–40% of patients with HGD have cancer in the resected specimen. Chromoendoscopy (the topical application of stains or pigments via the endoscope) magnification and narrow band imaging may aid the diagnosis of intestinal metaplasia, dysplasia and carcinoma.

Screening and surveillance in Barrett's oesophagus. There is little consensus about screening or surveillance, though many have strong opinions.

Screening does not identify those at risk for cancer (40% of patients with oesophageal adenocarcinoma (OA) have no history of reflux symptoms); survival for mainly elderly patients with Barrett's oesophagus is not significantly different from that for control subjects in the general population; no study has shown that surveillance is reliable in detecting curable dysplasia, as interval cancers occur during surveillance schemes; and treatment of the dysplasia has not been shown to prolong survival and/or improve quality of life, specifically by preventing oesophageal cancer. The estimated cost varies between $24 700 and $98 000 per life-year saved.

A consensus would be that patients without dysplasia do not require surveillance. Low-grade dysplasia requires extensive rebiopsy after 8–12 weeks of a PPI, rebiopsy 6 monthly if persistent and then 2–3 yearly if negative for dysplasia. For high-grade dysplasia – see treatment.

Treatment of Barrett's oesophagus. There is no evidence that treatment with PPIs or antireflux surgery leads to Barrett's regression, but symptomatic response does not correlate well with abolition of reflux. Ablation of high grade dysplastic Barrett's is under evaluation using photodynamic therapy, laser, endoscopic mucosal resection, argon plasma coagulation and wide-field diathermy, particularly in patients ineligible for surgery, but there is currently little long-term evidence of benefit. Because of the low rate of progression to cancer, ablation of non-dysplastic CLO seems unlikely to be of benefit.

FURTHER READING

Dallemagne B, Weerts J, Markiewicz S et al. Clinical results of laparoscopic fundoplication at ten years after surgery. *Surgical Endoscopy* 2006; **20**: 159–165.

Dulai GS, Guha S, Kahn KL et al. Preoperative prevalence of Barrett's esophagus in esophageal adenocarcinoma: a systematic review. *Gastroenterology* 2002; **122**: 26–33.

Goyal R, Shaker R (eds). Physiology of Oral Cavity, Pharynx and Upper Esophageal Sphincter. GI Motility On-line. www.nature.com/gimo/

Kahrilas PJ. Gastro oesophageal reflux disease. *New England Journal of Medicine* 2008; **359**: 1700–1707.

Moayyedi P, Talley NJ. Gastro-oesophageal reflux disease. *Lancet* 2006; **367**: 2086–2100.

Soni A, Sampliner RE, Sonnenberg A. Screening for high-grade dysplasia in gastroesophageal reflux disease: is it cost-effective? *American Journal of Gastroenterology* 2000; **95**: 2086–2093.

MOTILITY DISORDERS

Achalasia

Achalasia is characterized by oesophageal aperistalsis and impaired relaxation of the lower oesophageal sphincter.

Clinical features

Achalasia incidence is 1:100 000 equally in males and females. It occurs at all ages but is rare in childhood. Patients usually have a long history of intermittent dysphagia, characteristically for both liquids and solids from the onset. Regurgitation of food from the dilated oesophagus occurs, particularly at night, and aspiration pneumonia is a complication. Spontaneous chest pain occurs, said to be due to oesophageal 'spasm'. Dysphagia may be mild and accepted by the patient as normal. The pain may be misdiagnosed as cardiac. Weight loss is usually not marked.

Pathogenesis

The aetiology is unknown. Autoimmune, neurodegenerative and viral aetiologies have been implicated. A similar clinical picture is seen in chronic Chagas' disease (American trypanosomiasis, p. 158) where there is damage to the neural plexus of the gut.

Histopathology shows inflammation of the myenteric plexus of the oesophagus with reduction of ganglion cell numbers. Cholinergic innervation appears to be preserved. Reduction in nitric oxide synthase-containing neurones has been shown by immunohistochemical staining. Pharmacologic studies in patients with achalasia support the selective loss of *inhibitory*, nitrergic neurones. Genetically engineered animals with targeted disruption of the gene encoding for the neuronal form of nitric oxide synthase and pharmacologic administration of NO donors substantiate the role of nitrergic denervation in achalasia. Thus the relaxation of the sphincter is impaired in the absence of nitric oxide. Some patients have autoantibodies to a dopamine-carrying protein on the surface of the cells in the myenteric plexus. Degenerative lesions are also found in the vagus. The differential diagnosis of achalasia world-wide includes genetic syndromes, infectious diseases, neoplasms and chronic inflammatory conditions.

Investigations

- **Chest X-ray** shows a dilated oesophagus, sometimes with a fluid level seen behind the heart. The fundal gas shadow is absent.

- **Barium swallow** shows lack of peristalsis and often synchronous contractions in the body of the oesophagus, sometimes with dilatation. The lower end shows a 'swan neck deformity' due to failure of the sphincter to relax (Fig. 6.11).
- **Oesophagoscopy** is performed to exclude a carcinoma at the lower end of the oesophagus which can produce a similar X-ray appearance. When there is marked dilatation, a 24-hour liquid-only diet and a washout prior to endoscopy is useful to remove food debris. In true achalasia the endoscope passes through the lower oesophageal sphincter with little resistance.
- **CT scan** excludes distal oesophageal cancer.
- **Manometry** shows aperistalsis of the oesophagus and failure of relaxation of the lower oesophageal sphincter (Fig. 6.5).

Treatment

All current forms of treatment for achalasia are palliative. Drug therapy rarely produces satisfactory or durable relief; nifedipine (20 mg sublingually) or sildenafil can be tried initially.

The treatment of choice is endoscopic dilatation of the LOS using a hydrostatic balloon under X-ray control. This weakens the sphincter and is successful initially in 80% of cases. About 50% of patients require a second or third dilatation in the first 5 years. There is a low but significant risk of perforation. Intrasphincteric injection of botulinum toxin A produces satisfactory initial results but the effects wear off within months. Further injections can be given. It is safer and simpler than dilatation, so may be valuable in patients at risk of death if a perforation occurs. Neither pneumatic dilatation nor botulinum toxin work as well in younger patients.

Fig. 6.11 **Barium swallow showing achalasia** with atonic body of the oesophagus and a narrowed distal end. Note food residue in dilated oesophagus.

If these measures fail, surgical division of the LOS (cardiomyotomy, or Heller's operation) is performed.

Reflux oesophagitis complicates all procedures and the aperistalsis of the oesophagus remains.

Complications

There is a slight increase in the incidence of squamous carcinoma of the oesophagus in both treated and untreated cases (7% after 25 years).

Systemic sclerosis (see also p. 545)

The oesophagus is involved in almost all patients with this disease. Diminished peristalsis and oesophageal clearance, detected manometrically (Fig. 6.5) or by barium swallow, is due to replacement of the smooth muscle by fibrous tissue. LOS pressure is decreased, allowing reflux with consequent mucosal damage. Strictures may develop. Initially there are no symptoms, but dysphagia and heartburn occur as the oesophagus becomes more severely involved. Similar motility abnormalities may be found in other connective tissue disorders, particularly if Raynaud's phenomenon is present. Treatment is as for reflux (see p. 252) and benign stricture.

Diffuse oesophageal spasm

This is a severe form of oesophageal dysmotility that can sometimes produce retrosternal chest pain and dysphagia. It can accompany GORD. Swallowing is accompanied by bizarre and marked contractions of the oesophagus without normal peristalsis (Fig. 6.5). On barium swallow the appearance may be that of a 'corkscrew' oesophagus. However, asymptomatic oesophageal 'dysmotility' is not infrequent, particularly in patients over the age of 60 years.

A variant of diffuse oesophageal spasm is the 'nutcracker' oesophagus, which is characterized by very high-amplitude peristalsis (pressures > 200 mmHg) within the oesophagus. Chest pain is commoner than dysphagia.

Treatment

True oesophageal spasm producing severe symptoms is uncommon and treatment is often difficult. PPIs may be successful if reflux is a factor. Antispasmodics, nitrates or calcium-channel blockers – such as sublingual nifedipine 10 mg three times daily – may be tried. Occasionally, balloon dilatation or even longitudinal oesophageal myotomy is necessary.

Miscellaneous motility disorders

Abnormalities of motility that occasionally produce dysphagia are found in the elderly, in diabetes mellitus, myotonic dystrophy, oculopharyngeal muscular dystrophy and myasthenia gravis, as well as neurological disorders involving the brainstem.

OTHER OESOPHAGEAL DISORDERS

Oesophageal diverticulum

Diverticula occur:

- immediately **above** the upper oesophageal sphincter (pharyngeal pouch – Zenker's diverticulum) (see p. 1084).

- near the **middle** of the oesophagus (traction diverticulum due to inflammation, or associated with diffuse oesophageal spasm or mediastinal fibrosis)
- just above the **lower** oesophageal sphincter (epiphrenic diverticulum – associated with achalasia).

Usually detected incidentally on a barium swallow performed for other reasons, these are often asymptomatic. Dysphagia and regurgitation can occur with a pharyngeal pouch (see p. 1084).

Rings and webs

An oesophageal web is a thin, membranous tissue flap covered with squamous epithelium. Most acquired webs are located anteriorly in the postcricoid region of the cervical oesophagus and are well seen on barium swallow. They may produce dysphagia. In the *Plummer–Vinson syndrome* (or *Paterson–Brown–Kelly syndrome*) a web is associated with chronic iron deficiency anaemia, glossitis and angular stomatitis. This rare syndrome affects mainly women and its aetiology is not understood. The web may be difficult to see at endoscopy and may be ruptured unintentionally by the passage of the endoscope. Dilatation of the web is rarely necessary. Iron is given for the iron deficiency.

Cricopharyngeal dysfunction

There is poor relaxation of the cricopharyngeal muscle during swallowing so that a prominent indentation or 'bar' is seen on a barium swallow. It occurs in the elderly. Dysphagia is treated with dilatation but surgical myotomy may be necessary.

Lower oesophageal rings

Lower oesophageal rings are of two types:

1 *Mucosal (Schatzki's ring, also called B ring)* (Fig. 6.12), which is located at the squamocolumnar mucosal junction; it is common, and is associated with characteristic history of intermittent bolus obstruction. Barium swallow with a distended oesophagus shows the abnormality which may be difficult to see at endoscopy.
2 *Muscular (A ring)*, which is located proximal to the mucosal ring; it is uncommon, covered by squamous epithelium and may cause dysphagia.

Treatment for these rings is usually with reassurance and dietary advice, but dilatation is occasionally necessary. After a single dilatation, 68% of patients with Schatzki's rings are symptom free at 1 year, 35% remain symptom free after 2 years, but only 11% are symptom free at 3 years. Many also respond to oral PPI, either alone or with dilatation.

Benign oesophageal stricture (Fig. 6.13)

Peptic stricture secondary to reflux is the most common cause of benign strictures (for treatment, see p. 252). They also occur after the ingestion of corrosives, after radiotherapy, after sclerosis of varices, and following prolonged nasogastric intubation. Dysphagia is usually treated by endoscopic dilatation. Surgery is sometimes required.

Oesophageal infections

Infection is a cause of painful swallowing and is seen particularly in immunosuppressed (e.g. on chemotherapy) and debilitated patients and patients with AIDS. Infection can occur with:

- *Candida*
- herpes simplex
- cytomegalovirus.

It is occasionally difficult to distinguish between these either on barium swallow or oesophagoscopy, as only widespread ulceration is seen. In candidiasis the characteristic white plaques are frequently found; oral candidiasis is not always present. The diagnosis of *Candida* infection can be confirmed by examining a direct smear taken at endoscopy, but often infections are mixed and cultures and biopsies must be performed.

Treatment

Most patients on large doses of immunosuppressive agents are treated prophylactically with nystatin or amphotericin. Other antifungal or antiviral treatment is given appropriately (see Ch. 4).

Mallory–Weiss syndrome

This is described on page 269.

Eosinophilic oesophagitis

Eosinophils can be seen in the oesophageal mucosa (which is usually devoid of eosinophils microscopically) due to a variety of causes such as eosinophilic (or allergic) oesophagitis and GORD.

Eosinophilic oesophagitis is being increasingly recognized but its pathogenesis is unknown. There may be a personal or family history of allergic disorders. Food allergy is thought to be a pathogenetic factor.

Patients present with a long history of dysphagia, food impaction, 'heartburn' and oesophageal pain. Usually the patient is male, white, and with an average age at diagnosis of 35, but it also occurs in children.

Typical endoscopic abnormalities include mucosal furrowing, loss of vascular pattern due to a thickened mucosa, plaques of eosinophilic surface exudate and prominent circular folds (Fig. 6.14), but the oesophagus may appear macroscopically normal. Reflux oesophagitis and Schatzki's rings may co-exist. Endoscopic forceps biopsies (×6) should be taken from upper, mid and lower oesophagus for histology with eosinophil counts.

The eosinophilic infiltration of the oesophagus due to *reflux disease* can be excluded by rebiopsy after a 6-week course of full dose PPI.

Treatment is with topical (swallowing inhaled steroid preparations, e.g. fluticasone, with inhalers without spacers) or systemic steroids. Children may respond dramatically to elimination diets but relapse when a normal diet is resumed. The response in adults is less certain. Dilatation is sometimes necessary but a high risk of perforation is reported.

Oesophageal perforation or rupture

Oesophageal perforation most commonly occurs at the time of endoscopic dilatation and, rarely, following insertion of a

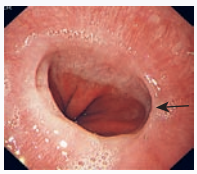


Fig. 6.12 **Schatzki's ring** – endoscopic view.

Fig. 6.13 **Benign oesophageal reflux stricture** – endoscopic view.

Fig. 6.14 **Eosinophilic oesophagitis** – endoscopic view.

Ask the authors
www.studentconsult.com

nasogastric tube. Patients with malignant, corrosive or post-radiotherapy strictures are more likely to suffer perforation than those with a benign peptic stricture.

Management usually involves placement of an oesophageal stent (see p. 257), which dilates the (usually malignant) stricture and seals the hole. A water-soluble contrast X-ray is performed after 2–3 days to check the perforation has sealed.

'Spontaneous' oesophageal rupture occurs with violent vomiting (Boerhaave's syndrome), producing severe chest pain and collapse in typical cases. Diagnosis can be difficult because classic symptoms are absent in about a third of cases and delays in presentation for medical care are common. It may follow alcohol ingestion. A chest X-ray shows a hydropneumothorax. The diagnosis is made with a water-soluble contrast swallow or on CT. The mortality rate is approximately 35%, making it the most lethal perforation of the GI tract. The best outcomes are associated with early diagnosis and definitive surgical management within 12 hours of rupture. If intervention is delayed longer than 24 hours, the mortality rate (even after surgery) rises to above 50% and to nearly 90% after 48 hours.

Table 6.5 Risk factors for cancer of the oesophagus

Squamous cell carcinoma	Adenocarcinoma
Tobacco smoking	Longstanding, severe GORD
High alcohol intake	Barrett's oesophagus
Plummer–Vinson syndrome	Tobacco smoking
Achalasia	Obesity
Corrosive strictures	Breast cancer treated with radiotherapy
Coeliac disease	
Breast cancer treated with radiotherapy	
Tylosis*	
Diet deficient in vitamins; high dietary carotenoids and vitamin C probably decreases the risk	

*Tylosis is a rare autosomal dominant condition with hyperkeratosis of the palms and soles.

OESOPHAGEAL TUMOURS

Cancer of the oesophagus

This is the sixth most common cancer world-wide. Squamous tumours occurring in the middle third account for 40% of tumours, and in the upper third, 15%. Adenocarcinomas occur in the lower third of the oesophagus and at the cardia and represent approximately 45%. Primary small cell cancer is extremely rare.

Epidemiology and aetiological factors

Squamous cell carcinoma (SCC)

The geographic variation in incidence is greater than for any other carcinoma – often in regions very close to one another. It is common in Ethiopia, China, South and east Africa and in the Caspian regions of Iran. By contrast, north, middle and west Africa have low rates (www.cancerresearchuk.org).

In the UK the incidence is 5–10 per 100 000 and represents 2.2% of all malignant disease. The incidence of SCC is decreasing, in contrast to adenocarcinoma. SCC of the oesophagus is more common in men (2:1). Risk factors are shown in Table 6.5.

Diets rich in fibre, carotenoids, folate, vitamin C and non-starchy vegetables probably *decrease* the risk of oesophageal cancer whereas diets high in saturated fat and cholesterol and refined cereals have been associated with an increased risk. Red and processed meat intake have been associated with an increased risk of both oesophageal SCC and adenocarcinoma. Conversely, fish and white meat consumption have been inversely associated with risk of oesophageal SCC in case-control studies from Italy, Switzerland and Uruguay.

Adenocarcinoma

These tumours primarily arise in columnar-lined epithelium in the lower oesophagus (see also Barrett's oesophagus, p. 252). The incidence of this tumour is increasing in western industrialized countries – by three- to fivefold over the last 40 years in Europe and the USA. Incidence rates in the year 2000 were: Great Britain (5.0–8.7/100 000), Australia (4.8/100 000) and the Netherlands (4.4/100 000), followed by the US (3.7/100 000). Extension of an adenocarcinoma of the gastric cardia into the oesophagus can present with the same symptoms. Previous reflux symptoms increase the risk up to eightfold and the risk is proportional to their severity.

Clinical features

Carcinoma of the oesophagus occurs mainly in those aged 60–70 years. Dysphagia is progressive and unrelenting. Initially there is difficulty in swallowing solids, but typically dysphagia for liquids follows within weeks. Impaction of food causes pain, but more persistent pain implies infiltration of adjacent structures.

The lesion may be ulcerative, proliferative or scirrhous, extending variably around the wall of the oesophagus to produce a stricture. Direct invasion of the surrounding structures and metastases to lymph nodes are commoner than disseminated metastases. Weight loss, due to the dysphagia as well as to anorexia, is frequent. Oesophageal obstruction eventually causes difficulty in swallowing saliva, and coughing and aspiration into the lungs is common.

Weight loss, anorexia and lymphadenopathy are the commonest physical signs.

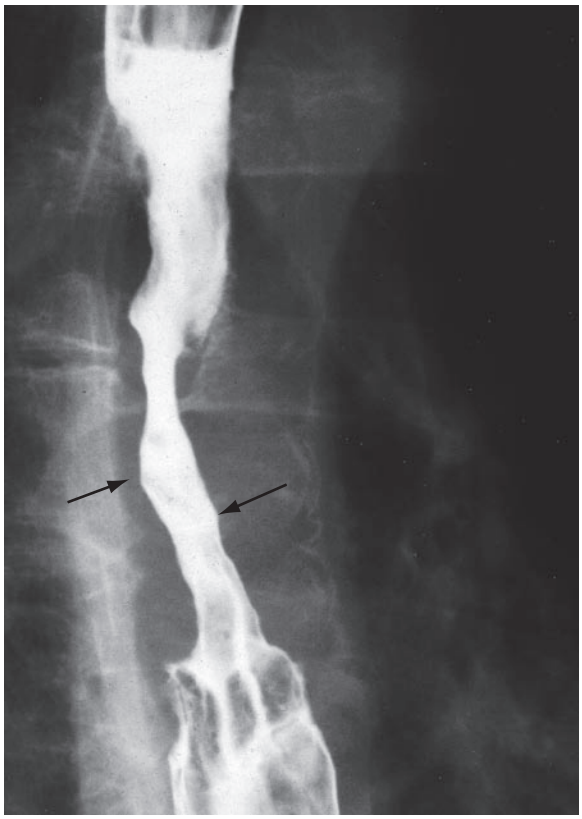
Investigation

Diagnostic

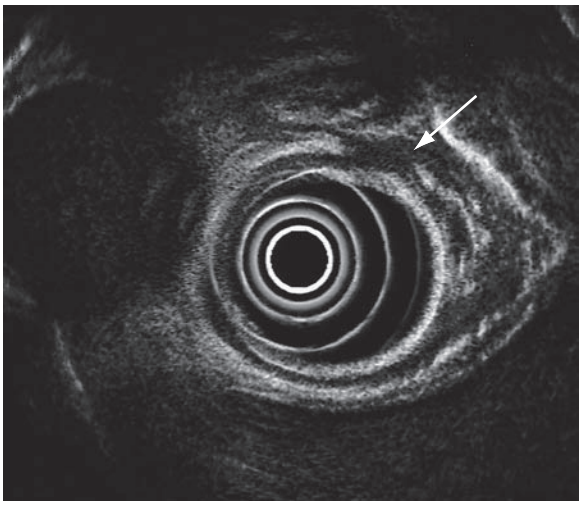
- *Endoscopy* provides histological or cytological proof of the carcinoma; 90% of oesophageal carcinomas can be confirmed with this technique.
- *Barium swallow* can be useful where the differential diagnosis of dysphagia includes a motility disorder such as achalasia (Fig. 6.15a).

Staging

The TNM staging system is used (see p. 456). Tumour invasion of the wall of the oesophagus (T), presence of tumour in lymph nodes (N) or metastases (M) are combined into stage categories.



(a)



(b)

Fig. 6.15 **Carcinoma of the oesophagus. (a) Barium swallow**, showing an irregular narrowed area (arrows) at the lower end of the oesophagus. **(b) Endoscopic ultrasound.** The central concentric circles are the probe. The arrow points to a break in the muscle layer and the soft tissue mass of the carcinoma.

- *CT scan* of the thorax and upper abdomen shows the volume of the tumour, local invasion, peritumoral and coeliac lymph node involvement and metastases in the lung and elsewhere. *MRI* is equivalent to CT in local staging but not as good for pulmonary metastases.
- *Endoscopic ultrasound* has an accuracy rate of nearly 90% for assessing depth of tumour and infiltration and 80% for staging lymph node involvement. It is useful if CT does not show a lesion already too advanced for surgery. A fine needle aspiration (FNA) of lymph nodes improves staging accuracy.
- *Laparoscopy* is useful if the tumour is at the cardia, to look for peritoneal and node metastases.
- *Positron emission tomography (PET)* after fluorodeoxyglucose is used principally to confirm distant metastases suspected on CT.

Treatment

This depends on the age and performance status of the patient and the stage of the disease. Five-year survival with stage 1 is 80% (T_1/T_2, N_0, M_0), stage 2 is 30%, stage 3 is 18% and stage 4 is 4%. Seventy per cent of patients present with stage 3 or greater disease, so that overall survival is 27% at 1 year and around 10% at 5 years. Management should be undertaken by multidisciplinary teams.

- *Surgery* provides the best chance of a cure and should be used only when imaging (see above) has shown that the tumour has not infiltrated outside the oesophageal wall (Stage 1). Surgery in this group shows an 80% survival rate if the postoperative pathology confirms the staging. The patients must be carefully evaluated preoperatively, particularly with regard to performance status (see p. 456) and surgery undertaken in designated units.
- *Chemoradiation.* Pre-operative ('neo-adjuvant') chemoradiation therapy may benefit patients with stage 2b and 3 disease. Prolongation of survival has been shown in some studies.
- *Palliative therapy* is often the only realistic possibility. Dilatation is only of short-term benefit. Insertion of an expanding metal stent allows liquids and soft foods to be eaten. Brachytherapy (see p. 465) prolongs luminal patency after stenting. Endoscopic laser destruction, or alcohol injections can be useful in proliferative tumours for relieving the dysphagia but repeated treatments are necessary.
- *Photodynamic therapy* (PDT) can be useful. Chemoradiation is sometimes given, but evidence of benefit is poor.
- *Nutritional support*, as well as support for the patient and their family, is vital in this distressing condition.

Other oesophageal tumours

Most other tumours are rare. Gastrointestinal stromal tumours (see p. 266) are found usually by chance. Ten per cent cause dysphagia or bleeding. Surgical removal is performed for symptomatic lesions or those over 3 cm which are more likely to harbour malignancy. Small benign tumours are relatively common and often do not require treatment.

Kaposi's sarcoma is found in the oesophagus as well as the mouth (see p. 204) and hypopharynx in patients with AIDS.

FURTHER READING

Oesophageal reflux/motility
Lagergren J, Bergström R, Lindgren A et al. Symptomatic gastro-oesophageal reflux as a risk factor for oesophageal adenocarcinoma. *New England Journal of Medicine* 1999; **340**: 825–831.

Oesophageal cancer
Allum WH. Guidelines for the management of oesophageal and gastric cancer. *Gut* 2002; **50** (suppl V): 1–17.

Fiorica F, Di Bona D, Schepis F et al. Preoperative chemoradiotherapy for oesophageal cancer – meta-analysis. *Gut* 2004; **53**: 925–930.

Wang KK, Wongkeesong M, Buttar NS. American Gastroenterological Association Technical Review on the role of the gastroenterologist in the management of esophageal carcinoma. *Gastroenterology* 2005; **128**: 1471–1505.

World Cancer Research Fund/American Institute for Cancer Research. *Food, Nutrition and Physical Activity in the Prevention of Cancer: A Global Perspective*. Washington DC: AICR, 2007.

THE STOMACH AND DUODENUM

Structure

The stomach consists of a small area immediately distal to the oesophagus (the cardia), the upper region (the fundus, under the left diaphragm), the mid-region or body, and the antrum, which extends to the pylorus (Fig. 6.6).

The smooth muscle of the wall of the stomach has three layers – outer longitudinal, inner circular and innermost oblique layers. There are two sphincters, the gastro-oesophageal sphincter and the pyloric sphincter. The latter is largely made up of a thickening of the circular muscle layer and controls the exit of gastric contents into the duodenum.

The duodenum has outer longitudinal and inner smooth muscle layers. It is C-shaped and the pancreas sits in the concavity. It terminates at the duodenojejunal flexure where it joins the jejunum.

- The *mucosal lining* of the stomach can stretch in size with feeding. The greater curvature of the undistended stomach has thick folds or rugae. The mucosa of the upper two-thirds of the stomach contains *parietal cells* that secrete hydrochloric acid, and *chief cells* that secrete pepsinogen (which initiates proteolysis). There is often a colour change at the junction between the body and the antrum of the stomach that can be seen macroscopically, and confirmed by measuring surface pH.
- The *antral mucosa* secretes bicarbonate and contains mucus-secreting cells and *G cells*, which secrete gastrin. There are two major forms of gastrin, G17 and G34, depending on the number of amino-acid residues. G17 is the major form found in the antrum. Somatostatin is also produced by specialized antral cells (*D cells*).
- *Mucus-secreting cells* are present throughout the stomach and secrete mucus and bicarbonate. The mucus is made of glycoproteins called mucins.
- The *'mucosal barrier'*, made up of the plasma membranes of mucosal cells and the mucus layer, protects the gastric epithelium from damage by acid and, for example, alcohol, aspirin, NSAIDs and bile salts. Prostaglandins stimulate secretion of mucus, and their synthesis is inhibited by aspirin and NSAIDs, which inhibit cyclo-oxygenase (see Fig. 14.32).
- The *duodenal mucosa* has villi like the rest of the small bowel, and also contains Brunner's glands that secrete alkaline mucus. This, along with the pancreatic and biliary secretions, helps to neutralize the acid secretion from the stomach when it reaches the duodenum.

Function

Factors controlling *acid secretion* are shown in Figure 6.16. Acid is not essential for digestion but does prevent some food-borne infections. Acid secretion is under neural and hormonal control. Both stimulate acid secretion through the direct action of histamine on the parietal cell. Acetylcholine and gastrin also release histamine via the enterochromaffin cells. Somatostatin inhibits both histamine and gastrin release and therefore acid secretion.

Other major gastric functions are:

- reservoir for food
- emulsification of fat and mixing of gastric contents
- secretion of intrinsic factor
- absorption (of only minimal importance).

Gastric emptying depends on many factors. There are osmoreceptors in the duodenal mucosa that control gastric emptying by local reflexes and the release of gut hormones. In particular, intraduodenal fat delays gastric emptying by negative feedback through duodenal receptors.

HELICOBACTER PYLORI INFECTION

H. pylori is a slow-growing spiral Gram-negative flagellate urease-producing bacterium (Fig. 6.17) which plays a major role in gastritis and peptic ulcer disease. Its complete genomic sequence is known. It colonizes the mucous layer in the gastric antrum but is found in the duodenum only in areas of gastric metaplasia. *H. pylori* is found in greatest numbers under the mucous layer in gastric pits, where it adheres specifically to gastric epithelial cells. It is protected from gastric acid by the juxtamucosal mucous layer which traps bicarbonate secreted by antral cells, and ammonia produced by bacterial urease.

Epidemiology of *Helicobacter* infection

The prevalence of *H. pylori* is high in developing countries (80–90% of the population), and much lower (20–50%) in developed countries. Infection rates are highest in lower income groups. Infection is usually acquired in childhood; although the exact route is uncertain, it may be faecal–oral or oral–oral, either by kissing or ingestion of contaminated vomit. Once acquired, the infection persists for life unless treated. The incidence increases with age, probably due to acquisition in childhood when hygiene was poorer (cohort effect) and not due to infection in adult life which is probably far less than 1% per year in developed countries.

Pathogenesis of *Helicobacter* infection

The pathogenetic mechanisms are not fully understood. *H. pylori* adheres by BabA adhesin to Lewis antigen expressed on the surface of gastric mucosal cells and causes gastritis in all infected subjects. Ulcers are commonest when the infecting strain expresses CagA (cytotoxic-associated protein) and VacA (vacuolating toxin) genes. CagA is a marker for a section of the bacterial DNA that contains genes responsible for a secretion system (Cag pathogenicity island, Pa_I) which injects CagA product into epithelial cells via a

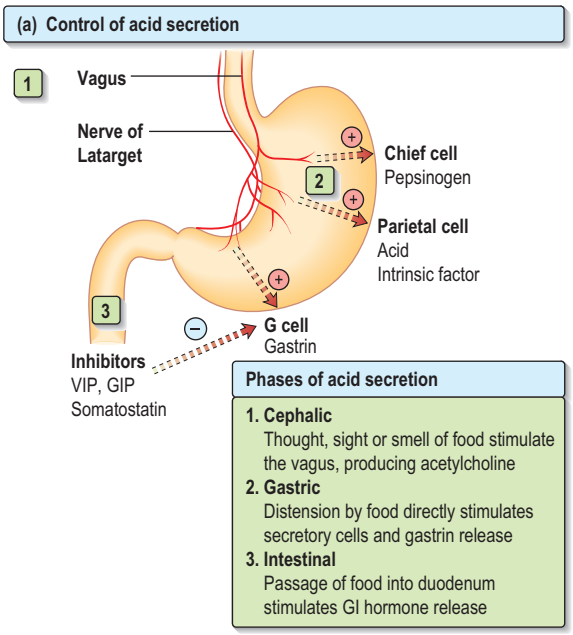


Fig. 6.16 **(a) Control of acid secretion. (b) Mechanisms involved in acid secretion.**

- Histamine stimulates G_s via the H_2 receptors and acts via cyclic AMP.
- Prostaglandin E_2 activates the G_i protein and INHIBITS acid secretion.
- Acetylcholine (ACh) acts via the vagus M_3 receptors.
- ACh also acts via the ECL cell.
- Gastrin acts via the gastrin receptor, increasing the intracellular free calcium, and also via the ECL cell stimulating histamine.

G_s and G_i, stimulating and inhibiting G-proteins; ECL, enterochromaffin cells.

pilus. This causes alterations in cell morphology, replication and apoptosis. VacA is a pore-forming protein which increases host cell permeability, induces apoptosis and suppresses local immune mechanisms. Genetic variations in the host are also thought to be involved; for example, polymorphisms leading to increased levels of IL-1 are associated with atrophic gastritis and cancer.

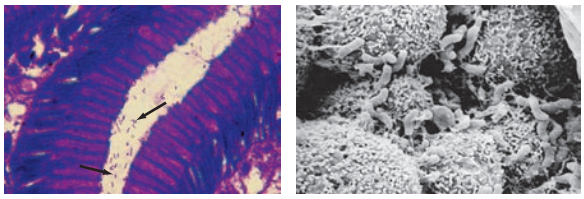


Fig. 6.17 ***Helicobacter pylori*. (a) Organisms (arrowed) are shown on the gastric mucosa** (cresyl fast violet (modified Giemsa) stain). Courtesy of Dr Alan Phillips, Department of Paediatric Gastroenterology, Royal Free Hospital. **(b) Scanning electron microscopy,** showing the spiral-shaped bacterium.

Results of infection

- Antral gastritis
- peptic ulcers (duodenal and gastric)
- gastric cancer.

Antral gastritis is the usual effect of *H. pylori* infection. It is usually asymptomatic, although occasionally patients without ulcers claim relief of dyspeptic symptoms after *Helicobacter* eradication. Antral gastritis causes hypergastrinaemia due to gastrin release from antral G cells. The subsequent increase in acid output is usually asymptomatic.

Duodenal ulcer (DU) (see Fig. 6.19a). *H. pylori* is causally associated with DU disease because in patients with DU:

- 95% are infected with *H. pylori* in the antrum (antral gastritis).
- Cure of the infection heals the ulcer and stops duodenal ulcer recurrence.

The precise mechanism of duodenal ulceration is unclear, as only 15% of patients infected with *H. pylori* (50–60% of the adult population world-wide) develop duodenal ulcers. Factors that have been implicated include:

- Increased acid secretion because of:
  - increased parietal cell mass
  - increased gastrin secretion.
- Smoking impairing mucosal healing.
- Virulence factors such as Vac A (vacuolating toxin) and CagA (cytotoxic-associated protein) as well as urease and adherence factors (see above).
- Decreased inhibition of acid secretion; *H. pylori*-induced gastritis reduces somatostatin production in the antrum with loss of the negative feedback on gastrin secretion.
- Genetic susceptibility: duodenal ulcers are more common in patients who have blood group O and are non-secretors of blood group substances in saliva.
- Duodenal bicarbonate secretion is decreased by *H. pylori* inflammation and the damage and repair leads to gastric metaplasia which *H. pylori* colonizes, causing local release of cytokines and further damage.

Gastric ulcer (GU) (see Fig. 6.19b). Gastric ulcers are associated with a gastritis affecting the body as well as the antrum of the stomach (pangastritis) causing parietal cell loss and reduced acid production. The ulcers are thought to occur because of reduction of gastric mucosal resistance due to cytokine production by the infection or perhaps to alterations in gastric mucus.

Epidemiology of peptic ulcer disease

Duodenal ulcers affect 10–15% of the adult population and are two to three times more common than gastric ulcers.

Ulcer rates are declining rapidly for younger men and increasing for older individuals, particularly women. Both DUs and GUs are common in the elderly. There is considerable geographical variation, with peptic ulcer disease being more prevalent in developing countries related to the high *H. pylori* infection.

Pathology of peptic ulcer disease

A *peptic ulcer* consists of a break in the superficial epithelial cells penetrating down to the muscularis mucosa; there is a fibrous base and an increase in inflammatory cells. *Erosions*, by contrast, are superficial breaks in the mucosa alone. Most DUs are found in the duodenal cap; the surrounding mucosa appears inflamed, haemorrhagic or friable (duodenitis). GUs are most commonly seen on the lesser curve near the incisura, but can be found in any part of the stomach. Peptic ulcers are seen without *H. pylori*, e.g. in patients on NSAIDs and in Zollinger–Ellison syndrome (see p. 385).

Clinical features of peptic ulcer disease

The characteristic feature of peptic ulcer is burning epigastric pain. It has been shown that if a patient points with a single finger to the epigastrium as the site of the pain this is strongly suggestive of peptic ulcer disease. The relationship of the pain to food is variable and on the whole not helpful in diagnosis. The pain of a DU classically occurs at night (as well as during the day) and is worse when the patient is hungry, but this is not reliable. The pain of both gastric and duodenal ulcers may be relieved by antacids.

Nausea may accompany the pain; vomiting is infrequent but often relieves pain. Anorexia and weight loss may occur, particularly with GUs. Persistent and severe pain suggests complications such as penetration into other organs. Back pain suggests a penetrating posterior ulcer. Severe ulceration can occasionally be symptomless, as many who present with acute ulcer bleeding or perforation have no preceding ulcer symptoms.

Untreated, the symptoms of a DU relapse and remit spontaneously. The natural history is for the disease to remit over many years due to the onset of atrophic gastritis and a decrease in acid secretion.

Examination is usually unhelpful; epigastric tenderness is quite common in non-ulcer dyspepsia.

Diagnosis of *Helicobacter pylori* infection

Non-invasive methods

- *Serological tests* detect IgG antibodies and are reasonably sensitive (90%) and specific (83%). They have been used in diagnosis and in epidemiological studies. IgG titres may take up to 1 year to fall by 50% after eradication therapy and therefore are not useful for confirming eradication or the presence of a current infection. Antibodies can also be found in the saliva, but tests are not as sensitive or specific as serology.
- *^{13}C-Urea breath test* (Fig. 6.18). This is a quick and reliable test for *H. pylori* and can be used as a screening test. The measurement of $^{13}CO_2$ in the breath after ingestion of ^{13}C urea requires a mass spectrometer. The test is very sensitive (97%) and specific (96%). This test is suitable for testing for eradication of the organism, but may be falsely negative if patients are taking PPIs at the time. A rapid release tablet which produces a result in 15 mins is becoming available.
- *Stool antigen test.* A specific immunoassay using monoclonal antibodies for the qualitative detection of *H. pylori* antigen is widely available. The overall sensitivity is 97.6% with a specificity of 96%. It is useful in the diagnosis of *H. pylori* infection and for monitoring efficacy of eradication therapy. (Patients should be off PPIs for 1 week but can continue with H_2 blockers.)

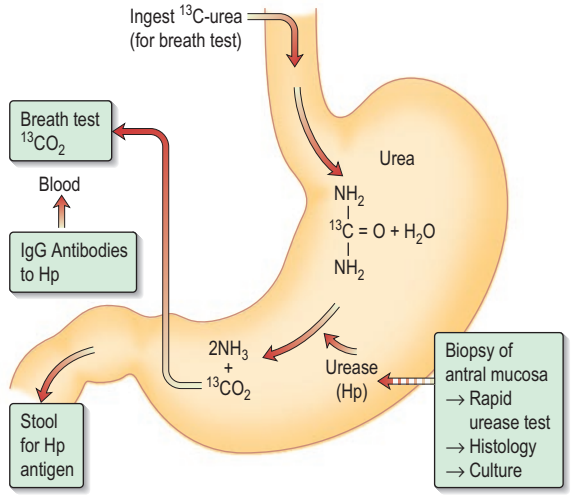


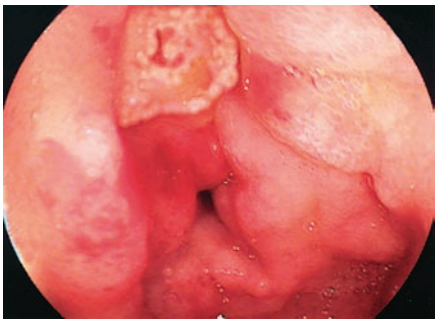
Fig. 6.18 **Metabolism of urea by *Helicobacter pylori* (Hp), showing the different tests that are available for the detection of *H. pylori*.**

Invasive (endoscopy)

- *Biopsy urease test.* Gastric biopsies are added to a substrate containing urea and phenol red. If *H. pylori* are present, the urease enzyme that they produce splits the urea to release ammonia which raises the pH of the solution and causes a rapid colour change (yellow to red). The test may be falsely negative if patients are taking PPIs or antibiotics at the time.
- *Culture.* Biopsies obtained can be cultured on a special medium, and *in vitro* sensitivities to antibiotics can be tested.
- *Histology. H. pylori* can be detected histologically on routine (Giemsa) stained sections of gastric mucosa obtained at endoscopy.

Investigation of suspected peptic ulcer disease

- *Patients under 55 years of age* with typical symptoms of peptic ulcer disease who are *H. pylori* positive can start eradication therapy without investigation.



(a)

(b)

Fig. 6.19 **Endoscopic views. (a)** Duodenal ulcer with inflamed duodenal folds. **(b) Benign gastric ulcer**.

- Confirmation of the diagnosis and exclusion of cancer is required in older patients.
- Endoscopy is the preferred investigation (Fig. 6.19). All GUs must be biopsied.
- Endoscopy is required in all patients with 'alarm symptoms' (see p. 241).

Eradication therapy

Current recommendations are that all patients with duodenal and gastric ulcers should have *H. pylori* eradication therapy. Many patients have incidental *H. pylori* infection with no gastric or duodenal ulcer. Whether all such patients should have eradication therapy is controversial (see Functional dyspepsia, p. 311). However, eradication therapy is advised in the hope that symptoms will be reduced and because of the link between *H. pylori* and gastric cancer.

The increase in the prevalence of GORD and adenocarcinoma of the lower oesophagus in the last few years is currently unexplained, but has been postulated to be linked to eradication of *H. pylori*. This seems unlikely but is not disproven.

Standard eradication therapies are successful in approximately 90% of patients.

Reinfection is very uncommon (1%) in developed countries. In developing countries reinfection is more common, compliance with treatment may be poor and metronidazole resistance is high (>50%) (as it is frequently used for parasitic infections) so failure of eradication is common.

There are many regimens for eradication, but all must take into account that:

- good compliance is essential
- there is a high incidence of resistance to metronidazole, particularly in some populations
- oral metronidazole has frequent side-effects
- bismuth chelate is unpleasant to take, even as tablets.

Metronidazole, clarithromycin, amoxicillin, tetracycline and bismuth are the most widely used agents. Resistance to amoxicillin (1–2%) and tetracycline (<1%) is low except in countries where they are available without prescription where resistance may exceed 50%. Quinolones such as ciprofloxacin, furazolidone and rifabutin are also used when standard regimens have failed ('rescue therapy'). Bismuth suppresses *H. pylori* effectively. None of these drugs is effective alone; eradication regimens therefore usually comprise two antibiotics given with powerful acid suppression in the form of a PPI, all given for 7 days.

Example regimes are:

- omeprazole 20 mg + clarithromycin 500 mg and amoxicillin 1 g – all twice daily
- omeprazole 20 mg + metronidazole 400 mg and clarithromycin 500 mg – all twice daily.

Box 6.3 Suspected perforated peptic ulcer

Look for:

- Other acute gastrointestinal conditions (e.g. cholecystitis, pancreatitis – check serum amylase)
- Non-GI conditions (e.g. myocardial infarction)
- Silent perforation in the elderly or patients on steroids
- Remember:
  - there is harm in leaving an undiagnosed perforation
  - avoid laparotomy if pancreatitis is diagnosed

In some regimens, H_2-receptor antagonists, e.g. ranitidine or ranitidine bismuth citrate, are included instead of a PPI. Bismuth chelate is not usually given in initial regimens because of the more complex dosing regimen and side effects. In eradication failures bismuth chelate (120 mg 4× daily), metronidazole (400 mg 3× daily), tetracycline (500 mg 4× daily) and a PPI 20–40 mg 2× daily) for 14 days is used. Sequential courses of therapy are being used in areas where resistance is high.

Prolonged therapy with a PPI after a course of PPI-based 7-day triple therapy is not necessary for ulcer healing in most *H. pylori*-infected patients. The effectiveness of treatment for uncomplicated duodenal ulcer should be assessed symptomatically. If symptoms persist, breath or stool testing should be performed to check eradication.

Patients with a risk of bleeding or those with complications, i.e. haemorrhage or perforation, should always have a ^{13}C urea breath test or stool test for *H. pylori* 6 weeks after the end of treatment to be sure eradication is successful. Long-term PPIs may be necessary if a rebleed would be likely to be fatal.

General measures

Stopping smoking should be strongly encouraged as smoking slows mucosal healing.

Patients with gastric ulcers should be routinely re-endoscoped at 6 weeks to exclude a malignant tumour by confirming healing with biopsy if necessary.

Complications of peptic ulcer

Haemorrhage

See page 267.

Perforation (Box 6.3; see also p. 316).

The frequency of perforation of peptic ulceration is decreas-

FURTHER READING

Gerrits MM, van Vliet AHM, Kuipers EJ et al. Helicobacter and antimicrobial resistance: molecular mechanisms and clinical implications. *Lancet Infectious Diseases* 2006; **6**: 699–709.

ing, partly attributable to medical therapy. DUs perforate more commonly than GUs, usually into the peritoneal cavity; perforation into the lesser sac also occurs. Detailed management of perforation is described on page 316. Surgery is usually performed to close the perforation and drain the abdomen. Conservative management using nasogastric suction, intravenous fluids and antibiotics is occasionally used in elderly and very sick patients.

Gastric outlet obstruction

The obstruction may be prepyloric, pyloric or duodenal. The obstruction occurs either because of an active ulcer with surrounding oedema or because the healing of an ulcer has been followed by scarring. However, obstruction due to peptic ulcer disease and gastric malignancy are now uncommon; Crohn's disease or external compression from a pancreatic carcinoma are more common causes. Adult hypertrophic pyloric stenosis is a rare cause.

The stomach becomes full of gastric juice and ingested fluid and food, giving rise to the main symptom of vomiting, usually without pain as the characteristic ulcer pain has abated owing to healing.

Vomiting is infrequent, projectile, large in volume, and the vomitus contains particles of previous meals. On examination of the abdomen there may be a succussion splash. The diagnosis is made by endoscopy but can be suspected by the nature of the vomiting; by contrast, psychogenic vomiting is frequent, small volume and usually noisy.

Severe or persistent vomiting causes loss of acid from the stomach and a metabolic alkalosis (see p. 678). Vomiting will often settle with intravenous fluid and electrolyte replacement, gastric drainage via a nasogastric tube and potent acid suppression therapy. Endoscopic dilatation of the pyloric region is useful and, overall, 70% of patients can be managed without surgery.

Surgical treatment and its long-term consequences

Once the mainstay of treatment, surgery is now used in peptic ulcer disease only for complications including:

- recurrent uncontrolled haemorrhage: the bleeding vessel is ligated
- perforation, which is oversewn.

No other procedure, such as gastrectomy or vagotomy, is required.

In the past, two types of operation were performed:

- *Partial gastrectomy.*
- *Vagotomy.* Initially this was a truncal vagotomy and required a gastric drainage procedure such as pyloroplasty or gastro-jejunostomy. In later years this was usually highly selective vagotomy or proximal gastric vagotomy, in which only the nerves supplying the parietal cells were transected, and therefore no drainage of the stomach was required.

Long-term complications of surgery which are still seen occasionally include:

- *Recurrent ulcer.* If this occurs, check for *H. pylori*; rule out Zollinger–Ellison syndrome (see p. 385).
- *Dumping.* This term describes a number of upper abdominal symptoms (e.g. nausea and distension associated with sweating, faintness and palpitations) that occur in patients following gastrectomy or gastroenterostomy. It is due to 'dumping' of food into the jejunum, causing rapid fluid shifts from plasma to dilute the high osmotic load with reduction of blood volume. The symptoms are usually mild and patients adapt to them. It is rare for it to be a long-term problem, and if so, the symptoms usually have a functional element. Hypoglycaemia can also occur.
- *Diarrhoea* was chiefly seen after vagotomy. Recurrent severe episodes occurred in about 1% of patients. Antidiarrhoeals are the usual treatment.
- *Nutritional complications*: in the long-term almost any gastric surgery, but particularly gastrectomy, may be followed by:
  - iron deficiency, due to poor absorption
  - folate deficiency, usually due to poor intake
  - vitamin B_{12} deficiency, due to intrinsic factor deficiency
  - weight loss, usually due to reduced intake.

Other *H. pylori*-associated diseases

- *Gastric adenocarcinoma.* The incidence of distal (but not proximal) gastric cancer parallels that of *H. pylori* infection in countries with a high incidence of gastric cancer. Serological studies show that people infected with *H. pylori* have a higher incidence of distal gastric carcinoma (see p. 264).
- *Gastric B cell lymphoma.* Over 70% of patients with gastric B cell lymphomas (mucosal-associated lymphoid tissue – MALT) have *H. pylori*. *H. pylori* gastritis has been shown to contain the clonal B cell that eventually gives rise to the MALT lymphoma (see p. 266).

GASTROPATHY AND GASTRITIS

'*Gastritis*' indicates inflammation associated with mucosal injury (although the term is often used loosely by endoscopists to describe 'redness'), and '*gastropathy*' indicates epithelial cell damage and regeneration without inflammation. Several classifications of gastritis (e.g. Sydney classification) have been proposed but are controversial due to lack of correlation, for example between endoscopic and histological findings. *H. pylori* infection is the commonest cause of gastritis (80%). Autoimmune gastritis is seen in 5% while the remaining causes include viruses (e.g. cytomegalovirus and herpes simplex), duodenogastric reflux and specific causes, e.g. Crohn's.

Gastropathy is usually caused by irritants (drugs, NSAIDs and alcohol), bile reflux, hypovolaemia and chronic congestion. Acute erosive/haemorrhagic gastropathy can also be seen after severe stress (stress ulcers) and secondary to burns (Curling ulcers), trauma, shock, renal failure or in portal hypertension (called portal gastropathy). The underlying mechanism for these ulcers is unknown but may be related to an alteration in mucosal blood flow.

NSAIDs, *Helicobacter* and ulcers

Aspirin and other NSAIDs deplete mucosal prostaglandins by inhibiting the cyclo-oxygenase (COX) pathway, which leads to mucosal damage. Cyclo-oxygenase occurs in two main

The stomach and duodenum
Gastropathy and gastritis
Management of dyspepsia in the community
Gastric tumours

forms: COX-1, the constitutive enzyme, and COX-2, inducible by cytokine stimulation in areas of inflammation. NSAIDs more specific for COX-2 are available and these drugs have less effect on the COX-1 enzyme in the gastric mucosa. They still produce gastric mucosal damage but less than with other conventional NSAIDs.

Fifty per cent of patients taking regular NSAIDs will develop gastric mucosal damage and approximately 30% will have ulcers on endoscopy. Only a small proportion of patients have symptoms (about 5%) and only 1–2% have a major problem, i.e. GI bleed. Because of the large number of patients on NSAIDs including low-dose aspirin for vascular prophylaxis, this is a significant problem, particularly in the elderly.

H. pylori and NSAIDs are independent and synergistic risk factors for the development of ulcers. In a meta-analysis, the odds ratio (OR) for the incidence of peptic ulcer was 61.1 in patients infected with *H. pylori* and also taking NSAIDs, compared with uninfected controls not taking NSAIDs.

Alcohol in high concentration also damages the gastric mucosal barrier and is associated with acute gastric mucosal erosions and subepithelial haemorrhage which can lead to upper GI bleeding.

Treatment

- Stop the ingestion of NSAIDs.
- A PPI should be given.
- *H. pylori* eradication therapy if *H. pylori* positive.

In many patients with severe arthritis, stopping NSAIDs may not be possible. Therefore use:

- An NSAID with low GI side-effects at lowest dose possible (see p. 518) or if there is no cardiovascular risk, a COX-2 NSAID can be used (see p. 518).
- *Prophylactic cytoprotective therapy*, e.g. PPI or misoprostol (a synthetic analogue of prostaglandin E1 – 800 µg/day) for all high-risk patients, i.e. over 65 years; those with a peptic ulcer history, particularly with complications, and patients on therapy with corticosteroids or anticoagulants. Double dose H_2-receptor antagonists (H_2RAs) and PPIs reduce the risk of endoscopic duodenal and gastric ulcers and are better tolerated than misoprostol, which causes diarrhoea.

Autoimmune gastritis

This affects the fundus and body of the stomach (pangastritis), leading to atrophic gastritis and loss of parietal cells with achlorhydria and intrinsic factor deficiency causing the clinical syndrome of 'pernicious anaemia'. Serum autoantibodies to gastric parietal cells are common and non-specific: antibodies to intrinsic factor are rarer and more significant (see p. 399).

Ménétrier's disease

This is a rare condition with characteristic giant gastric folds, mainly in the fundus and the body of the stomach. Histologically there is hyperplasia of the gastric pits, atrophy of glands and an overall increase of mucosal thickness. Hyperchlorhydria is usually present. The patient may complain of epigastric pain. Peripheral oedema may occur because of hypoalbuminaemia due to protein loss through the gastric mucosa. It is probably premalignant but the rarity of the condition makes this uncertain. Treatment is unclear; some patients have responded to *Helicobacter* eradication and some improve spontaneously. Anti-secretory drugs are usually given. A few patients will require surgery.

MANAGEMENT OF DYSPEPSIA IN THE COMMUNITY

Meal-related upper abdominal symptoms (commonly known by the unhelpful term 'dyspepsia') are very common. Over-the-counter antacids and H_2-receptor antagonists (H_2RAs) are available and are widely used.

A history suggestive of gastro-oesophageal reflux disease (GORD) should be treated with antacids, lifestyle measures, and H_2RAs or PPIs if necessary. Treatment may be on-demand or continuous, and may need to be long-term.

In people under 45 years, gastric cancer or any pathology is very rare. Initial investigation with endoscopy is therefore not recommended. Blind *Helicobacter* eradication or 'test and treat' strategies are both used; their effectiveness depends on the prevalence of *Helicobacter* and local resistance patterns. Neither endoscopy nor 'test and treat' are cost effective. Further investigation may be necessary to reassure those *H. pylori*-negative subjects and those who remain symptomatic after successful *H. pylori* eradication, though organic disease is uncommon.

Older people (over 45 years) with new-onset dyspepsia and patients with alarm symptoms are traditionally investigated with endoscopy to exclude cancer and ulcers. Much investigation in this group is driven by the fear of missing a treatable cancer, though functional dyspepsia is by far the commonest diagnosis.

Treatments used for dyspepsia (often over-the-counter) include:

- *Antacids* (see p. 252).
- *H_2-receptor antagonists*. These have molecular structures that fit the H_2-receptors on the parietal cells. They can produce up to 80% reduction in nocturnal acid production but are less effective by day. There is little difference between the several second-generation H_2RAs such as ranitidine and famotidine, but they have fewer side-effects and cross-reactivity with other medication, e.g. warfarin.
- *Proton pump inhibitors* (see p. 252) are very effective for reflux symptoms, ulcer pain and healing but do not cure ulcer disease alone. They are also used empirically in functional dyspepsia, despite their cost.

GASTRIC TUMOURS

Adenocarcinoma

Gastric cancer is currently the fourth most common cancer found world-wide and the second leading cause of cancer-related mortality. The incidence increases with age (peak incidence 50–70 years), and it is rare under the age of 30 years. The highest incidences of the disease are found in Eastern Asia, Eastern Europe and South America. The incidence in men is twice that in women and varies throughout the world, being high in Japan (M: 53/100 000, F: 21.3/100 000) and Chile and relatively low in the USA (M: 7/100 000, F:

2.9/100 000). In the UK carcinoma of the stomach (see Fig. 6.1) is the seventh commonest cancer (M: 14/100 000, F 6/100 000) and the eighth most common fatal cancer. Although the overall world-wide incidence of gastric carcinoma is falling even in Japan, probably due to reductions in incidence of *Helicobacter*, proximal gastric cancers are increasing in the West.

Epidemiology and pathogenesis

- There is a strong link between *H. pylori* infection and distal gastric cancer. *H. pylori* is recognized by the International Agency for Research in Cancer (IARC) as a Group 1 (definite) gastric carcinogen. *H. pylori* infection causes chronic gastritis which eventually leads to atrophic gastritis and premalignant intestinal metaplasia (Fig. 6.20). Much of the earlier epidemiological data (i.e. the increase of cancer in lower socio-economic groups) can be explained by the intrafamilial spread of *H. pylori*.
- *Dietary factors* may also be involved (as both initiators and promoters) and have separate roles in carcinogenesis. Diets high in salt probably increase the risk. Dietary nitrates can be converted into nitrosamines by bacteria at neutral pH, and nitrosamines are known to be carcinogenic in animals but the evidence in human carcinogenesis is limited. Nitrosamines are also present in the stomach of patients with achlorhydria, who have an increased cancer risk. Consumption of diets high in non-starchy vegetables, allium vegetables (e.g. garlic), fruits and low in salt protect against cancer.
- *Smoking tobacco* is associated with an increased incidence of stomach cancer.
- The commonest *genetic abnormality* is a loss of heterozygosity (LOH) of tumour suppressor genes such as p53 (in 70% of cancers as well as in pre-cancerous states) and the APC gene (in over a third of gastric cancers). These abnormalities are similar to those found in colorectal cancers. Some rare families with diffuse gastric cancer have been shown to have mutations in the E-cadherin gene (CDH-1). There is a higher incidence of gastric cancer in blood group A patients.

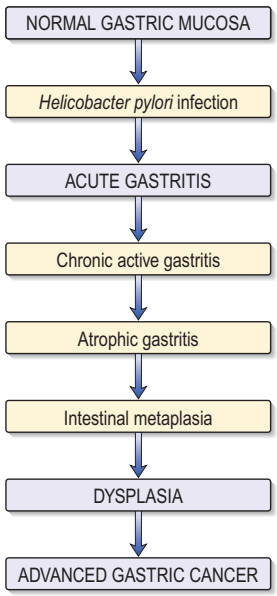


Fig. 6.20 **Flow diagram showing the development of gastric cancer associated with *H. pylori* infection.**

- *First-degree relatives* of patients with gastric cancer have two- to three-fold increased relative risk of developing the disease, but this may be environmental rather than inherited.
- *Pernicious anaemia* carries a small increased risk of gastric carcinoma due to the accompanying atrophic gastritis.
- There is an increased *risk of gastric cancer* after a partial gastrectomy (postoperative stomach) whether performed for a gastric or a duodenal ulcer, probably due to untreated *H. pylori* infection.

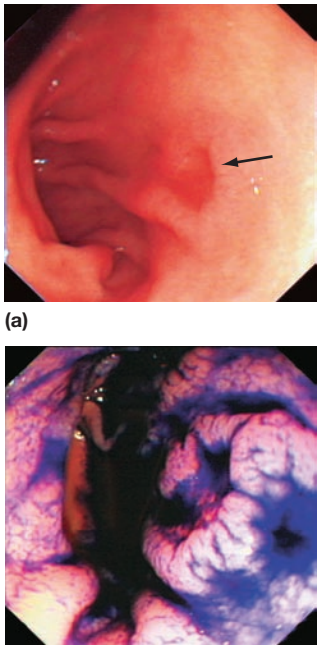
Benign gastric ulcers have not been shown to develop into gastric cancer. It can, however, be difficult to differentiate a benign ulcer from a malignant ulcer, as 10–15% of malignant ulcers can heal on medical treatment. Thus it was originally thought that gastric ulcers could become malignant.

Screening

Earlier diagnosis has been advocated in an attempt to improve the poor prognosis of gastric cancer. Screening is discussed on page 454. Although the incidence of gastric cancer is falling in Japan where aggressive screening by barium meal and endoscopy is pursued, there is no evidence that screening has had an effect on overall mortality. Similarly, early investigation of dyspepsia has had little effect on mortality, possibly because of the relatively low incidence of cancer.

Early gastric cancer (Fig. 6.21)

Early gastric cancer is defined as a carcinoma that is confined to the mucosa or submucosa, regardless of the presence of lymph node metastases. It is associated with 5-year survival rates of approximately 90%, but many of these patients would have survived 5 years without treatment. In Japan, mass screening with mobile X-ray units has increased



(a)

(b)

Fig. 6.21 **Early gastric cancer: (a)** endoscopic view – showing subtle changes in mucosa; **(b)** same view following an ink spray.

the proportion of early gastric cancers (EGC) diagnosed. In a large series of patients with gastric cancer from the UK, only 0.7% were identified as having EGC. They are usually detected by chance as although EGC exists in western populations, endoscopists do not readily recognize it at present.

Pathology

There are two major types of gastric cancer:

- *Intestinal (type 1)*, with well-formed glandular structures. The tumours are polypoid or ulcerating lesions with heaped-up, rolled edges. Intestinal metaplasia is seen in the surrounding mucosa, often with *H. pylori*. This type is more likely to involve the distal stomach and occur in patients with atrophic gastritis. It has a strong environmental association.
- *Diffuse (type 2),* with poorly cohesive cells that tend to infiltrate the gastric wall, may involve any part of the stomach, especially the cardia, and have a worse prognosis than the intestinal type. Unlike type 1 gastric cancers, type 2 cancers have similar frequencies in all geographic areas.

50% of gastric cancer in western countries now occurs in the proximal stomach.

Clinical features

Symptoms

Fifty per cent of patients with EGC discovered at screening have no symptoms. Most patients with carcinoma of the stomach have advanced disease at the time of presentation. The most common symptom of advanced disease is epigastric pain, indistinguishable from the pain of peptic ulcer disease, it may be relieved by food and antacids. The pain can vary in intensity, but may be constant and severe, and there may also be nausea, anorexia and weight loss. Vomiting is frequent and can be severe if the tumour encroaches on the pylorus. Dysphagia can occur with tumours involving the fundus. Gross haematemesis is unusual, but anaemia from occult blood loss is frequent. No pattern of symptoms is suggestive of early gastric cancer.

Widely spreading submucosal gastric cancer causes diffuse thickening and rigidity of the stomach wall and is called 'linitis plastica'.

Patients can present at a late stage with malignant ascites or jaundice due to liver involvement. Metastases also occur in bone, brain and lung, producing appropriate symptoms.

Signs

Weight loss is often the dominant feature. Nearly 50% of patients have a palpable epigastric mass with abdominal tenderness. A palpable lymph node is sometimes found in the supraclavicular fossa (Virchow's node, usually on the left side) and metastases are present in up to one-third of patients at presentation. This cancer is the most frequently associated with dermatomyositis (see p. 547) and acanthosis nigricans.

Diagnosis

- **Gastroscopy** (Fig. 6.22). Gastroscopy is usually performed so that biopsies can be taken for histological assessment and to exclude lymphoma. Positive biopsies can be obtained in almost all cases of obvious carcinoma, but a negative biopsy does not necessarily rule out the diagnosis. For this reason, 8–10 biopsies should be taken from suspicious lesions. Superficial brushings for cytology further improve the diagnostic rate.
- **Full blood count** and liver biochemistry.
- **Barium meal**. Although used as a screening test in Japan, this has largely been replaced by gastroscopy. A good quality double-contrast barium meal has a diagnostic accuracy of up to 90%. The carcinoma is usually seen as a filling defect or an irregular ulcer with rolled edges. With a diffuse (linitis plastica) infiltrating cancer, the X-ray may show a narrowed and rigid stomach.

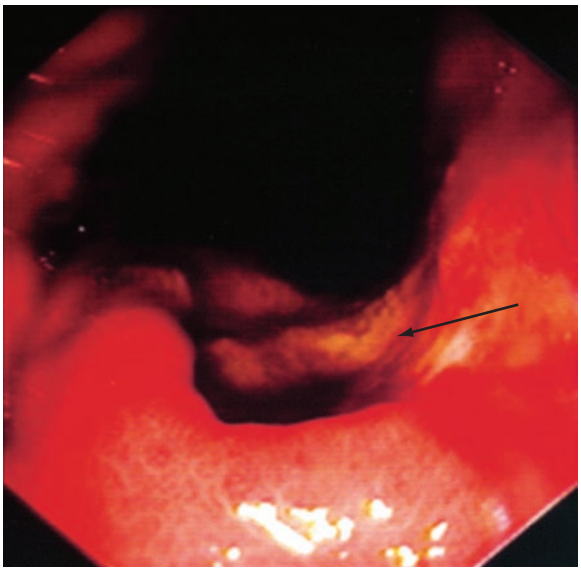


Fig. 6.22 **Carcinoma of the stomach.** Endoscopic picture showing a large irregular ulcer with rolled edges.

Staging

- **CT scan** of the chest and abdomen: CT with a gastric water load can demonstrate gastric wall thickening, lymphadenopathy and lung and liver secondaries, but has limited ability to determine the depth of local tumour invasion.
- **Transabdominal ultrasound** can demonstrate masses, gastric wall thickening, lymphadenopathy and liver secondaries.
- **Endoscopic ultrasound** is useful for local staging to demonstrate the depth of penetration of the cancer through the gastric wall and extension into local lymph nodes. It complements CT and ultrasound but is only relevant in the absence of metastatic disease.
- **Laparoscopy** is useful in patients being considered for surgery to exclude serosal disease.
- **PET and CT/PET** can be helpful in further delineation of the cancer.

The TNM classification is used (see p. 456). The tumour grade (T) indicates depth of tumour invasion, N denotes the presence or absence of lymph nodes, M indicates presence or absence of metastases. TNM classification is then combined into stage categories 0–4. At presentation, two-thirds of patients are at stage 3 or 4, i.e. advanced disease (Table 6.6). The histological grade of the tumour also determines survival.

Table 6.6 Gastric cancer – staging and 5-year survival rates

Stage	TNM stage	5-year survival (%)
1	T1N0M0, T1N1M0 or T2N0M0	88
2	T1N2M0, T2N1M0 or T3N0M0	65
3a	T2N2M0, T3N1M0 or T4N0M0	35
3b	T3N2M0	35
4	T4N1–3M0, TxN3M0 or TxNxM1*	5

T, tumour; N, nodes; M, metastases.
*Tx indicates any T stage; Nx, any N stage.

Treatment

Early non-ulcerated mucosal lesions can be removed endoscopically, but suitable lesions are rare outside Japan.

Surgery remains the most effective form of treatment if the patient is operable. Careful selection has reduced the numbers undergoing surgery and has improved the overall surgical 5-year survival rates to around 30%. Five-year survival rates in 'curative' operations are as high as 50%. Surgery and combined chemo-radiotherapy and treatment of advanced disease is described on page 490. Peri-operative chemotherapy with epirubicin, cisplatin and infusional 5-fluorouracil (ECF) (see Table 9.10) has improved 5-year survival in operable gastric and lower oesophageal adenocarcinomas from 23% to 36%. An alternative regimen is oral epirubicin, oxalplatin and capecitabine. Despite the improved results, the overall survival rate for a patient with gastric carcinoma has not dramatically improved, with a maximum 10% 5-year survival rate overall. Palliative care with relief of pain and counselling are usually required.

Gastrointestinal stromal tumours (GIST)

GI stromal tumours (GISTs) are a subset of GI mesenchymal tumours of varying differentiation. They are usually asymptomatic and found by chance but they can occasionally ulcerate and bleed. There are 200–900 new cases each year in the UK. GISTs mostly affect people between 55 and 65.

These tumours used to be classified as GI leiomyomas, leiomyosarcomas, leiomyoblastomas or Schwannomas. Truly benign leiomyomas do occur, mainly in the oesophagus. GISTs are now recognized as a distinct group of mesenchymal tumours and comprise about 80% of GI mesenchymal tumours. They are of stromal origin, thought to share a common ancestry with the interstitial cells of Cajal. They have varying differentiation. Mutations occur in the cellular proto-oncogene KIT (which leads to activation and cell-surface expression of the tyrosine kinase KIT (CD 117)) in 80% and also in platelet-derived growth factor receptor alpha (PDGFRA) in up to 10% of patients.

Treatment is surgical as far as possible. These tumours generally grow slowly but may be malignant. Imatinib, a tyrosine kinase inhibitor (p. 464), is used for unresectable or metastatic disease, but some patients are resistant; sunitinib is an alternative.

Primary gastric lymphoma

Mucosa-associated lymphatic tissue lymphomas are indolent B cell marginal zone lymphomas primarily involving sites other than lymph nodes (gastrointestinal tract, thyroid, breast or skin). They constitute about 10% of all types of non-Hodgkin's lymphoma (NHL).

Presentation

Most patients are diagnosed in their sixties with stage I or stage II disease outside the lymph nodes. Patients have stomach pain, ulcers or other localized symptoms, but rarely have systemic complaints such as fatigue or fever.

Causes

About 90% of cases are due to *H. pylori* infection. Chromosome abnormalities t(1;14)(p22;q32) and t(11;18)(q21;q21) have also been noted in this form of NHL.

Treatment

Eradication of *H. pylori* infection may resolve cases of local gastric involvement. After standard antibiotic regimens, 50% of patients show resolution at 3 months. Other patients may resolve after 12–18 months of observation. Stage III or IV disease is treated with surgery or CHOP chemotherapy with or without radiation. The prognosis is good, with an estimated 90% 5-year survival.

Gastric polyps

Gastric polyps are found in about 1% of endoscopies, usually by chance. They rarely produce symptoms.

Endoscopic excision is the usual approach to diagnosis and treatment. Occasional large or multiple polyps may require surgery.

Hyperplastic polyps are by far the most common. Most are less than 2 cm. The polyps are rarely premalignant, but may be accompanied by premalignant atrophic gastritis.

Adenomatous polyps are usually solitary lesions in the antrum. Approximately 3% progress to gastric cancer, especially if greater than 2 cm in diameter, but they are not a common cause of gastric cancer (cf colorectal cancer).

Cystic gland polyps contain microcysts that are lined by fundic-type parietal and chief cells. They are located in the fundus and body of the stomach. They are found in otherwise normal subjects, but are especially common in familial polyposis syndromes and have no malignant potential.

Inflammatory fibroid polyps are benign spindle cell tumours infiltrated by eosinophils. Excision of these polyps is indicated because of their propensity to enlarge and cause obstruction.

FURTHER READING

Delaney B, Ford AC, Forman D et al. Initial management strategies for dyspepsia. *Cochrane Database of Systematic Reviews* 2005; Oct **19**(4): CD001961.

Gerrits MM, van Vliet A, Kuipers E et al. *Helicobacter pylori* and antimicrobial resistance. *Lancet Infectious Diseases* 2006; **6**: 699–709.

Rostom A. Prevention of NSAID induced gastrointestinal ulcers. *Cochrane Rev Abstract* 2007. The Cochrane Collaboration: posted July 01, 2007.

Rubin BP, Heinrich MC, Corless CL et al. Gastrointestinal stromal tumour. *Lancet* 2007; **369**: 1731–1741.

Stenning SP, Cunningham D, Allum H. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *New England Journal of Medicine* 2006; **355**: 11–20.

World Cancer Research Fund/American Institute for Cancer Research. *Food, Nutrition, Physical Activity and the Prevention of Cancer: A Global Perspective*. Washington DC: AICR, 2007.

Wündisch T, Thiede C, Morgner A et al. Long-term follow-up of gastric malt lymphoma after *Helicobacter pylori* eradication. *Journal of Clinical Oncology* 2005; **23**: 8018–8024.

ACUTE AND CHRONIC GASTROINTESTINAL BLEEDING

This section should be read in conjunction with the descriptions of the specific conditions mentioned.

Acute upper gastrointestinal bleeding

The cardinal features are haematemesis (the vomiting of blood) and melaena (the passage of black tarry stools; the black colour due to blood altered by passage through the gut). Melaena can occur with bleeding from any lesion proximal to the right colon. Following a bleed from the upper GI tract, unaltered blood can appear per rectum, but the bleeding must be massive and is almost always accompanied by shock. The passage of dark blood and clots without shock is always due to lower GI bleeding.

Aetiology (Fig. 6.23)

Peptic ulceration is the commonest cause of serious and life-threatening gastrointestinal bleeding. The relative incidence of causes depends on the patient population; overall the incidence has fallen. In the developing world haemorrhagic viral infections (see Table 4.22, p. 114) can cause significant gastrointestinal bleeding.

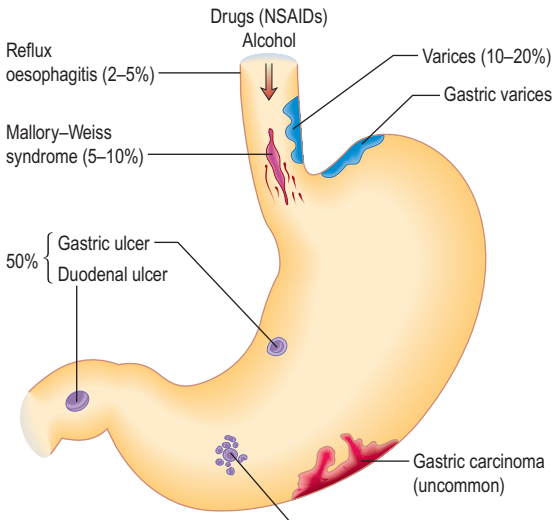


Fig. 6.23 **Causes of upper gastrointestinal haemorrhage.** The approximate frequency is also given.

Drugs. Aspirin (even 75 mg a day) and other NSAIDs can produce ulcers and erosions. These agents are also responsible for GI haemorrhage from both duodenal and gastric ulcers, particularly in the elderly. They are available over the counter and patients may not be aware they are taking aspirin or an NSAID. Corticosteroids in the usual therapeutic doses probably have no influence on GI haemorrhage. Anticoagulants do not cause acute GI haemorrhage per se but bleeding from any cause is greater if the patient is anticoagulated.

Clinical approach to the patient

All cases with a recent (i.e. within 48 hours) significant GI bleed should be seen in hospital. In many, no immediate treatment is required as there has been only a small amount of blood loss. Approximately 85% of patients stop bleeding spontaneously within 48 hours.

Scoring systems have been developed to assess the risk of rebleeding or death (Rockall) and need for intervention (Blatchford) such as transfusion or endoscopic therapy.

Table 6.7 shows the Rockall score which is based on clinical and endoscopy findings. The Blatchford score uses the level of plasma urea, haemoglobin and clinical markers but not endoscopic findings.

The following factors affect the risk of rebleeding and death:

- age
- evidence of co-morbidity, e.g. cardiac failure, ischaemic heart disease, renal disease and malignant disease
- presence of the classical clinical features of shock (pallor, cold peripheries, tachycardia and low blood pressure)
- endoscopic diagnosis
- ulcer with active bleeding or endoscopic stigmata of recent bleeding
- clinical signs of chronic liver disease.

Bleeding associated with liver disease is often severe and recurrent if it is from varices. Splenomegaly suggests portal hypertension but its absence does not rule out oesophageal varices. Liver failure can develop.

Immediate management

This is shown in Emergency Box 6.1. In addition, stop NSAIDs, aspirin and warfarin if patients are taking them.

Many hospitals have multidisciplinary specialist teams with agreed protocols and these should be followed carefully. Patients should be managed in high-dependency beds. Oxygen should be given by face mask and the patient should be kept nil by mouth until endoscopy has been performed.

Patients with large bleeds and clinical signs of shock require urgent resuscitation. Details of the management of shock are given in Figure 15.20.

Blood volume

The major principle is to rapidly restore the blood volume to normal. This can be best achieved by transfusion of red cell concentrates via one or more large-bore intravenous

FURTHER READING

Gralnek IM, Baskin AN, Bardon M. Management of acute bleeding from a peptic ulcer. *New England Journal of Medicine* 2008; **359**: 928–937.

Table 6.7 Risk assessment in non-variceal upper gastrointestinal haemorrhage

(a) Rockall risk assessment score

Variable	Score			
	0	1	2	3
Age (years)	<60	60–79	>79	–
Circulation	BP > 100 mmHg Pulse < 100 b.p.m.	BP > 100 mmHg Pulse > 100 b.p.m.	BP < 100 mmHg Pulse > 100 b.p.m.	–
Comorbidity	None	–	Cardiac disease, any other major comorbidity	Renal failure, liver failure, disseminated malignancy
Endoscopic diagnosis	Mallory–Weiss tear, no lesion	All other diagnoses	Malignancy of the upper GI tract	–
Major SRH	None, or dark spots	–	Blood in the upper GI tract, adherent clot or spurting vessel	–

BP, blood pressure; GI, gastrointestinal; SRH, stigmata of recent haemorrhage.

(b) Rebleed and mortality risk according to Rockall score

Risk score	Predicted rebleed (%)	Predicted mortality (%)
0	5	0
1	3	0
2	5	0
3	11	3
4	14	5
5	24	11
6	33	17
7	44	27
8+	42	41

Emergency Box 6.1

Management of acute gastrointestinal bleeding

- History and examination. Note co-morbidity
- Monitor the pulse and blood pressure half-hourly
- Take blood for haemoglobin, urea, electrolytes, liver biochemistry, coagulation screen, group and crossmatching (2 units initially)
- Establish intravenous access – 2 large bore i.v. cannulae; central line if brisk bleed
- Give blood transfusion/colloid if necessary. Indications for blood transfusion are:
  (a) SHOCK (pallor, cold nose, systolic BP below 100 mmHg, pulse > 100 b.p.m.)
  (b) haemoglobin < 10 g/dL in patients with recent or active bleeding
- Oxygen therapy
- Urgent endoscopy in shocked patients/liver disease
- Continue to monitor pulse and BP
- Re-endoscope for continued bleeding/hypovolaemia
- Surgery if bleeding persists

cannulae; plasma expanders or 0.9% saline is given until the blood becomes available (see p. 906).

Transfusion must be monitored to avoid overload leading to heart failure. The pulse rate and venous pressure are the best guides to adequacy of transfusion. A central venous pressure line is inserted for patients with organ failure who require blood transfusion, and in patients with severe hypotension.

Haemoglobin levels are generally a poor indicator of the need to transfuse because anaemia does not develop immediately as haemodilution has not taken place. However, if the Hb is less than 10 g/dL and the patient has either bled recently or is actively bleeding, transfusion is usually necessary. In most patients the bleeding stops, albeit temporarily, so that further assessment can be made.

Endoscopy

Endoscopy will usually make a diagnosis and endoscopic therapy can be performed if needed. Endoscopy should be performed within 24 hours in patients with significant bleeds. Patients with Rockall scores of 0 or 1 may be candidates for immediate (see over) discharge and outpatient endoscopy the following day, depending on local policy.

After adequate resuscitation, **urgent** endoscopy should be performed in patients with shock, suspected varices or with continued bleeding. Endoscopy can detect the cause of the haemorrhage in 80% or more of cases. In patients with a peptic ulcer, if the stigmata of a recent bleed are seen (i.e. a spurting artery, active oozing, fresh or organized blood clot or black spots) the patient is more likely to rebleed. Calculation of the postendoscopy Rockall score gives an indication of the risk of rebleeding and death.

At first endoscopy:

- Varices should be treated, usually with banding. Stenting for varices is a recent innovation, not yet widely available – see page 350 for management of varices.
- Bleeding ulcers and those with stigmata of recent bleeding should be treated with two haemostatic methods, usually injection with epinephrine (adrenaline) and thermal coagulation (with heater probe, bipolar

probe, laser or argon plasma coagulation) or endoscopic clipping because dual therapy is clearly more effective than monotherapy in reducing rebleeding.
- Antral biopsies should be taken to look for *H. pylori*. A positive biopsy urease test is valid, but a negative test is not reliable. If the urease test is negative, gastric histology should always be performed.

The above procedures reduce the incidence of rebleeding, although they do not significantly improve mortality.

Drug therapy

After diagnosis at endoscopy, intravenous omeprazole 80 mg followed by infusion 8 mg/h for 72 hours should be given to all ulcer patients as it reduces rebleeding rates and the need for surgery. PPI therapy has no effect on mortality in studies in the western world, but there is a positive effect in Oriental studies. H_2-receptor antagonists are of no value.

Uncontrolled or repeat bleeding

Endoscopy should be repeated to assess the bleeding site and to treat, if possible. Surgery is necessary if bleeding is persistent and/or uncontrollable and should aim primarily to control the haemorrhage.

Discharge policy

The patient's age, diagnosis on endoscopy, co-morbidity and the presence or absence of shock and the availability of support in the community should be taken into consideration. In general, all patients who are haemodynamically stable and have no stigmata of recent haemorrhage on endoscopy (Rockall Score pre-endoscopy 0, post-endoscopy ≤ 1) can be discharged from hospital within 24 hours. All shocked patients and patients with co-morbidity need inpatient observation.

Specific conditions

Chronic peptic ulcer. Eradication of *H. pylori* is started as soon as possible (see p. 261). A PPI is continued for 4 weeks to ensure ulcer healing. Eradication of *H. pylori* should always be checked in a patient who has bled and long-term acid suppression given if HP eradication is not possible. If bleeding is not controlled, surgery with ligation of the bleeding vessel is performed to control haemorrhage.

Gastric carcinoma. Most of these patients do not have large bleeds but surgery is occasionally necessary for uncontrolled or repeat bleeding. Usually surgery can be delayed until the patient has been fully evaluated (see p. 265). Oozing from gastric cancer is very difficult to control endoscopically.

Oesophageal varices. These are discussed on page 348.

Mallory–Weiss tear. This is a linear mucosal tear occurring at the oesophagogastric junction and produced by a sudden increase in intra-abdominal pressure. It often occurs after a bout of coughing or retching and is classically seen after alcoholic 'dry heaves'. There may, however, be no antecedent history of retching. Most bleeds are minor and discharge is usual within 24 hours. The haemorrhage may be large but most patients stop spontaneously. Early endoscopy confirms diagnosis and allows therapy if necessary. Surgery with oversewing of the tear is rarely needed.

Bleeding after percutaneous coronary intervention (PCI). This occurs in approximately 2% of patients undergoing PCI (who are on antiplatelet therapy), and has a high mortality of 5–10%. Urgent endoscopy should be performed with appropriate therapy for, for example, an ulcer. A proton pump inhibitor should be given intravenously. Management is difficult as cessation of antiplatelet therapy has a high risk of acute stent thrombosis and also an associated high mortality. Using a risk assessment score (e.g. Rockall, p. 268), a reasonable approach is to stop all antiplatelet therapy in high risk patients but continue in low risk ones. These patients should be under the combined care of a cardiologist and a gastroenterologist.

Prognosis

The mortality from gastrointestinal haemorrhage has not changed from 5–12% over the years, despite many changes in management, mainly because of a demographic shift to more elderly patients with co-morbidity. The lowest mortality rates are achieved in dedicated medical/surgical GI units. Early therapeutic endoscopy has not so far reduced the mortality, although rebleeding episodes are reduced.

Acute lower gastrointestinal bleeding

Massive bleeding from the lower GI tract is rare and usually due to diverticular disease or ischaemic colitis. Common causes of small bleeds are haemorrhoids and anal fissures. The causes of lower GI bleeding are shown in Figure 6.24.

Management

Most acute lower GI bleeds start and stop spontaneously. The few patients who continue bleeding and are haemodynamically unstable need resuscitation using the same principles as for upper GI bleeding (see p. 267). Surgery is rarely required.

A diagnosis is made using the history, examination including rectal examination and the following investigations as appropriate:

- proctoscopy (e.g. anorectal disease, particularly haemorrhoids)

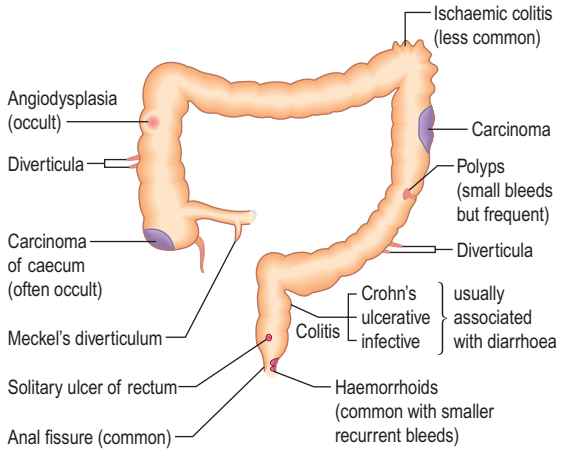


Fig. 6.24 **Causes of lower gastrointestinal bleeding.** The sites shown are illustrative – many of the lesions can be seen in other parts of the colon.

FURTHER READING

BSG Endoscopy Committee Guidelines on non-variceal gastrointestinal haemorrhage. *Gut* 2002; **51** (suppl 4).

FURTHER READING

Pennazlo M, Santucci R, Rondonotti E et al. Outcome of patients with obscure GI bleeding after capsule endoscopy. *Gastroenterology* 2004; **126**: 643–653.

- flexible sigmoidoscopy or colonoscopy (e.g. inflammatory bowel disease, cancer, ischaemic colitis, diverticular disease, angiodysplasia)
- angiography – vascular abnormality (e.g. angiodysplasia). The yield of angiography is low, so it is a test of last resort.

Isolated episodes of rectal bleeding in the young (<45 years) usually only require rectal examination and flexible sigmoidoscopy because the probability of a significant proximal lesion is very low unless there is a strong family history of colorectal cancer at a young age. Individual lesions are treated as appropriate.

Chronic gastrointestinal bleeding

Patients with chronic bleeding usually present with iron-deficiency anaemia (see Ch. 8).

Chronic blood loss producing iron deficiency anaemia in all men, and all women after the menopause, is always due to bleeding from the GI tract. The primary concern is to exclude cancer, particularly of the stomach or right colon, and coeliac disease. Occult blood tests are unhelpful (Box 6.4).

Diagnosis

Chronic blood loss can occur with any lesion of the GI tract that produces acute bleeding (Figs 6.23 and 6.24). However, oesophageal varices usually bleed obviously and rarely present as chronic blood loss. Although uncommon in developed countries, hookworm is the most common world-wide cause of chronic GI blood loss.

History and examination may indicate the most likely site of the bleeding, but if no clue is available it is usual to investigate both the upper and lower GI tract endoscopically at the same session ('top and tail').

Upper gastrointestinal endoscopy is usually performed first. Duodenal biopsies should always be taken to exclude coeliac disease which is a recognized cause of iron deficiency. *Colonoscopy* follows and any lesion should be biopsied or removed, though it is not safe to assume that colonic polyps are the cause of chronic blood loss. *Unprepared CT* scanning is a reasonable test to look for colon cancer in frail patients, and *CT colonography* can be used on the rare occasions colonoscopy fails to reach the caecum.

If gastroscopy, colonoscopy and duodenal biopsy have not revealed the cause, investigation of the small bowel is probably only justified if the anaemia is transfusion-dependent. The diagnostic yield of small bowel follow-through in this situation is very low. *Video capsule endoscopy* is the diagnostic investigation of choice if endoscopy fails to reveal the cause, but has no therapeutic ability. *Push enteroscopy* can examine part of the jejunum and new techniques of balloon-assisted enteroscopy can examine the whole small bowel though the procedure is time consuming. Thermal therapy can also be applied, e.g. to vascular lesions in the small bowel.

If these investigations fail to show the cause, *coeliac axis and mesenteric angiography* may be performed, although the yield is low. Occasionally, intravenous technetium-labelled colloid may be used to demonstrate a potential bleeding site in a Meckel's diverticulum.

Treatment

The cause of the bleeding should be treated, if found. Oral iron is given to treat anaemia (see p. 396). Some patients will require maintenance with regular transfusion as a last resort.

Box 6.4 Measurement of faecal occult blood

This is frequently performed unnecessarily. It is only of value in:

- Premenopausal women – if a history of menorrhagia is uncertain and the cause of iron deficiency is unclear
- Mass population screening for large bowel malignancy
  Advantages: cheap and easy to perform
  Disadvantages: high false-positive rate, leading to unnecessary investigations

THE SMALL INTESTINE

Structure

The small intestine extends from the duodenum to the ileocaecal valve. It is approximately 6 m in length, and 300 m^2 in surface area. The upper 40% is the duodenum and jejunum, the remainder is the ileum. Its surface area is enormously increased by circumferential mucosal folds that have on them multiple finger-like projections called villi. On the villi the surface area is further increased by microvilli on the luminal side of the epithelial cells (enterocytes) (Fig. 6.25).

Each villus consists of a core containing blood vessels, lacteals (lymphatics) and cells (e.g. plasma cells and lymphocytes). The lamina propria contains plasma cells, lymphocytes, macrophages, eosinophils and mast cells. The crypts of Lieberkühn are the spaces between the bases of the villi.

Enterocytes are formed at the bottom of the crypts and migrate toward the tops of the villi, where they are shed. This

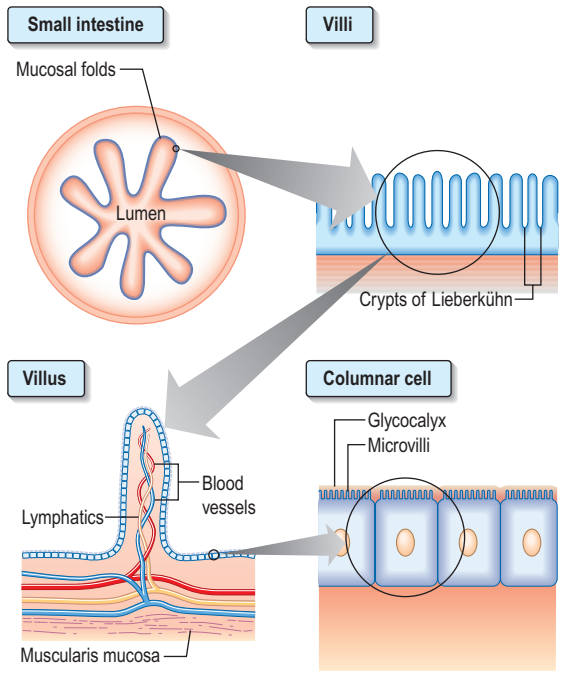


Fig. 6.25 **Structure of the small intestine.**

Ask the authors
www.studentconsult.com

process takes 3–4 days. On its luminal side, the enterocyte is covered by microvilli and a gelatinous layer called the glycocalyx. Intraepithelial lymphocytes lie between the enterocytes. Scattered throughout the gut are peptide-secreting cells. Most of the blood supply to the small intestine is via branches of the superior mesenteric artery. The terminal branches are end arteries – there are no local anastomotic connections.

Enteric nervous system (ENS)

This controls the functioning of the small bowel; it is an independent system that coordinates absorption, secretion, blood flow and motility. It is estimated to contain 10^8 neurones (as many as the spinal cord) contained in two major ganglionated plexuses; the myenteric plexus between the muscular layers of the intestinal wall, and the submucosal plexuses associated with the mucosa. The ENS communicates with the central nervous system via autonomic afferent and efferent pathways but can operate autonomously.

Coordination of small intestinal function involves a complex and poorly understood interplay between many neuroactive mediators and their receptors, ion channels, GI hormones, nitric oxide and other transmitters. For example, 5HT acts via numerous receptors, some of which also have their own subtypes. Tetrodotoxin (TTX) sensitive and TTX resistant sodium channels and acid sensitive ion channels are differentially expressed in different neurones. Acetylcholine, adrenaline, ATP, vasoactive intestinal peptide (VIP) and other hormones and opioids have been shown to have actions in the small bowel, but the exact role for each is far from being understood.

Gut motility

The contractile patterns of the small intestinal muscular layers are primarily determined by the enteric nervous system. The CNS and gut hormones also have a modulatory role on motility. The interstitial cells of Cajal which lie within the smooth muscle appear to govern rhythmic contractions.

During fasting, a distally migrating sequence of motor events termed the migrating motor complex (MMC) occurs in a cyclical fashion. The MMC consists of a period of motor quiescence (phase I) followed by a period of irregular contractile activity (phase II), culminating in a short (5–10 min) burst of regular phasic contractions (phase III). Each MMC cycle lasts for approximately 90 minutes. In the duodenum, phase III is associated with increased gastric, pancreatic and biliary secretions. The role of the MMC is unclear, but the strong phase III contractions propel secretions, residual food and desquamated cells towards the colon. It is named the 'intestinal housekeeper'.

After a meal, the MMC pattern is disrupted and replaced by irregular contractions. This seemingly chaotic pattern lasts typically for 2–5 hours after feeding, depending on the size and nutrient content of the meal. The irregular contractions of the fed pattern have a mixing function, moving intraluminal contents to and fro, aiding the digestive process.

Neuroendocrine peptide production

The hormone-producing cells of the gut are scattered diffusely throughout its length and also occur in the pancreas. The cells that synthesize hormones are derived from neural ectoderm and are known as APUD (amine precursor uptake and decarboxylation) cells. Many of these hormones have very similar structures. Although they can be detected by radioimmunoassay in the circulation, their action is often local.

Gut hormones play a part in the regulation and integration of the functions of the small bowel and other metabolic activities. Their actions are complex and interacting, both with each other and with the ENS (Table 6.8).

Physiology

In the small bowel digestion and absorption of nutrients and ions takes place, as does the regulation of fluid absorption and secretion. The epithelial cells of the small bowel form a physical barrier that is selectively permeable to ions, small molecules and macromolecules. Digestive enzymes such as proteases and disaccharidases are produced by intestinal cells and expressed on the surface of microvilli; others such as lipases produced by the pancreas are associated with the glycocalyx. Some nutrients are absorbed most actively in specific parts of the small intestine; iron and folate in the duodenum and jejunum, vitamin B_{12} and bile salts in the terminal ileum where they have specific receptors.

General principles of absorption

Simple diffusion

This process is non-specific, requires no carrier molecule or energy and takes place if there is a concentration gradient from the intestinal lumen (high concentration) to the bloodstream (low concentration). Vitamin B_{12} can be absorbed from the jejunum by this means.

Facilitated diffusion

Absorption takes place down a concentration gradient, but a membrane carrier protein is involved, conferring specificity on the process. Fructose is an example.

Active transport

Absorption occurs via a specific carrier protein, powered by cellular energy, and thus a substance can be transported against a concentration gradient. Many carrier proteins are powered by ion gradients across the enterocyte wall. For example, glucose crosses the enterocyte microvillous membrane from the lumen into the cell against a concentration gradient by using a *co-transporter* carrier molecule. This is the sodium/glucose co-transporter, SGLT1 (Fig. 6.26). The process is powered by the energy derived from the flow of Na^+ ions from a high concentration outside the cell to a low concentration inside. The sodium gradient across the cell wall is maintained by a separate ATP-consuming Na^+/K^+ exchanger in the basolateral membrane. Glucose leaves the cell on the serosal side by facilitated diffusion via a sodium-independent carrier (GLUT-2) in the basolateral membrane.

Another active transport mechanism operates for Na^+ absorption in the ileum using an Na^+/H^+ *exchange* mechanism powered by the outwardly directed gradient of H^+ across the cell membrane.

Absorption of nutrients in the small intestine

Carbohydrate

Dietary carbohydrate consists mainly of starch, some sucrose and a small amount of lactose. Starch is a polysaccharide made up of numerous glucose units. Its hydrolysis begins in the mouth by salivary amylase, though the majority

Table 6.8 Gut regulatory peptides

Peptide	Localization	Main actions
Gastrin/cholecystokinin family		
Cholecystokinin (CCK) Multiple forms from CCK8 (8 amino acids) to CCK83; 8, 33 and 58 are predominant. Terminal 5 amino acids same as gastrin	Duodenum and jejunum (I cells) Enteric nerves CNS	Causes gall bladder contraction and sphincter of Oddi relaxation. Trophic effects on duodenum and pancreas. Pancreatic secretion (minor role). Role in satiety – acting in CNS
Gastrin	G cells in gastric antrum and duodenum	Stimulates acid secretion. Trophic to mucosa
Secretin-glucagon family		
Secretin	Duodenum and jejunum (S cells)	Stimulates pancreatic bicarbonate secretion
Glucagon	Alpha cells of pancreas	Opposes insulin in blood glucose control
Vasoactive polypeptide (VIP)	Enteric nerves	Intestinal secretion of water and electrolytes. Neurotransmitter. Splanchnic vasodilatation, stimulates insulin release
Peptide-histidine methionine	Enteric nerves	As for VIP (major sequence homology; coded by same gene)
Glucose-dependent insulinotropic peptide (GIP)	Duodenum (K cells) Gastric antrum Ileum	Release by intraduodenal glucose causes greater insulin release by islets than i.v. glucose (incretin effect)
Glucagon-like peptide-1 (GLP-1) Glucagon-like peptide-2 (GLP-2)	Ileum and colon (L cells) Co-secreted with GLP1	Incretin. Stimulates insulin synthesis. Trophic to islet cells. Inhibits glucagon secretion and gastric emptying Stimulates growth of enterocytes
Glicentin	L cells	Stimulates insulin secretion and gut growth, inhibits gastric secretion.
Growth hormone-releasing factor (GHRF)	Small gut	Unclear
Pancreatic polypeptide family		
Pancreatic polypeptide (PP)	Pancreas (PP cells)	Inhibits pancreatic and biliary secretion
Peptide YY (PYY)	Ileum and colon (L cells)	Inhibits pancreatic exocrine secretion. Slows gastric and small bowel transit ('ileal brake'). Reduces food intake and appetite
Neuropeptide Y (NPY)	Enteric nerves	Stimulates feeding. Regulates intestinal blood flow
Other		
Motilin	Whole gut	Increases gastric emptying and small bowel contraction
Ghrelin	Stomach	Stimulates appetite, increases gastric emptying
Gastrin releasing-polypeptide (bombesin)	Whole gut and pancreas	Stimulates pancreatic exocrine secretion and gastric acid secretion
Somatostatin	Stomach and pancreas (D cells) Small and large intestine	Inhibits secretion and action of most hormones
Galanin	Enteric nerves	Inhibits insulin secretion. Regulates food intake
Pancreastatin (derived from chromogranin A)	Whole gut	Inhibits pancreatic exocrine and endocrine secretion
Substance P	Enteric nerves	Enhances gastric acid secretion, smooth muscle contraction
Calcitonin gene-related peptide	Sensory enteric neurones	Very potent vasodilator. Inhibits gastric and pancreatic secretion
Neurotensin	Ileum	Affects gut motility. Increases jejunal and ileal fluid secretion
Insulin	Pancreatic β cells	Increases glucose utilization
Chromogranins	Neuroendocrine cells	Precursor for other regulatory peptides that inhibit neuroendocrine secretion

takes place in the upper intestine by pancreatic amylase. Hydrolysis by amylase is limited because it cleaves α1–4 but not α1–6 glucose–glucose branching links. The breakdown products of the action of amylase, including maltose and maltotriose, together with sucrose and lactose, are hydrolysed on the microvillous membrane by their appropriate oligo- and disaccharidases to form the monosaccharides glucose, galactose and fructose. These monosaccharides are transported across the enterocytes into the blood (Fig. 6.26).

Protein

Dietary protein is digested by pancreatic proteolytic enzymes to amino acids and peptides prior to absorption. These

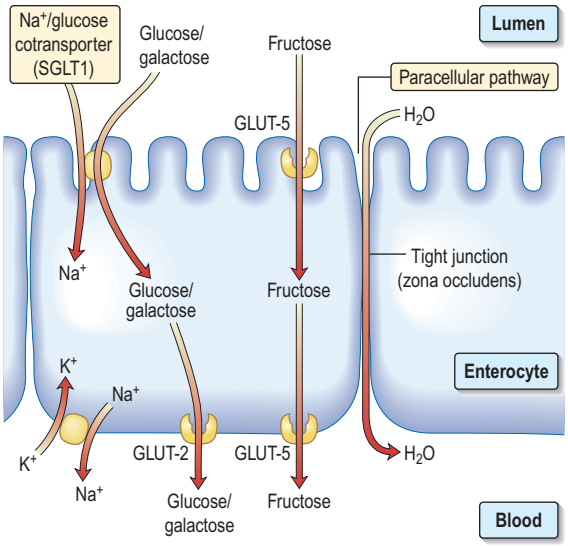


Fig. 6.26 **Solute (glucose, galactose, fructose) transport across the apical membrane** (glucose and galactose are transported via the sodium/glucose cotransporter – SGLT1). Fructose is transported by the facilitative transporter GLUT-5 across the apical and basolateral membranes. Water absorption is mainly cellular but paracellular absorption also occurs. The tight junction is a network of strands and is a dynamic structure. The sodium/potassium-ATPase pump is located in the basolateral membrane.

enzymes are secreted by the pancreas as pro-enzymes and transformed to active forms in the lumen by an enzymic cascade initiated by enterokinase (EK). Protein in the duodenal lumen stimulates the release of EK from duodenal cells. EK activates trypsinogen to trypsin, and this in turn activates the other pro-enzymes, chymotrypsin and elastase.

These enzymes break down protein into oligopeptides. Some di- and tri-peptides are absorbed intact by carrier-mediated processes, while the remainder are broken down into free amino acids by peptidases on the microvillous membranes of the enterocytes, prior to absorption into the cell by a variety of amino acid and peptide carrier systems.

Fat (Fig. 6.27)

Dietary fat consists mainly of triglycerides with some cholesterol and fat-soluble vitamins. Fat is emulsified by mechanical action in the stomach. Bile containing the amphipathic detergents bile acids and phospholipids enters the duodenum following gall bladder contraction. They act to solubilize fat and promote hydrolysis of triglycerides in the duodenum by pancreatic lipase to yield fatty acids and monoglycerides. Bile acids, phospholipids and the products of fat digestion cluster together with their hydrophilic ends on the outside to form aggregations called mixed micelles. Trapped in the centre of the micelles are the hydrophobic monoglycerides, fatty acids and cholesterol. At the cell membrane the lipid contents of the micelles are absorbed, while the bile salts remain in the lumen. Inside the cell the monoglycerides and fatty acids are re-esterified to triglycerides. The triglycerides

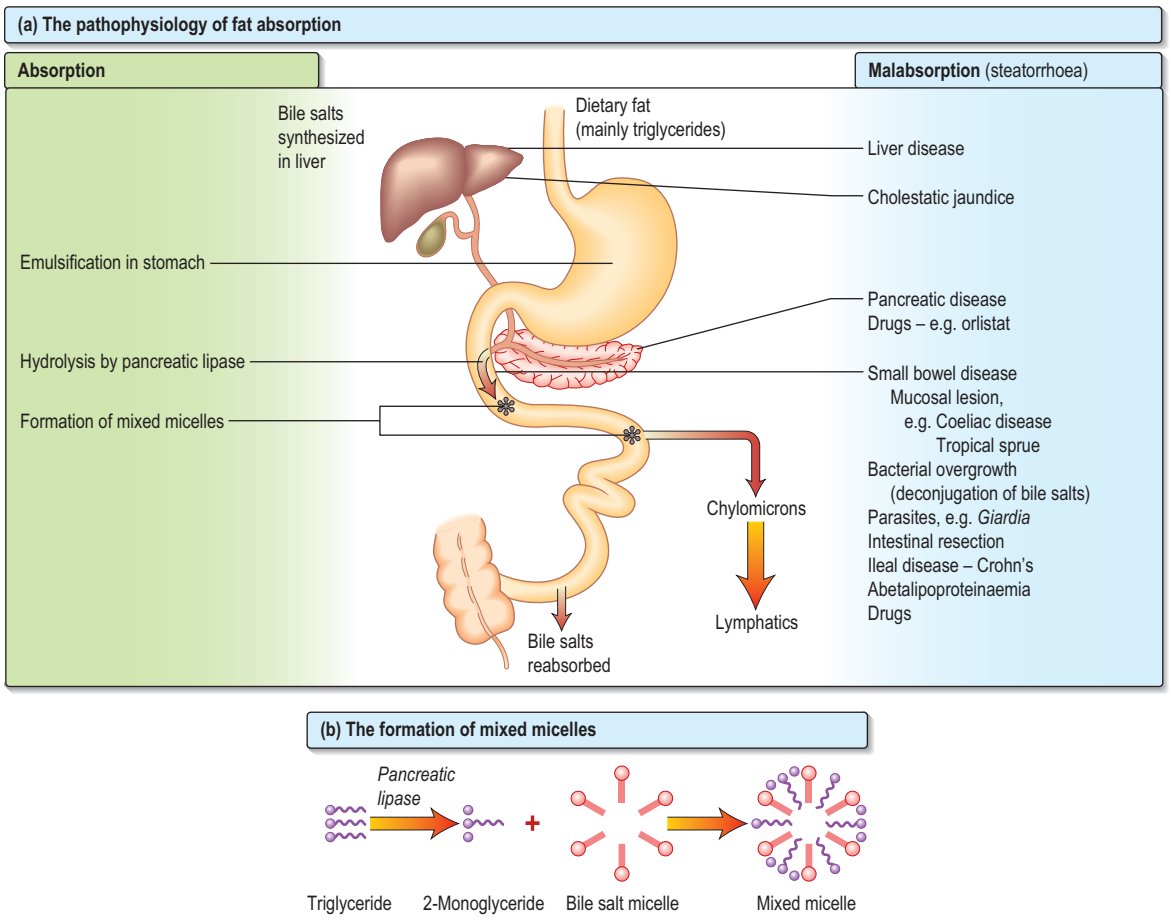


Fig. 6.27 **(a) The pathophysiology of fat absorption. (b) Diagram showing the formation of mixed micelles.**

and other fat soluble molecules (e.g. cholesterol, phospholipids) are then incorporated into chylomicrons to be transported into the lymph.

By contrast, another mechanism exists for medium-chain triglycerides (MCT; fatty acids of chain length 6–12) which are transported via the portal vein with a small amount of long-chain fatty acid. Patients with pancreatic exocrine or bile salt insufficiency can therefore supplement their fat absorption with MCT.

Bile salts are not absorbed in the jejunum, so that the intraluminal concentration in the upper gut is high. They pass down the intestine to be absorbed in the terminal ileum and are transported back to the liver. This enterohepatic circulation prevents excess loss of bile salts (see p. 323).

The pathophysiology of fat absorption is shown in Figure 6.27. Interference with absorption can occur at all stages, as indicated, giving rise to steatorrhoea (>17 mmol or 6 g of faecal fat per day).

Water and electrolytes

Large amounts of water and electrolytes, partly dietary, but mainly from intestinal secretions, are absorbed coupled with absorption of monosaccharides, amino acids and bicarbonate in the upper jejunum. Water and electrolytes are also absorbed paracellularly (between the enterocytes) down electrochemical and osmotic gradients. Additional water and electrolytes are absorbed in the ileum and colon, where active sodium transport is not coupled to solute absorption. Secretion of fluid and electrolytes occur together to maintain the normal functioning of the gut. Secretory diarrhoea (see p. 306) can occur because of defects in intestinal secretory mechanisms.

Water-soluble vitamins, essential metals and trace elements

These are all absorbed in the small intestine. Vitamin B_{12} (see p. 399) and bile salts are absorbed by specific transport mechanisms in the terminal ileum; malabsorption of both these substances often occurs following ileal resection.

Calcium absorption

Calcium absorption is discussed on page 558.

Iron absorption

Iron absorption is discussed on page 393.

Response of the small bowel to antigens and pathogens

The small bowel has a number of mechanisms to prevent colonization and invasion by pathogens while simultaneously preventing inappropriate responses to foreign antigens or the indigenous bacterial population. At the same time commensal bacteria maintain the integrity of the small bowel and play a major role in host physiology.

Mechanisms

Physical defence

- The mucus layer
- continuous shedding of surface epithelial cells
- the physical movement of the luminal contents
- colonization resistance – the ability of the indigenous microbiota to outcompete pathogens for a survival niche in the gut.

Innate chemical defence

- **Enzymes** such as lysozyme and phospholipase A_2 secreted by Paneth cells at the base of the crypts help ensure an infection-free environment in the gut, even in the presence of commensal bacteria.
- **Antimicrobial peptides** are secreted from enterocytes and Paneth cells in response to pathogenic bacteria. These include *defensins*, which are 15–20 amino acid peptides with potent activity against a broad range of pathogens including Gram-positive and Gram-negative bacteria, fungi and viruses. Other peptides that contribute to the defence against organisms include the antibacterial *cathelicidin, BPI,* (Bactericidal Permeability-

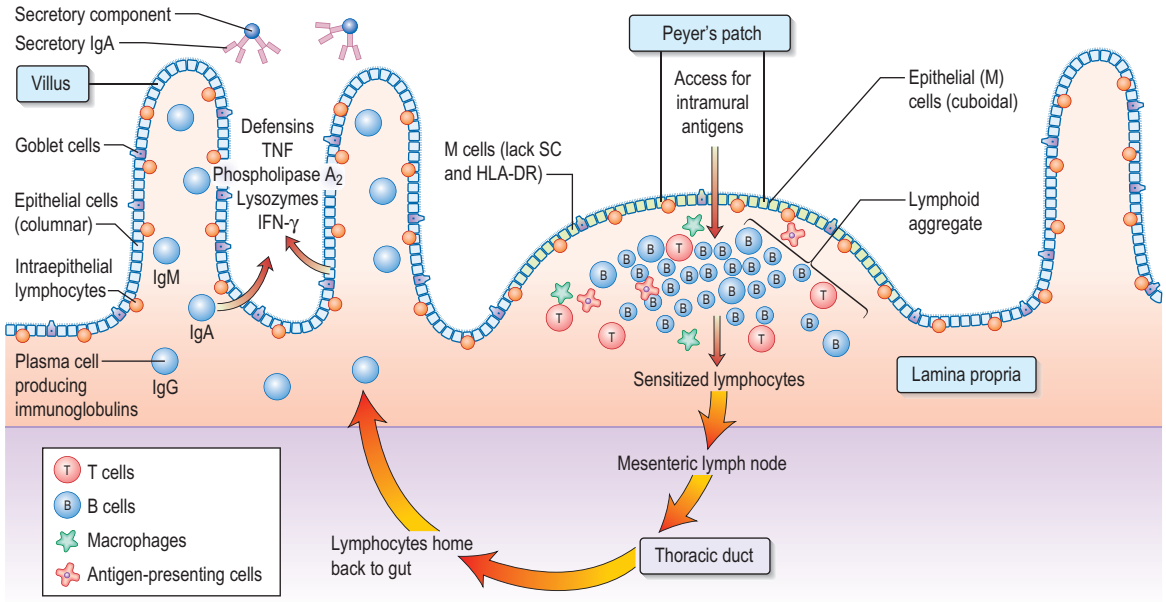


Fig. 6.28 **Small intestinal mucosa with a Peyer's patch, showing the gut-associated lymphoid tissue (GALT).** SC, secretory component; HLA-DR, human lymphocyte antigen-DR; TNF, tumour necrosis factor; IFN-γ, gamma-interferon; M cells, membrane cells.

inducing protein), *antimicrobial lectins*, and *β-Resistin*-like molecule (RELMb) which appears to have anthelminthic effects.

- **Trefoil peptides** are a family of small proteins secreted by goblet cells. They consist of a three-loop structure with intra-chain disuphide bonds which makes the molecules highly resistant to digestion. Their actions include stabilization of mucus, promotion of cell migration over injured areas, and promotion of repair. Three trefoil factors (TFF) are found in humans (TFF1, TFF2 and TFF3) all of which have been implicated in the response to gastrointestinal injury in experimental models. Their molecular mode of action is not yet known.

Innate immunological defence

- **Humoral. IgA** is the principal mucosal antibody. IgA mediates mucosal immunity by agglutinating and neutralizing pathogens in the lumen and preventing colonization of the epithelial surface (Fig. 6.28). Seventy to ninety per cent of the 10^{10} Ig-producing cells per metre of human small bowel are IgA plasma cells. IgA is secreted from immunocytes in the lamina propria as dimers joined by a protein called the 'Joining chain' (J-chain); in this form it is known as polymeric IgA (pIgA). This pIgA is internalized by endocytosis at the basolateral membrane of enterocytes expressing the polymeric IgA Receptor (pIgAR), which is the receptor for J-chain. It crosses the cell as a complex of pIgA/pIgAR and is secreted onto the mucosal surface of the enterocyte as a complex of pIgA with a fragment of pIgAR (secretory component, SC). This molecule is known as secretory IgA (sIgA).
- **B cell sensitization.** Antigens from the lumen of the bowel are transported by M (microfold) and dendritic cells in the specialized (follicle associated epithelium (FAE)). This covers Peyer's patches in the 'dome' region which contain abundant virgin B cells, helper T cells and antigen-presenting cells. Activated B cells then produce IgA locally and are programmed to home back to the lamina propria. They travel through mesenteric lymph nodes and then via the thoracic duct to the blood and back to the small bowel and other mucosal surfaces (such as the airways) where they undergo terminal differentiation into plasma cells. Homing back to the gut is facilitated by the $\alpha 4\beta 7$-integrin on gut-derived lymphocytes binding to MAdCAM-1, uniquely expressed on blood vessels in the gut.

Cellular defence Host defence is also provided by **T lymphocytes** that initiate, activate and regulate adaptive immune responses. Intestinal T lymphocytes occur principally in three major compartments:

1. Organized gut-associated lymphoid tissue (GALT) such as Peyer's patches where mucosal T cell responses are generated, and after which cells leave the organized lymphoid tissue and home back to the mucosa.
2. The lamina propria, containing mostly CD4 cells.
3. The surface epithelium where these lymphocytes are known as intraepithelial lymphocytes (IELs) and are mostly CD8 cells.

T cells are sensitized to antigen in the Peyer's patch lymphoid tissue in a similar fashion to B cells, pass through mesenteric lymph nodes into the thoracic duct and into the circulation, homing back to the small bowel to end up in the lamina propria or the epithelium. It is probable that IELs are cytotoxic cells, capable of killing virally or bacterially infected epithelial cells. CD4 cells in the lamina propria of healthy individuals are highly activated cells, probably protecting against low-grade infections, since loss of these cells, as in HIV infection, leads to colonization of the gut by protozoa such as cryptosporidia.

Commensal bacteria

The relationship between the hundred thousand billion microbes in the human gut and the host are only beginning to be appreciated. Germ-free mice have essentially no mucosal immune system showing that the abundant and activated immune system seen in healthy individuals is driven by the flora, without adverse effects. Bacteria also release chemical signals such as LPS and lipoteichoic acid that are recognized by Toll-like receptors (TLRs) (see p. 59) present on a variety of intestinal cells, priming repair processes and enhancing the ability of the epithelium to respond to injury. Thus, genetically engineered animals with absent TLR signalling pathways develop exaggerated responses to intestinal injury.

Oral tolerance

The immune system must guard against pathogens and toxins while avoiding an excessive response to the multiplicity of food antigens and commensal bacteria. The mechanisms by which tolerance occurs are undoubtedly multiple, including maintaining barrier function to prevent excess antigen uptake, active inhibition via regulatory T cells, and dendritic cells which promote tolerogenic rather than immunogenic T cell responses. All of these are likely to play a role in diseases such as coeliac disease, caused by an excessive T cell response to gluten, or Crohn's disease, where tolerance to the indigenous bacterial population is defective.

Presenting features of small bowel disease

Regardless of the cause, the common presenting features of small bowel disease are:

- *Diarrhoea*. Although a common feature of small bowel disease, 10–20% of patients will have no diarrhoea or any other gastrointestinal symptoms. *Steatorrhoea* occurs when the stool fat is > 17 mmol/day (or 6 g/day). The stools are pale, bulky, offensive, float (because of their increased air content), leave a fatty film on the water in the pan and are difficult to flush away.
- *Abdominal pain* and discomfort. Abdominal distension can cause discomfort and flatulence. The pain has no specific character or periodicity and is not usually severe.
- *Weight loss*. Weight loss is largely due to the anorexia that invariably accompanies small bowel disease. The calorie deficit due to malabsorption is small relative to the reduction in intake.
- *Nutritional deficiencies*. Deficiencies of iron, B_{12}, folate or all of these, leading to anaemia, are the only common deficiencies. Occasionally malabsorption of other vitamins or minerals occurs, causing bruising (vitamin K deficiency), tetany (calcium deficiency), osteomalacia (vitamin D deficiency), or stomatitis, sore tongue and

aphthous ulceration (multiple vitamin deficiencies). Ankle oedema due to hypoproteinamia is due to low intake and intestinal loss of albumin (protein-losing enteropathy).

Physical signs are few and non-specific. If present, they are usually associated with anaemia and the nutritional deficiencies described above.

Abdominal examination is often normal, but sometimes distension and, rarely, hepatomegaly or an abdominal mass are found. Gross weight loss, oedema and muscle wasting is seen only in severe cases. A neuropathy, not always due to B_{12} deficiency, can be present.

Investigation of small bowel disease (Fig. 6.29)

The emphasis in the investigation of malabsorption is on the structural features of the underlying disorder, rather than on the documentation of malabsorption itself.

Blood tests

- **Full blood count and film**. Anaemia can be microcytic, macrocytic or normocytic and the blood film may be dimorphic. Other abnormal cells (e.g. Howell–Jolly bodies, p. 423) may be seen in splenic atrophy associated with coeliac disease.
  - *If the MCV is low*, serum ferritin and the serum soluble transferrin receptor should be measured to differentiate iron deficiency from anaemia of chronic disorder (see p. 397).
  - *If the MCV is high*, serum B_{12}, and serum and red cell folate should be measured.
  - However, with mixed deficiencies, the MCV may be normal. The red cell folate is a good indicator of the presence of small bowel disease. It is frequently low in both coeliac disease and Crohn's disease, which are the two most common causes of small bowel disease in developed countries.
- **Serum albumin** gives some indication of the nutritional status and the presence of intestinal protein loss. Always check urine protein loss to ensure that albumin is not being lost from the kidneys.
- Low **serum calcium** and raised alkaline phosphatase may indicate the presence of osteomalacia due to vitamin D deficiency.
- **Immunological tests**. Measurement of serum antibodies to endomysium and tissue transglutaminase are useful for the diagnosis of coeliac disease.
- **HLA testing** is useful in coeliac disease.

Small bowel anatomy

- **Small bowel follow-through** (see p. 245). This detects gross anatomical defects such as diverticula, strictures and Crohn's disease. Dilatation of the bowel and a changed fold pattern may suggest malabsorption but these are non-specific findings. Gross dilatation is seen in myopathic pseudo-obstruction.

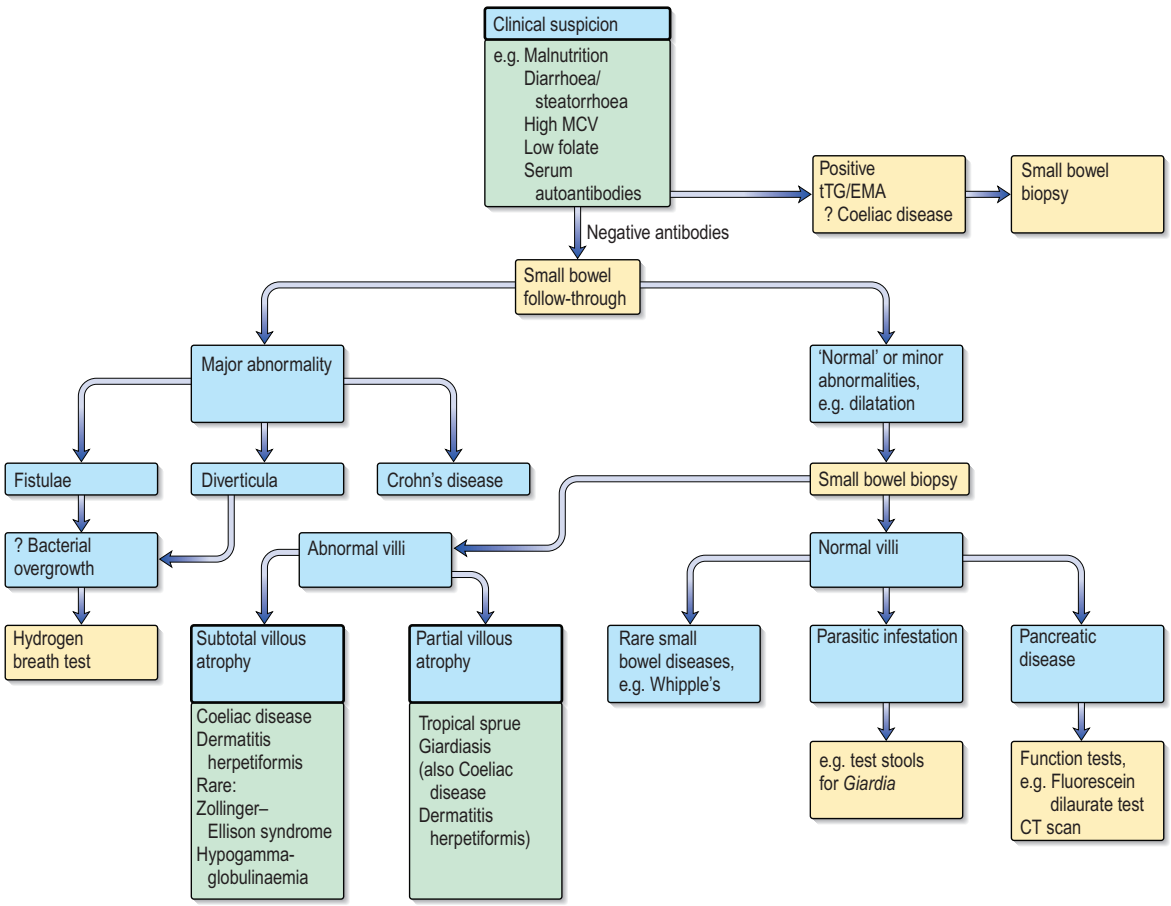


Fig. 6.29 **Flow diagram for investigation of patients with suspected small bowel disease.** EMA, endomysial antibody; tTG, tissue transglutaminase.

- **Small bowel biopsy**. This is used to assess the microanatomy of the small bowel mucosa. Biopsies are usually obtained via an endoscope passed into the duodenum and should be well-orientated for correct evaluation. The histological appearances are described in the sections on individual diseases. A smear of the jejunal juice or a mucosal impression should also be made when *Giardia intestinalis* is suspected.
- **CT scanning** is used to look for small bowel wall thickening, diverticula and for extraintestinal features such as abscesses (e.g. in Crohn's disease).
- **MRI enteroclysis**, which does not involve radiation, is also used to assess and monitor changes in small bowel anatomy.

Tests of absorption

These are required only in *complicated* cases.

- **Fat malabsorption**. The confirmation of the presence of steatorrhoea is only occasionally necessary. Three-day faecal fat analysis, triglyceride breath tests and serum β carotene are now rarely performed. In rare cases when it is essental to confirm steatorrhoea, Sudan III staining of a faecal sample can be used.
- **Lactose tolerance test**. Testing is of little use in adults because lactose intolerance is rarely a clinical problem; patients who are upset by milk usually avoid it. Formal testing involves giving an oral dose of 50 g of lactose and serial measurement of blood glucose over 2 hours (N.B. 500 mL of milk contain 20 g of lactose). There is a high incidence of lactase deficiency in many parts of the world (e.g. the Mediterranean countries, and parts of Africa and Asia).

Other tests

- **Hydrogen breath test**. This is frequently used as a screening test to measure transit time and to detect small bowel bacterial overgrowth. Bacteria are present in the oral cavity so the mouth should be rinsed out with an antiseptic mouthwash beforehand. The appearance of a breath hydrogen peak after oral lactulose is used to estimate mouth to caecum transit time. An earlier rise in the breath hydrogen after lactulose indicates bacterial breakdown in the small intestine. This test is simple to perform and it does not involve radioisotopes. However, interpretation is often difficult with a low sensitivity and specificity.
- **^{14}C-glycocholic acid breath test**. Some bacteria can deconjugate an orally administered ^{14}C labelled bile salt, releasing ^{14}C-glycine, which is metabolized and appears in the breath as $^{14}CO_2$. This test has largely been replaced by the hydrogen breath test.
- **Direct intubation**. Aspiration of intestinal juices can detect bacterial contamination, but it is seldom used. Bacterial counts are performed on aerobic and anaerobic cultures. Chromatography of bile salts can be performed on the aspirate to detect evidence of deconjugation.
- **Tests for pancreatic insufficiency** are used in the differential diagnosis of steatorrhoea.
- **Other blood tests**. Serum immunoglobulins are measured to exclude immune deficiencies. Gut peptides (e.g. VIP) are measured in high-volume secretory diarrhoea, and chromogranins A and B are raised in endocrine tumours.
- **Tests for protein-losing enteropathy (PLE)** (see p. 282). These tests are rarely required unless a low serum albumin is a major clinical feature.
- **Measurement of α_1 antitrypsin clearance** does not require an isotope (see p. 376). α_1 antitrypsin is a large molecule (>50 000 daltons), which is resistant to proteolysis. Simultaneous measurements of serum and stool concentration (24-hour collection) are made.
- **Bile salt loss**. This can be demonstrated by giving oral ^{75}Se-homocholyl taurine (SeHCAT – a synthetic taurine conjugate) and measuring the retention of the bile acid by whole-body counting at 7 days.

MALABSORPTION

In many small bowel diseases, malabsorption of specific substances occurs, but these deficiencies do not usually dominate the clinical picture. An example is Crohn's disease, in which malabsorption of vitamin B_{12} can be demonstrated, but this is not usually the major problem; diarrhoea and general ill-health are the major features.

The major disorders of the small intestine that cause malabsorption are shown in Table 6.9.

Coeliac disease (gluten-sensitive enteropathy)

Coeliac disease (CD) is a condition in which there is inflammation of the mucosa of the upper small bowel that improves when gluten is withdrawn from the diet and relapses when gluten is reintroduced. Up to 1% of many populations are affected, though most have clinically silent disease.

Aetiology

Gluten is the entire protein content of the cereals wheat, barley and rye. Prolamins (gliadin from wheat, hordeins from barley, secalins from rye) are damaging factors. These proteins are resistant to digestion by pepsin and chymotrypin because of their high glutamine and proline content and remain in the intestinal lumen triggering immune responses.

Immunology. Gliadin peptides pass through the epithelium and are deaminated by tissue transglutaminase which increases their immunogenicity. Gliadin peptides then bind to antigen-presenting cells which interact with CD4+ T cells in the lamina propria via HLA class II molecules DQ2 or DQ8. These T cells produce pro-inflammatory cytokines, particularly interferon γ. $CD4^+$ T cells also interact with B cells to produce endomysial and tissue transglutaminase antibodies. Gliadin peptides also cause release of interleukin 15 from enterocytes, activating intraepithelial lymphocytes with a natural killer cell marker. This inflammatory cascade releases metalloproteinases and other mediators that contribute to

Table 6.9 Disorders of the small intestine causing malabsorption

Coeliac disease
Dermatitis herpetiformis
Tropical sprue
Bacterial overgrowth
Intestinal resection
Whipple's disease
Radiation enteritis
Parasite infestation (e.g. *Giardia intestinalis*)

the villous atrophy and crypt hyperplasia which are typical of the disease.

The mucosa of the proximal small bowel is predominantly affected, the mucosal damage decreasing in severity towards the ileum as gluten is digested into smaller 'non-toxic' fragments.

Inheritance. There is an increased incidence of CD within families but the exact mode of inheritance is unknown; 10–15% of first-degree relatives will have the condition, although it may be asymptomatic. The concordance rate in identical twins is about 70%.

HLA-DQ2 (DQAI*0501, DQBI*0201) and HLA-DQ8 (DQAI*0301, DQBI*0302) are associated with CD. Over 90% of patients will have HLA-DQ2, compared with 20–30% of the general population. Studies in twins and siblings indicate that HLA genes are responsible for less than 50% of the genetic cause of the disease. Many unaffected people also carry these genes, so other factors must also be involved. Non-HLA genes may also contribute to coeliac disease, e.g. chromosome regions 19p13.1, 11q, 5q31-33 and 6q21-22. The CD28/CTLA4/ILO5 gene cluster has also shown linkage with coeliac disease.

Environmental factors. Breast-feeding and the age of introduction of gluten into the diet was thought to be significant but this remains controversial.

Rotavirus infection in infancy also increases the risk, and adenovirus-12 which has sequence homology with α-gliadin has been suspected as a causative agent.

Clinical features

Coeliac disease can present at any age. In infancy it sometimes appears after weaning onto gluten-containing foods. The peak period for diagnosis in adults is in the fifth decade, with a female preponderance. Many patients are asymptomatic (silent) and come to attention because of routine blood tests, e.g. a raised MCV, or iron deficiency in pregnancy. The symptoms are very variable and often non-specific, e.g. tiredness and malaise often associated with anaemia.

GI symptoms may be absent. Diarrhoea or steatorrhoea, abdominal discomfort, bloating or pain and weight loss suggest more severe disease. Mouth ulcers and angular stomatitis are frequent and can be intermittent. Infertility and neuropsychiatric symptoms of anxiety and depression occur.

Rare complications include tetany, osteomalacia, or gross malnutrition with peripheral oedema. Neurological symptoms such as paraesthesia, ataxia (due to cerebellar calcification), muscle weakness or a polyneuropathy occur; the prognosis for these symptoms is variable. There is an increased incidence of atopy and autoimmune disease, including thyroid disease, type 1 diabetes and Sjögren's syndrome. Other associated diseases include inflammatory bowel disease, primary biliary cirrhosis, chronic liver disease, interstitial lung disease and epilepsy. IgA deficiency is more common than in the general population. Long-term problems include osteoporosis which occurs even in patients on long-term gluten-free diets.

Physical signs are usually few and non-specific and are related to anaemia and malnutrition.

Diagnosis

Small bowel biopsy is the standard for diagnosis, and is essential because treatment involves a life-long diet that is both expensive and socially limiting. Because the disease is sometimes patchy and it can be difficult to orientate endoscopic biopsies for histological section, four to six forceps biopsies should be taken from the second part of the duodenum. Endoscopic signs including absence of mucosal folds, mosaic pattern of the surface and scalloping of mucosal folds are often present; however, their absence is not conclusive because they are markers of relatively severe disease. Biopsies should always be taken if CD is a possibility.

Histology (Fig. 6.30). Histological changes are of variable severity and, though characteristic, are not specific. Villous atrophy can be caused by many other conditions, but CD is the commonest cause of subtotal villous atrophy. The villous architecture is almost normal in mild cases, but there are abnormal numbers of intraepithelial lymphocytes. In severe cases there is an absence of villi, with flattening of the mucosal surface.

Histological examination shows crypt hyperplasia with chronic inflammatory cells in the lamina propria, and villous atrophy. The enterocytes become cuboidal with an increase in the number of intraepithelial lymphocytes. In the lamina propria there is an increase in lymphocytes and plasma cells. The most severe histological change with mucosal atrophy and hypoplasia is seen in patients who do not respond to a gluten-free diet.

Serology. Persistent diarrhoea, folate or iron deficiency, a family history of CD and associated autoimmune disease are indications for serological testing.

The most sensitive tests are for endomysial and anti-tissue transglutaminase antibodies. The sensitivity of these tests is >90% though both are not always positive in the same subject. Titres of either correlate with the severity of mucosal damage so they can be used for dietary monitoring. Standard tests use IgA class antibodies. Selective IgA deficiency occurs in 2.5% of CD patients but only 0.25% of normals and renders these tests falsely negative. IgG-based tests should then be used.

HLA typing. HLA-DQ2 is present in 90–95% of CD patients and HLA-DQ8 in about 8%, i.e. most of the rest. The *absence* of both alleles has a high negative predictive value for coeliac disease. HLA typing is useful for ruling out the disease, for example in patients already on a gluten free diet.

Other investigations

- **Haematology**. Mild or moderate anaemia is present in 50% of cases. Folate deficiency is common, often causing macrocytosis. B_{12} deficiency is rare. Iron deficiency due to malabsorption of iron and increased loss of desquamated cells is common. A blood film may therefore show microcytes and macrocytes as well as hypersegmented polymorphonuclear leucocytes and Howell–Jolly bodies (see p. 423) due to splenic atrophy.
- **Biochemistry**. *In severe cases*, biochemical evidence of osteomalacia may be seen (low calcium and high phosphate) and hypoalbuminaemia.
- **Radiology**. A small bowel follow-through may show dilatation of the small bowel with slow transit. Folds become thicker and in severe disease total effacement is seen. Radiology is now mainly used when a complication, e.g. lymphoma, is suspected.
- **Bone densitometry** (DXA) should be performed on all patients because of the risk of osteoporosis.

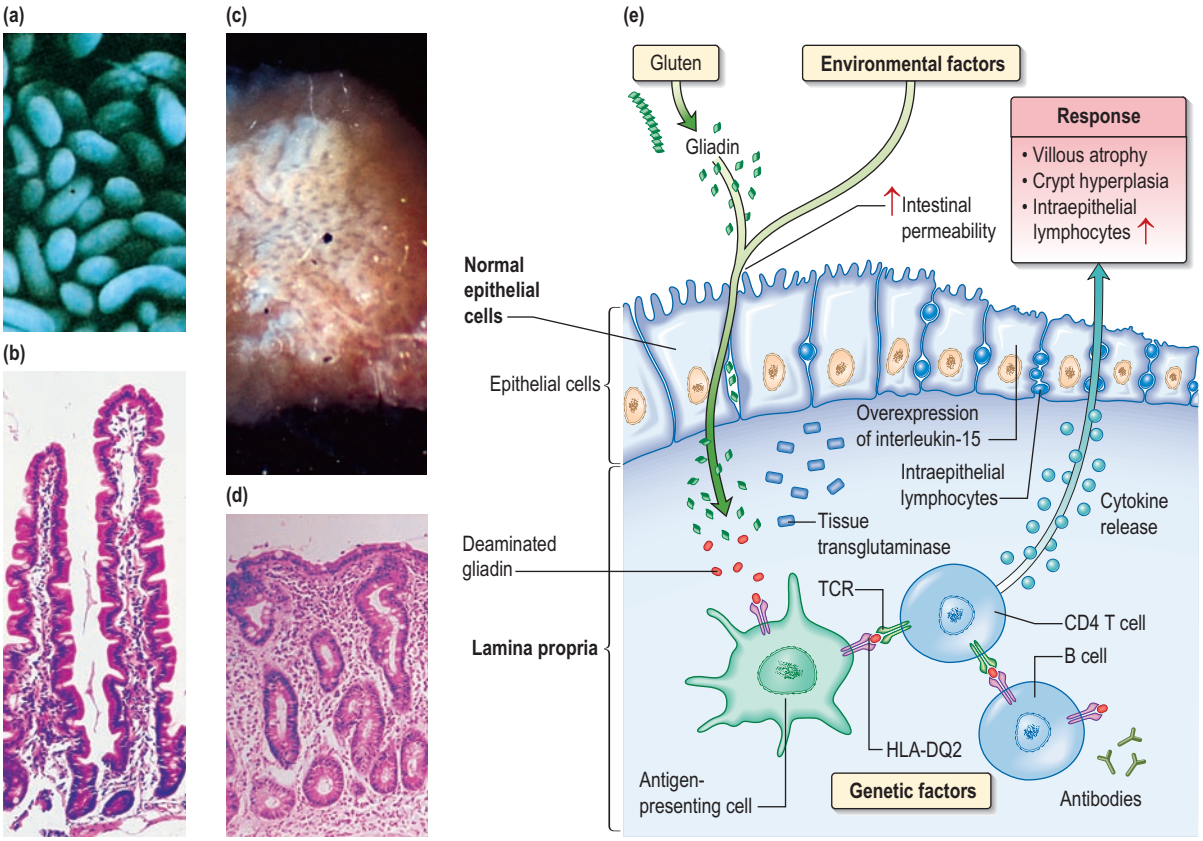


Fig. 6.30 **Small bowel mucosal appearances – macroscopic and microscopic. (a)** Normal mucosa under the dissecting microscope (DM). **(b)** Normal mucosal histology. **(c)** Coeliac disease (DM) – flattened mucosa. **(d)** Coeliac disease – showing subtotal villous atrophy. **(e) Immune and genetic mechanisms in coeliac disease.** After Green PHR, Cellier C. Coeliac disease. *New England Journal of Medicine* 2007; **357**: 1731–1743.

- **Wireless capsule endoscopy** (see p. 246) is used to look for gut abnormalities when a complication is suspected.
- **Absorption tests** are often abnormal (see p. 277) but are seldom performed because they are not crucial to the diagnosis.

Treatment and management

Replacement minerals and vitamins, e.g. iron, folic acid, calcium, vitamin D, may be needed initially to replace body stores.

Treatment is with a gluten-free diet for life. Dietary elimination of wheat, barley and rye usually produces a clinical improvement within days or weeks. Morphological improvement often takes months, especially in adults. Oats are tolerated by most coeliacs, but must not be contaminated with flour during their production. Meat, dairy products, fruits and vegetables are naturally gluten free and are all safe.

Gluten-free products can be expensive, unless subsidised by national health services. Patient support organizations such as The Coeliac Society are valuable as information sources and for advice about diet, recipes and gluten-free processed foods. Despite advice, many patients do not keep to a strict diet but maintain good health. The long-term effects of this low gluten intake are uncertain but osteoporosis can occur even in treated cases.

The usual cause for *failure to respond* to the diet is poor compliance. Dietary adherence can be monitored by serial tests for endomysial antibody (EMA) and tissue transglutaminase (tTG). If clinical progress is suboptimal then a repeat intestinal biopsy should be taken. If the diagnosis is equivocal a gluten challenge, i.e. reintroduction of gluten with evidence of jejunal morphological change, confirms the diagnosis.

Patients should have *pneumococcal vaccinations* (because of splenic atrophy) once every 5 years (see p. 423).

Complications

A few patients do not improve on a strict diet and are said to have *unresponsive coeliac disease*. Often no cause is found, but enteropathy-associated T cell lymphoma (EATCL), ulcerative jejunitis or carcinoma are sometimes responsible. The incidence of EATCL (see p. 264) and small bowel adenocarcinoma is increased in coeliac disease.

Ulcerative jejunitis presents with fever, abdominal pain, perforation and bleeding. Diagnosis for these conditions is with barium studies but laparotomy with full-thickness biopsies is often required. Steroids and immunosuppressive agents, e.g. azathioprine, are used.

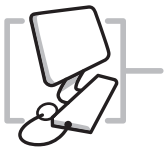
Carcinoma of the oesophagus as well as extra-gastrointestinal cancers are also increased in incidence. Malignancy seems to be unrelated to the duration of the disease but the incidence is reduced by a gluten-free diet.

Dermatitis herpetiformis (see also p. 1256)

This is an uncommon blistering subepidermal eruption of the skin associated with a gluten-sensitive enteropathy. Rarely

there may be gross malabsorption, but usually the jejunal morphological abnormalities are not as severe as in coeliac disease. The inheritance and immunological abnormalities are the same as for coeliac disease. The skin condition responds to dapsone but a gluten-free diet improves both the enteropathy and the skin lesion.

Tropical sprue

www.studentconsult.com
More online


This is a condition presenting with chronic diarrhoea and malabsorption that occurs in residents or visitors to affected tropical areas. The disease is endemic in most of Asia, some Caribbean islands, Puerto Rico and parts of South America. Epidemics occur, lasting up to 2 years, and in some areas repeated epidemics occur at varying intervals of up to 10 years.

The term tropical sprue is reserved for severe malabsorption (of two or more substances) accompanied by diarrhoea and malnutrition. Malabsorption of a mild degree, sometimes following an enteric infection, is quite common in the tropics; it is usually asymptomatic and is sometimes called tropical malabsorption.

Aetiology

The aetiology is unknown, but is likely to be infective because the disease occurs in epidemics and patients improve on antibiotics. A number of agents have been suggested but none has been unequivocally shown to be responsible. Different agents could be involved in different parts of the world.

Clinical features

These vary in intensity and consist of diarrhoea, anorexia, abdominal distension and weight loss. The onset is sometimes acute and occurs either a few days or many years after being in the tropics. Epidemics can break out in villages, affecting thousands of people at the same time. The onset can also be insidious, with chronic diarrhoea and evidence of nutritional deficiency. The clinical features of tropical sprue vary in different parts of the world, particularly as different criteria are used for diagnosis.

Diagnosis

Acute infective causes of diarrhoea must be excluded (see p. 307), particularly *Giardia*, which can produce a syndrome very similar to tropical sprue. Malabsorption should be demonstrated, particularly of fat and B_{12}. The jejunal mucosa is abnormal, showing some villous atrophy (partial villous atrophy). In most cases the lesion is less severe than that found in coeliac disease, although it affects the whole small bowel. Mild mucosal changes can be seen in asymptomatic individuals in the tropics.

Treatment and prognosis

Many patients improve when they leave the sprue area and take folic acid (5 mg daily). Most patients also require an antibiotic to ensure a complete recovery (usually tetracycline 1 g daily for up to 6 months).

Severely ill patients require resuscitation with fluids and electrolytes for dehydration, and nutritional deficiencies should be corrected. Vitamin B_{12} (1000 μg) is also given to all acute cases.

The prognosis is excellent. Mortality is usually associated with water and electrolyte depletion, particularly in epidemics.

Bacterial overgrowth

The gut contains many resident bacteria in the terminal ileum and colon. Anaerobic bacteria, e.g. *Bacteroides*, bifidobacteria, are 100–1000 times more abundant than aerobic (facultative anaerobes), e.g. *Escherichia*, *Enterobacter*, *Enterococcus*. This gut microflora has major functions including metabolic, e.g. fermentation of non-digestible dietary residues into short chain fatty acids as an energy source in the colon.

The microflora which influences epithelial cell proliferation, is involved in the development and maintenance of the immune system and protects the gut mucosa from colonization by pathogenic bacteria. Bacteria also initiate vitamin K production.

The upper part of the small intestine is almost sterile, containing only a few organisms derived from the mouth. Gastric acid kills some ingested organisms and intestinal motility keeps bacterial counts in the jejunum low. The normal terminal ileum contains faecal-type organisms, mainly *Escherichia coli* and anaerobes, and the colon has abundant bacteria.

Bacterial overgrowth is normally found associated with a structural abnormality of the small intestine such as a stricture or diverticulum, although it can occur occasionally in the elderly without. *E. coli* and/or *Bacteroides*, both in concentrations greater than 10^6/mL, are found as part of a mixed flora. These bacteria are capable of deconjugating and dehydroxylating bile salts, so that unconjugated and dehydroxylated bile salts can be detected in small bowel aspirates.

The *clinical features* of overgrowth are chiefly diarrhoea and steatorrhoea. There may also be symptoms due to the underlying small bowel pathology. Steatorrhoea (see p. 275) occurs because of conjugated bile salt deficiency. Some bacteria can metabolize vitamin B_{12} and interfere with its binding to intrinsic factor, leading to mild B_{12} deficiency (see p. 399) rarely severe enough to produce a neurological deficit. Some bacteria produce folic acid giving a high serum folate. Bacterial overgrowth has only minimal effects on the absorption of other substances. Confirmation of bacterial overgrowth is with the hydrogen breath test (see p. 277); aspiration studies are not routinely performed.

Treatment

If possible, the underlying lesion should be corrected (e.g. a stricture should be resected). Where this is not possible, rotating courses of antibiotics are necessary, such as metronidazole, a tetracycline or ciprofloxacin. The response to antibiotics is unpredictable.

Intestinal resection

Small intestinal resection is usually well tolerated, but massive resection leaving less than 1 m of small bowel in continuity is followed by the *short-bowel syndrome*. The effects of resection depend on the amount and location of the resection and the presence or absence of the colon. Resection of the jejunum is better tolerated than ileal resection, where there is less adaptation, probably due to low levels of glucagon-like peptide 2 (GLP-2), which is a specific growth hormone for the enterocyte.

Ileal resection

The ileum is the site of specific mechanisms for the absorption of bile salts and vitamin B_{12}. Relatively small resections

lead to malabsorption of these substances. Removal of the ileocaecal valve increases the incidence of diarrhoea (Fig. 6.31).

The following occur after ileal resection:

- *Bile-salt induced diarrhoea*: bile salts and fatty acids enter the colon and cause malabsorption of water and electrolytes (see p. 307).
- *Steatorrhoea and gallstone formation*: increased bile salt synthesis can compensate for loss of approximately one-third of the bile salts in the faeces. Greater loss than this results in decreased micelle formation and steatorrhoea, and lithogenic bile and gallstone formation.
- *Oxaluria and oxalate stones*: bile salts in the colon cause increased oxalate absorption with oxaluria leading to urinary stone formation.
- *B_{12} deficiency*: low serum B_{12}, macrocytosis and other effects of B_{12} deficiency.

Investigations include a small bowel follow-through, measurement of B_{12}, bile salt retention (SeHCAT test) and occasionally fat absorption (see p. 277). A hydrogen breath test may show rapid transit (see p. 277). Many patients require B_{12} replacement and some need a low-fat diet if there is steatorrhoea. Diarrhoea is often improved by colestyramine which binds bile salts and reduces the level of diarrhoeogenic bile salts in the colon.

Jejunal resection

The ileum can compensate for loss of jejunal absorptive function. Jejunal resection may lead to gastric hypersecretion with high gastrin levels; the exact mechanism of this is unclear. Structural and functional intestinal adaptation take place over the course of a year, with an increase in the absorption per unit length of bowel.

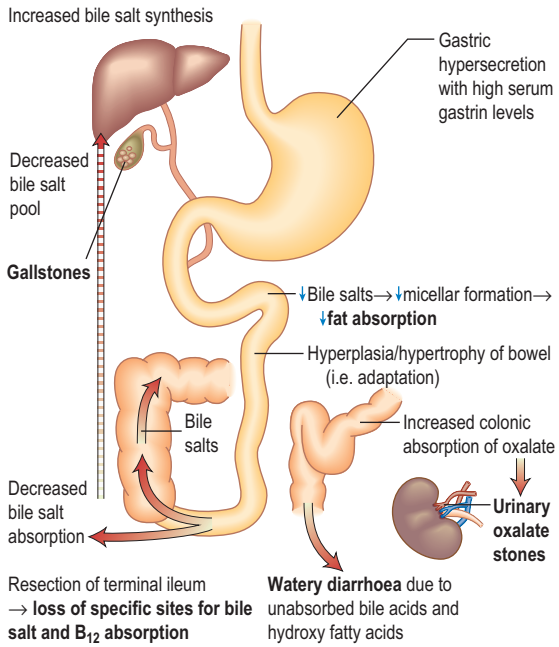


Fig. 6.31 The effects of resection of distal small bowel.

Massive intestinal resection (short-bowel syndrome)

This most often occurs following resection for Crohn's disease, mesenteric vessel occlusion (see p. 283), radiation enteritis (see p. 282) or trauma. There are two common situations:

Shortened small intestine ending at a terminal small bowel stoma

The major problem is of sodium and fluid depletion and the majority of patients with 100 cm or less of jejunum remaining will require parenteral supplements of fluid and electrolytes, often with nutrients. Sodium losses can be minimized by increasing salt intake, restricting clear fluids between meals and administering oral glucose–electrolyte mixture with a sodium concentration 90 mmol/L. Jejunal transit time can be increased and stomal effluent loss of fluids and electrolytes reduced by treatment with the somatostatin analogue octreotide, and to a much lesser extent, with loperamide, codeine phosphate or co-phenotrope. There is no benefit from a low-fat diet, but fat assimilation can be increased on treatment with colestyramine and synthetic bile acids.

Shortened small intestine in continuity with colon

Only a small proportion of these patients require parenteral supplementation of fluid, electrolytes and nutrients because of the absorptive capacity of the colon for fluid and electrolytes. Unabsorbed fat results in impairment of colonic fluid and electrolyte absorption so patients should be on a low-fat diet. A high carbohydrate intake is advised as unabsorbed carbohydrate is metabolized anaerobically to short-chain fatty acids (SCFAs) which are absorbed; they also stimulate fluid and electrolyte absorption in the colon and act as an energy source (1.6 kcal/g). Patients are often treated with colestyramine to reduce diarrhoea and colonic oxalate absorption.

Whipple's disease

Whipple's disease is a rare infectious bacterial disease caused by *Tropheryma whipplei*. About 1000 cases have been described; 87% are males, usually white and middle-aged. It usually presents with arthritis and arthralgia, progressing over years to weight loss and diarrhoea with abdominal pain, systemic symptoms of fever and weight loss. Peripheral lymphadenopathy and involvement of the heart, lung, joints and brain may occur, simulating many neurological conditions. 15% of patients have atypical symptoms.

Blood tests may show features of chronic inflammation and malabsorption. Endoscopy typically shows pale, shaggy duodenal mucosa with eroded, red, friable patches.

Diagnosis is made by small bowel biopsy. Periodic acid–Schiff (PAS)-positive macrophages are present but are non-specific. On electron microscopy, the characteristic trilaminar cell wall of *T. whipplei* can be seen within macrophages. *T. whipplei* antibodies can be identified by immunohistochemistry. A confirmatory PCR-based assay is available.

Treatment is with antibiotics which cross the blood–brain barrier, such as 160 mg trimethoprim and 800 mg sulphamethoxazole (co-trimoxazole) daily for 1 year. This is often preceded by a 2-week course of streptomycin and penicillin

or ceftriaxone. Treatment periods of less than a year are associated with relapse in about 40%.

Radiation enteritis

Radiation of more than 40 Gy will damage the intestine. The chronic effects of radiation are muscle fibre atrophy, ulcerative changes due to ischaemia, and obstruction due to radiation-induced fibrotic strictures.

Pelvic irradiation is frequently used for gynaecological and urinary tract malignancies, so the ileum and rectum are the areas most often involved.

At the time of the irradiation, there may be nausea, vomiting, diarrhoea and abdominal pain, usually improving within 6 weeks of completion of therapy.

Chronic radiation enteritis is diagnosed if symptoms persist for 3 months or more. The prevalence is more than 15%. Abdominal pain due to obstruction is the main symptom. Malabsorption can be due to bacterial overgrowth in dilated segments and mucosal damage.

Many patients suffer from increased bowel frequency.

Treatment is symptomatic, although often unsuccessful in chronic radiation enteritis. Surgery should be avoided if possible, being reserved for obstruction or perforation.

Acute radiation damage to the rectum produces a *radiation proctitis* with diarrhoea and tenesmus, with or without blood. Local steroids sometimes help initially. When the acute phase heals, mucosal telangiectases form and may cause persistent bleeding. They can be treated with argon plasma coagulation or, under a light anaesthetic, by packing the rectum with a formalin-soaked swab for 2 minutes, both of which destroy the telangiectases.

Parasite infestation

Giardia intestinalis (see p. 161) not only produces diarrhoea but can produce malabsorption with steatorrhoea. Minor changes are seen in the jejunal mucosa and the organism can be found in the jejunal fluid or mucosa.

Cryptosporidiosis (see p. 162) can also produce malabsorption.

Patients with HIV infection are particularly prone to parasitic infestation (see Table 6.24).

Other causes of malabsorption

- *Drugs that bind bile salts* (e.g. colestyramine) and some antibiotics (e.g. neomycin) produce steatorrhoea.
- *Orlistat* (see p. 232) is used in obesity to reduce fat absorption by inhibiting gastric and pancreatic lipase causing diarrhoea and steatorrhoea.
- *Thyrotoxicosis*: diarrhoea, rarely with steatorrhoea, occurs in thyrotoxicosis owing to increased gastric emptying and increased motility.
- *Zollinger–Ellison syndrome*.
- *Intestinal lymphangiectasia* produces diarrhoea and rarely steatorrhoea (see p. 283).
- *Lymphoma* that has infiltrated the small bowel mucosa causes malabsorption.
- *Diabetes mellitus*: diarrhoea, malabsorption and steatorrhoea may occur, sometimes due to bacterial overgrowth from autonomic neuropathy causing small bowel stasis.
- *Hypogammaglobulinaemia*, which is seen in a number of conditions including lymphoid nodular hyperplasia, causes steatorrhoea due either to an abnormal jejunal mucosa or to secondary infestation with *Giardia intestinalis*.

MISCELLANEOUS INTESTINAL DISEASES

Protein-losing enteropathy

Protein-losing enteropathy refers to intestinal conditions causing protein loss, usually manifest by hypoalbuminaemia. The causes are protean, including Crohn's disease, tumours, Ménétrier's disease, coeliac disease and lymphatic disorders (e.g. lymphangiectasia).

Usually protein-losing enteropathy forms a minor part of the generalized disorder, but occasionally hepatic synthesis of albumin cannot compensate for the protein loss, and peripheral oedema dominates the clinical picture. The investigations are described on page 277 and treatment is that of the underlying disorder.

Meckel's diverticulum

This is the most common congenital abnormality of the GI tract, affecting 2–3% of the population. The diverticulum projects from the wall of the ileum approximately 60 cm from the ileocaecal valve. It is usually symptomless, but 50% contain gastric mucosa that secretes hydrochloric acid. Peptic ulcers can occur and may bleed (see p. 267) or perforate.

Acute inflammation of the diverticulum also occurs and is indistinguishable clinically from acute appendicitis. Obstruction from an associated band rarely occurs.

Treatment is surgical removal, often laparoscopically.

Tuberculosis (see also p. 863)

Tuberculosis (TB) can affect the intestine as well as the peritoneum (see p. 318). In developed countries, most patients are from ethnic minority groups, or are immunocompromised due to HIV or drugs. Intestinal tuberculosis is due to reactivation of primary disease caused by *Mycobacterium tuberculosis*. Bovine TB occurs in areas where milk is unpasteurized and is rare in western countries.

Clinical features are abdominal pain, weight loss, anaemia, fever with night sweats, obstruction, right iliac fossa pain or a palpable mass. The ileocaecal area is most commonly affected, but the colon, and rarely other parts of the gastrointestinal tract, can be involved. One-third of patients present acutely with intestinal obstruction or generalized peritonitis. 50% have X-ray evidence of pulmonary tuberculosis.

Diagnosis

TB must be differentiated from Crohn's disease and should always be a differential diagnosis in Asian immigrants in developed countries. A caecal carcinoma can present with similar symptoms.

- A small bowel follow-through will show transverse ulceration, diffuse narrowing of the bowel with shortening of the caecal pole.

- Ultrasound or CT shows additional mesenteric thickening and lymph node enlargement.
- Histology and culture of tissue is desirable, but it is not always possible. Specimens can be obtained by colonoscopy or laparoscopy but laparotomy is required in some cases.

Treatment

Drug treatment is similar to that for pulmonary TB – rifampicin, isoniazid and pyrazinamide (see p. 866) – but treatment should last 1 year. Treatment should be started if there is a high degree of suspicion.

Amyloid (see also p. 1078)

In systemic amyloidosis there is usually diffuse involvement that may affect any part of the GI tract. Occasionally amyloid deposits occur as polypoid lesions. The symptoms depend on the site of involvement; amyloidosis in the small intestine gives rise to diarrhoea.

Rheumatic autoimmune disorders

Systemic sclerosis (see p. 545) most commonly affects the oesophagus (see p. 250), although the small bowel and colon are often found to be involved if the appropriate radiological studies are performed. Frequently there are no symptoms of this involvement, but diarrhoea and steatorrhoea can occur. This is usually due to bacterial overgrowth of the small bowel as a result of reduced motility, dilatation and the presence of diverticula.

In *rheumatoid arthritis* (see p. 523) and systemic lupus erythematosus (see p. 541), gastrointestinal symptoms may occur, but rarely predominate.

Intestinal ischaemia

Intestinal ischaemia results from occlusion of arterial inflow, occlusion of venous outflow or failure of perfusion; these factors may act alone or in combination and usually occur in the elderly. Ischaemia can be due to:

- **Arterial inflow occlusion:**
  - atheroma
  - thrombosis
  - embolus (cardiac arrhythmia) including cholesterol emboli (see p. 609)
  - aortic disease (occluding ostia of mesenteric vessels)
  - vasculitis (see p. 548), thromboangiitis and Takayasu's syndrome (see p. 809)
  - neoplasia (occlusion of vessels – rare).
- **Venous outflow occlusion.** Occurs in 5–15% of cases and usually in sick patients with circulatory failure.
- **Infarction without occlusion.** Approximately one-third of patients dying with acute ischaemic necrosis of the small intestine have no demonstrable occlusion of a major vessel. Reduced cardiac output, hypotension and shock are the main causes of reduced intestinal blood flow leading to non-occlusive infarction.

Acute small intestinal ischaemia

Patients present with sudden abdominal pain and vomiting. An embolus from the heart in a patient with atrial fibrillation is the commonest cause, usually occluding the superior mesenteric artery. The abdomen is usually distended and tender, and bowel sounds are absent. The patient is hypotensive and ill. Surgery is necessary to resect the gangrenous bowel. Mortality is high (up to 90%) and is related to coexisting disease, the development of multiorgan failure (MOF) (see p. 909) and massive fluid and electrolyte losses in the postoperative period. Survivors may go on to develop nutritionally inadequate short-bowel syndrome (see p. 281).

Ischaemic colitis

See page 297.

Chronic small intestinal ischaemia

This is due to atheromatous occlusion or cholesterol emboli of the mesenteric vessels, particularly in the elderly. Such an occlusion does not always produce clinical effects because of the collateral circulation. The characteristic symptom is abdominal pain occurring after food. This may be followed by acute mesenteric vascular occlusion. Loud bruits may be heard but, as these are heard in normal subjects, they are of doubtful significance. The diagnosis is made using angiography. The term 'coeliac axis compression syndrome' has been used in young patients with chronic abdominal pain, bruits and minor angiographic changes. It is not an organic syndrome, but results from the false correlation of pain and bruits.

Eosinophilic gastroenteritis

In this condition of unknown aetiology there is eosinophilic infiltration and oedema of any part of the gastrointestinal mucosa. The gastric antrum and proximal small intestine are usually involved either as a localized lesion (eosinophilic granuloma) or diffusely with sheets of eosinophils seen in the serosal and submucosal layers. There is an association with asthma, eczema and urticaria.

The condition occurs mainly in the third decade. The clinical presentation depends on the site of gut involvement. Abdominal pain, nausea and vomiting and upper GI bleeding occur. Peripheral eosinophilia occurs in only 20% of patients. Endoscopic biopsy is useful for making the diagnosis histologically. Radiology may demonstrate mass lesions.

Treatment is with steroids for the widespread infiltration, particularly if peripheral eosinophilia is present.

In some adults the condition appears to be allergic (allergic gastroenteritis) and is associated with peripheral eosinophilia and high levels of plasma and tissue IgE. Eosinophilic oesophagitis is described on page 255, but its relationship to eosinophilic gastritis is unclear.

Intestinal lymphangiectasia

Dilatation of the lymphatics may be primary or secondary to lymphatic obstruction, such as occurs in malignancy or constrictive pericarditis. Hypoproteinaemia with ankle oedema is the main feature. The rare primary form may be detected incidentally as dilated lacteals on a jejunal biopsy or it can produce steatorrhoea of varying degrees. Serum immunoglobulin levels are reduced, with low circulating lymphocytes. Treatment is with a low-fat diet, MCT and fat-soluble vitamin supplements as required. Octreotide has a dramatic effect in a few primary cases, although the mechanism of action is unknown.

Abetalipoproteinaemia

This is rare and is due to a failure of apo B-100 synthesis in the liver and apo B-48 in the intestinal cell, so that chylomicrons are not formed. This leads to fat accumulation in the intestinal cells, giving a characteristic histological appearance to the jejunal mucosa. Clinical features include acanthocytosis (spiky red cells owing to membrane abnormalities), a form of retinitis pigmentosa, and mental and neurological abnormalities. The latter can be prevented by vitamin E injections.

TUMOURS OF THE SMALL INTESTINE

The small intestine is relatively resistant to the development of neoplasia and only 3–6% of all GI tumours and fewer than 1% of all malignant lesions occur here. The reason for the rarity of tumours is unknown. Explanations include the fluidity and relative sterility of small bowel contents and the rapid transit time, reducing the time of exposure to potential carcinogens. It is also possible that the high population of lymphoid tissue and secretion of IgA in the small intestine protect against malignancy.

Adenocarcinoma of the small intestine is rare and found most frequently in the duodenum (in the periampullary region) and in the jejunum. It is the most common tumour of the small intestine, accounting for up to 50% of primary tumours.

Lymphomas are most frequently found in the ileum. These are of the non-Hodgkin's type and must be distinguished from peripheral or nodal lymphomas involving the gut secondarily.

In developed countries, the most common type of lymphoma is the B cell type arising from MALT (see p. 262). These lymphomas tend to be annular or polypoid masses in the distal or terminal ileum, whereas most T cell lymphomas are ulcerated plaques or strictures in the proximal small bowel.

A tumour similar to Burkitt's lymphoma also occurs and commonly affects the terminal ileum of children in North Africa and the Middle East.

Predisposing factors for adenocarcinoma and lymphoma

Coeliac disease

There is an increased incidence of lymphoma of the T cell type and adenocarcinoma of the small bowel, as well as an unexplained increase in all malignancies both in the GI tract and elsewhere. The reason for the local development of malignancy is unknown. It is now accepted that coeliac disease is a premalignant condition, but there is no association with the length of the symptoms. Treatment with a gluten-free diet reduces the risk of both lymphoma and carcinoma.

Crohn's disease

There is a small increase in the incidence of adenocarcinoma of the small bowel in Crohn's disease.

Immunoproliferative small intestinal disease (IPSID)

IPSID is a B cell disorder in which there is proliferation of plasma cells in the lamina propria of the upper small bowel producing truncated monoclonal heavy chains, without associated light chains. The α heavy chains are found in the gut mucosa on immunofluorescence and can also be detected in the serum. It occurs usually in countries surrounding the Mediterranean, but it has also been found in other developing countries in South America and the Far East. IPSID predominantly affects people in lower socio-economic groups in areas with poor hygiene and a high incidence of bacterial and parasitic infection of the gut. IPSID presents as a malabsorptive syndrome associated with diffuse lymphoid infiltration of the small bowel and neighbouring lymph nodes, progressing in some cases to a lymphoma. The condition has also been documented in the developed world.

Clinically, patients present with abdominal pain, diarrhoea, anorexia, weight loss and symptoms of anaemia. There may be a palpable mass, and a small bowel follow-through may detect a mass lesion. Endoscopic biopsy is useful where lesions are within reach. Ultrasound and CT may show bowel wall thickening and the involvement of lymph nodes, which is common with lymphoma. Wireless capsule endoscopy can be used where obstruction by the capsule is not likely, but cannot deliver histology.

Treatment of small intestinal tumours

Adenocarcinoma. Most patients are treated surgically with a segmental resection. The overall 5-year survival rate is 20–35%; this varies with the histological grade and the presence or absence of lymph node involvement. Radiotherapy and chemotherapy are used in addition.

IPSID. If there is no evidence of lymphoma, antibiotics, e.g. tetracycline, should be tried initially. In the presence of lymphoma, combination chemotherapy is used; in one series the 3- to 5-year survival rate was 58%.

Lymphoma. Most patients require surgery and radiotherapy with chemotherapy for more extensive disease. The prognosis varies with the type. The 5-year survival rate for T cell lymphomas is 25%, but is better for B cell lymphomas, varying from 50% to 75%, depending on the grade of lymphoma.

Carcinoid tumours

These originate from the enterochromaffin cells (APUD cells) of the intestine. They make up 10% of all small bowel neoplasms, the most common sites being in the appendix, terminal ileum and the rectum. It is often difficult to be certain histologically whether a particular tumour is benign or malignant. Ten per cent of carcinoid tumours in the appendix present as acute appendicitis, secondary to obstruction. Surgical resection of the tumour is usually performed.

Most carcinoids do not secrete hormones or vasoactive compounds, and may present with liver enlargement due to metastases.

Carcinoid syndrome occurs in only 5% of patients with carcinoid tumours and only when there are liver metastases. Patients complain of spontaneous or induced bluish-red flushing, predominantly on the face and neck, sometimes leading to permanent changes with telangiectases.

Gastrointestinal symptoms consist of abdominal pain and recurrent watery diarrhoea. Cardiac abnormalities are found in 50% of patients and consist of pulmonary stenosis or tricuspid incompetence. Examination of the abdomen reveals

hepatomegaly. The tumours secrete a variety of biologically active amines and peptides, including serotonin (5-hydroxytryptamine – 5-HT), bradykinin, histamine, tachykinins and prostaglandins. The diarrhoea and cardiac complications are probably caused by 5-HT itself, but the cutaneous flushing is thought to be produced by one of the kinins, such as bradykinin. This is known to cause vasodilatation, bronchospasm and increased intestinal motility.

Diagnosis and treatment

- *Ultrasound examination* confirms the presence of liver secondary deposits.
- *Urine* shows a high concentration of 5-hydroxyindoleacetic acid (5-HIAA) which is the major metabolite of 5-HT.

Treatment is with octreotide and lanreotide; both are octapeptide somatostatin analogues that inhibit the release of many gut hormones. They alleviate the flushing and diarrhoea and can control a carcinoid crisis. Octreotide is given subcutaneously in doses up to 200 µg three times daily initially; a depot preparation 30 mg every 4 weeks can then be used. Lanreotide 30 mg is given every 7–10 days, or as a gel 60 mg every 28 days. Long-acting octreotide also sometimes inhibits tumour growth and, since its introduction, other therapy has become usually unnecessary. Interferon and other chemotherapeutic regimens occasionally reduce tumour growth, but have not been shown to increase survival.

Most patients survive for 5–10 years after diagnosis.

Peutz–Jeghers syndrome

This consists of mucocutaneous pigmentation (circumoral (95% of patients), hands (70%) and feet (60%)) and gastrointestinal polyps. It has an autosomal dominant inheritance. The gene *STK11* (also known as *LKB1*) responsible for Peutz–Jeghers codes for a serine protein kinase and can be used for genetic analysis. The brown buccal pigment is characteristic of the condition. The polyps, which are hamartomas, can occur anywhere in the GI tract but are most frequent in the small bowel. They may bleed or cause small bowel obstruction or intussusception (50% of patients).

Treatment is by endoscopic polypectomy. Balloon enteroscopy may be necessary to reach all the small bowel polyps. Bowel resection should be avoided if possible, but may be necessary in patients presenting with gangrenous bowel due to intussusception. Follow-up is with yearly pan-endoscopy. There is an increased incidence of GI cancers. Non-GI cancers also occur with increased frequency, so yearly screening for uterine, ovarian and cervical cancer should start in the teens, and breast and testicular screening by the age of 20.

Other tumours

Adenomas, lipomas and stromal tumours (see p. 266) are rarely found and are usually asymptomatic and picked up incidentally. They occasionally present with iron deficiency anaemia. In familial adenomatous polyposis (FAP) (p. 301) duodenal adenomas form in one-third of patients and may progress to adenocarcinoma. This is the commonest cause of death in FAP patients who have been treated by prophylactic coloectomy.

FURTHER READING

Craig D, Robbins G, Howdle PD. Advances in coeliac disease. *Current Opinions in Gastroenterology* 2007; **23**: 142–148.

Fenollar F, Puéchal X, Raoult D. Whipple's disease (review). *New England Journal of Medicine* 2007; **356**: 55–66.

Green PHR, Cellier C. Coeliac disease. *New England Journal of Medicine* 2007; **357**: 1731–1743.

Scott EM, Gaywood B, Scott BB for the British Society for Gastroenterology. Guidelines for osteoporosis in coeliac disease and inflammatory bowel disease. *Gut* 2000; **46**(suppl I): i1–i8.

INFLAMMATORY BOWEL DISEASE (IBD)

Two major forms of *non-specific* inflammatory bowel disease are recognized: Crohn's disease (CD), which can affect any part of the GI tract, and ulcerative colitis (UC), which affects only the large bowel.

There is overlap between these two conditions in their clinical features, histological and radiological abnormalities; in 10% of cases of colitis a definitive diagnosis of either UC or CD is not possible. Currently it is necessary to distinguish between these two conditions because of certain differences in their management. It is possible that these conditions represent two aspects of the same disease.

Three additional forms of inflammatory bowel disease are also recognized, namely microscopic ulcerative, microscopic lymphocytic and microscopic collagenous colitis (see p. 294).

Epidemiology

The incidence of CD varies from country to country but is approximately 4–10 per 100 000 annually, with a prevalence of 27–106 per 100 000. The incidence of UC is stable at 6–15 per 100 000 annually, with a prevalence of 80–150 per 100 000.

Although both conditions have a world-wide distribution, the highest incidence rates and prevalence have been reported from northern Europe, the UK and North America. Both race and ethnic origin affect the incidence and prevalence of CD and UC. Thus in North America prevalence rates of CD are lower in Hispanic and Asian people (4.1, 5.6 per 100 000 respectively) compared to white individuals (43.6 per 100 000). Jewish people are more prone to inflammatory bowel disease than any other ethnic group. Prevalence rates also change after migration (e.g. higher rates in Chinese people living in Hong Kong compared with mainland China).

Crohn's disease is slightly commoner in females (M : F = 1 : 1.2) and occurs at a younger age (mean 26 years) than UC (M : F = 1.2 : 1; mean 34 years).

Aetiopathogenesis

Although the aetiology of IBD is unknown, it is becoming clear that IBD represents the outcome of three essential interactive co-factors: genetic susceptibility, the environment and host immune response (Fig. 6.32), with the environmental factors representing both the local microenvironment (enteric microflora) and also the nutritional environment.

- *Familial*. A positive family history is the largest independent risk factor for the disease. Thus up to 1 in 5 patients with CD and 1 in 6 patients with UC will have a first-degree relative with the disease.
- *Genetic*. There is increased concordance for the disease (CD more than UC) in monozygotic twins in comparison with dizygotic twins who demonstrated familial aggregation. UC and CD are polygenic diseases. There is no single locus but susceptibility loci have been identified on chromosomes 16 (IBD1), 12, 6, 14, 5, 19, 1, 16 (IBD8) and 3, and these have been renamed IBD1–9 respectively. Significant loci have also been found on chromosome 13q and in Jewish families on 1p and 3q. Linkage mutations have been found on Card 15 (NOD 2), the underlying gene on chromosome 16 (IBD1), and also genes underlying the IBD5 and IBD3 loci. The Card 15 (NOD 2) gene is associated with CD in white populations, and has been associated with stricturing small bowel CD. Recent SNP (single nucleotide polymorphism) scans have identified a locus for UC and CD at ECM1 (extracellular matrix protein 1). Several other risk loci for both diseases have been found on IL23R, IL12B, NKX2-3 and MST1. The autophagy genes *ATG16L1* and *IRGM*, which control an intracellular degradation process (along with *CARD15*), are specific for CD.

  Apart from susceptibility, HLA genes on chromosome 6 also appear to have a role in modifying the disease. Thus the *DRB*0103* allele, which is uncommon, is linked to a particularly aggressive course of UC and the need for surgery, as well as with colonic CD. *DRB*0103* and *MICA*010* are associated with perianal disease and *DRB*0701* with ileal CD. For the extraintestinal disease complications and HLA links see page 287.
- *Environmental factors*. Good domestic hygiene has been shown to be a risk factor for CD but not for UC. Thus poor and large families living in crowded conditions with no tap or hot water and consuming contaminated food have a lower risk of developing CD.

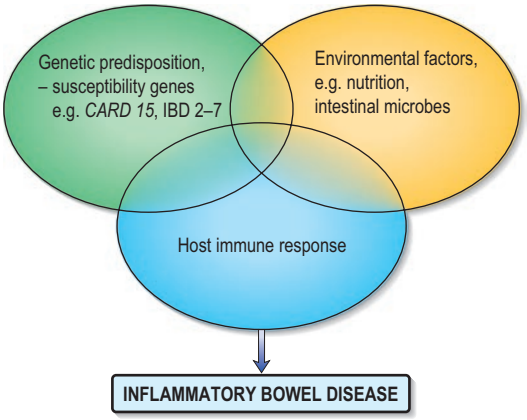


Fig. 6.32 **Inflammatory bowel disease.** Schematic diagram showing the aetiopathogenesis. Modified from Cashman KD, Shanahan F. Is nutrition an aetiological factor for inflammatory bowel disease? *European Journal of Gastroenterology and Hepatology* 2003; **15**: 607–613, with permission.

H. pylori sero-prevalence is high in developing countries but low in patients with CD. A 'clean' environment may not expose the intestinal immune system to pathogenic or non-pathogenic microorganisms, particularly helminthic parasites, and therefore be 'untrained' to confront minor infections. Helminth infections are associated with a type 2 helper T cell response (Th2), which would counterbalance the type-1 helper T cell response (Th1) that is characteristic of CD. If such a mechanism is operative, it would explain why there is a frequent association of a recent intestinal infection with the first presentation and subsequent flare-ups of CD.

- *Lifestyle*. Breast feeding may provide protection against inflammatory bowel disease developing in offspring. Breast feeding per se does not contribute to disease exacerbation.
  - *Nutritional factors*. Many foods and food components have been suggested to play a role in the aetiopathogenesis of IBD (e.g. high sugar and fat intake) but unfortunately the results of numerous studies have been equivocal.
  - *Smoking*. Patients with CD are more likely to be smokers, and smoking has been shown to exacerbate CD. By contrast, there is an increased risk of UC in non- or ex-smokers and nicotine has been shown to be an effective treatment of UC.
  - *Adverse life events and psychological factors* such as chronic stress and depression seem to increase relapses in patients with quiescent disease.
- *Appendicectomy* is 'protective' for the development of UC, particularly if performed for appendicitis or for mesenteric lymphadenitis before the age of 20. It also influences the clinical course of UC, with a lower incidence of colectomy and need for immunosuppressive therapy. By contrast, appendicectomy may increase the risk of development of CD and may result in more aggressive disease.
- *Intestinal microflora*. The gut is colonized by 10 times more bacterial organisms than there are host cells, there being 300–400 distinct bacterial species. Evidence supports a hypothesis that IBD is characterized by an overaggressive immune response to luminal bacterial antigens and other products, occurring against a background of genetic susceptibility (see above).
  - *Bacterial flora*. There is an alteration in the bacterial flora, with an increase in anaerobic bacteria in CD and an increase in aerobic bacteria in UC.
  - *Bacterial antigens*. Bacteria may exert their pro-inflammatory influence by producing toll-like receptor ligands such as peptidoglycan-polysaccharides (PG-PS), lipopolysaccharides (LPS), which interact in the normal intestine with surface toll-like receptors (TLR). The disruption in TLR signalling could prevent the mucosa withstanding bacterial insult.
  - *Intestinal mucosal invasion*. The intestinal wall in IBD patients is contaminated by adherent and invading bacteria. Recently it has been shown that there is increased *E. coli* adherence to the ileal-epithelial cells in CD. This occurs via *E. coli*'s type 1 pilli to a protein called carcinoma embryonic antigen-related cell adhesion molecule 6 (CEACAM 6). The latter may become a marker for inflammation.
  - *Defective chemical barrier or intestinal defensins* (see p. 274). Evidence suggests a decrease in human α

www.studentconsult.com
Ask the authors

defensin-1 (HD-1) in both CD and UC and lack of induction of HD-2 and HD-3, HD-5 in CD.
  - *Impaired mucosal barrier function* may explain the presence of unusual and potentially pathogenic bacteria, e.g. *Mycobacterium paratuberculosis* (MAP), *Listeria* and mucosal adherent *E. coli*. Their presence does not necessarily imply causation of the disease. However, MAP has recently been found in the blood of patients with CD and further studies are awaited.
  - *Butyrate*. Sulphate-producing bacteria increase luminal levels of hydrogen sulphide (H_2S), which leads to a reduction of butyrate oxidation in colonic mucosa, producing an energy-deficient state and leading to mucosal inflammation. H_2S and methane-ethiol may produce the malodorous flatus that some patients complain of prior to a flare-up.
- *Pathogenesis*. IBD results from a defective mucosal immune system producing an inappropriate response to luminal antigens, such as bacteria which enter the intestine via a leaky epithelium. Bacterial ligands interact with toll-like receptors expressed on the epithelial cell membrane via co-stimulatory molecules which enable the epithelial cells to act as antigen presenting cells to myeloid dendritic cells. There then follows a stimulation of naïve Th 0 cells into effector T cells (Th1 (IL-12, INFδ), Th2 (IL-5) and Th17 (IL-17)), which predominate over the regulatory T cells (T reg).

  The pro-inflammatory cytokines released by these activated T cells stimulate macrophages to secret tumour necrosis factor α (TNF-α), IL-1 and IL-6 in large quantities.

  Macrophages are also stimulated by plasmacytoid dendritic cells via natural killer cells which themselves can cause direct cytotoxic effects on cells and secrete inflammatory cytokines.

  These mechanisms all result in leucocytes leaving the circulation to enter the tissue and release chemokines (lymphokines, arachidonic acid metabolites, neuropeptides and free oxygen radicals), all of which lead to tissue damage and also attract more inflammatory cells like a vicious circle.

Pathology

- Crohn's disease is a chronic inflammatory condition that may affect any part of the gastrointestinal tract from the mouth to the anus but has a particular tendency to affect the terminal ileum and ascending colon (ileocolonic disease). The disease can involve one small area of the gut such as the terminal ileum, or multiple areas with relatively normal bowel in between (skip lesions). It may also involve the whole of the colon (total colitis) sometimes without small bowel involvement.
- Ulcerative colitis can affect the rectum alone (proctitis), can extend proximally to involve the sigmoid and descending colon (left-sided colitis), or may involve the whole colon (total colitis). In a few of these patients there is also inflammation of the distal terminal ileum (backwash ileitis).

Macroscopic changes

In Crohn's disease the involved small bowel is usually thickened and narrowed. There are deep ulcers and fissures in the mucosa, producing a cobblestone appearance. Fistulae and abscesses may be seen in the colon. An early feature is aphthoid ulceration, usually seen at colonoscopy (Fig. 6.33); later, larger and deeper ulcers appear in a patchy distribution, again producing a cobblestone appearance.

In ulcerative colitis the mucosa looks reddened, inflamed and bleeds easily. In severe disease there is extensive ulceration with the adjacent mucosa appearing as inflammatory polyps.

In fulminant colonic disease of either type, most of the mucosa is lost, leaving a few islands of oedematous mucosa (mucosal islands), and toxic dilatation occurs. On healing, the mucosa can return to normal, although there is usually some residual glandular distortion.

Microscopic changes

In Crohn's disease the inflammation extends through all layers (transmural) of the bowel, whereas in UC a superficial inflammation is seen. In CD there is an increase in chronic inflammatory cells and lymphoid hyperplasia, and in 50–60% of patients granulomas are present. These granulomas are non-caseating epithelioid cell aggregates with Langhans' giant cells.

In ulcerative colitis the mucosa shows a chronic inflammatory cell infiltrate in the lamina propria. Crypt abscesses and goblet cell depletion are also seen.

The differentiation between these two diseases can usually be made not only on the basis of clinical and radiological data but also on the histological differences seen in the rectal and colonic mucosa obtained by biopsy (Table 6.10).

It is occasionally not possible to distinguish between the two disorders, particularly if biopsies are obtained in the acute phase, and such patients are considered to have an indeterminate inflammatory colitis. Serological testing may be of value in differentiating the two conditions (see p. 289).

Extragastrointestinal manifestations

These occur with both diseases. Joint complications are commonest, and the peripheral arthropathies are classified as type 1 (pauci-articular) and type 2 (polyarticular).

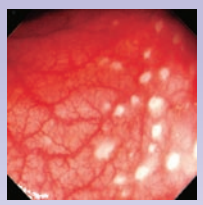
Type 1 attacks are acute, self-limiting (<10 weeks) and occur with IBD relapses; they are associated with other extraintestinal manifestations of IBD activity.

Type 2 arthropathy lasts longer (months to years), is independent of IBD activity and usually associated with uveitis.

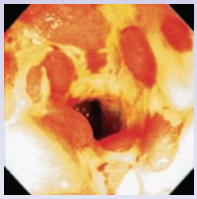
The incidence of joint and other extragastrointestinal manifestations is shown in Table 6.11. There is an association of HLA DRB1*0103 with pauci-articular large joint arthritis in UC and CD and small joint symmetrical arthritis with HLA-B44. HLA B27 is associated with sacroileitis.

Table 6.10 Histological differences between Crohn's disease and ulcerative colitis

	Crohn's disease	Ulcerative colitis
Inflammation	Deep (transmural) Patchy	Mucosal Continuous
Granulomas	++	Rare
Goblet cells	Present	Depleted
Crypt abscesses	+	++



(a)



(b)

Fig. 6.33
Crohn's disease. Colonoscopic appearances of **(a)** aphthoid ulcers typical of Crohn's disease, **(b)** cobblestone appearance.

Table 6.11 Extragastrointestinal manifestations of inflammatory bowel disease (percentage of cases)

	Ulcerative colitis	Crohn's disease
Eyes		
Uveitis	2	5
Episcleritis, conjunctivitis	5–8	3–10
Joints		
Type I (pauci-articular) arthropathy	4	6
Type II (polyarticular) arthropathy	2.5	4
Arthralgia	5	14
Ankylosing spondylitis	1	1.2
Inflammatory back pain	3.5	9
Skin		
Erythema nodosum	1	4
Pyoderma gangrenosum	1	2
Liver and biliary tree		
Sclerosing cholangitis	2.5–7.5	1–2
Fatty liver	Common	Common
Chronic hepatitis	Uncommon	Uncommon
Cirrhosis	Uncommon	Uncommon
Gallstones	As normal population	15–30
Nephrolithiasis	–	5–10 (oxalate stones in patients with small bowel disease or after resection)
Venous thrombosis	5	1

Differential diagnosis

All causes of diarrhoea should be excluded (see Table 6.22) and stool cultures should always be performed. Crohn's disease should be considered in all patients with evidence of malabsorption, e.g. megaloblastic anaemia, or malnourishment, as well as in children with small stature. Ileocolonic tuberculosis (see p. 282) is common in developing countries, e.g. India, which makes a diagnosis of CD difficult. Microscopy and culture for TB of any available tissue is essential in these countries. A therapeutic trial of antituberculosis therapy may be required.

Lymphomas can occasionally involve the ileum and caecum.

Crohn's disease

Clinical features

The major symptoms are diarrhoea, abdominal pain and weight loss. Constitutional symptoms of malaise, lethargy, anorexia, nausea, vomiting and low-grade fever may be present and in 15% of these patients there are no gastrointestinal symptoms. Despite the recurrent nature of this condition, many patients remain well and have an almost normal lifestyle. However, patients with extensive disease often have frequent recurrences, necessitating multiple hospital admissions.

The clinical features are very variable and depend partly on the region of the bowel that is affected. The disease may present insidiously or acutely. The abdominal pain can be colicky, suggesting obstruction, but it usually has no special characteristics and sometimes in colonic disease only minimal discomfort is present. Diarrhoea is present in 80% of all cases and in colonic disease it usually contains blood, making it difficult to differentiate from UC. Steatorrhoea can be present in small bowel disease. Diarrhoea can also be due to bile acid malabsorption occurring as a consequence of the effects of old and presently inactive disease rather than to active disease.

Crohn's disease can present as an emergency with acute right iliac fossa pain mimicking appendicitis. If laparotomy is undertaken, an oedematous reddened terminal ileum is found. There are other causes of an acute ileitis (e.g. infections such as *Yersinia*).

Up to 30% of patients presenting with acute ileitis turn out eventually to have CD. Crohn's disease can be complicated by anal and perianal disease and this is the presenting feature in 25% of cases, often preceding colonic and small intestinal symptoms by many years (Table 6.12).

Enteric fistulae, e.g. to bladder or vagina, occur in 20–40% of cases, equally divided between internal and external fistulae; the latter usually occurring after surgery.

Examination

Physical signs are few, apart from loss of weight and general ill-health. Aphthous ulceration of the mouth is often seen. Abdominal examination is often normal although tenderness and/or a right iliac fossa mass are occasionally found. The mass is due either to inflamed loops of bowel that are matted together or to an abscess. The anus should always be examined to look for oedematous anal tags, fissures or perianal abscesses.

Extragastrointestinal features of inflammatory bowel disease should be looked for (see Table 6.11).

Sigmoidoscopy should always be performed in patients with suspected CD. With small bowel involvement the rectum may appear normal, but a biopsy must be taken as non-specific histological changes can sometimes be found in the mucosa. Even with extensive colonic CD the rectum may be spared and be relatively normal, but patchy involvement with an oedematous haemorrhagic mucosa can be present.

Investigations

Blood tests

- *Anaemia* is common and is usually the normocytic, normochromic anaemia of chronic disease. Deficiency of iron and/or folate also occurs. Despite terminal ileal

Table 6.12 Anal and perianal complications of Crohn's disease

Fissure in ano (multiple and indolent)
Haemorrhoids
Skin tags
Perianal abscess
Ischiorectal abscess
Fistula in ano (may be multiple)
Anorectal fistulae

involvement in CD, megaloblastic anaemia due to B_{12} deficiency is unusual, although serum B_{12} levels can be below the normal range.

- *Raised ESR and CRP* and a raised white cell count.
- *Hypoalbuminaemia* is present in severe disease.
- *Liver biochemistry* may be abnormal.
- *Blood cultures* are required if septicaemia is suspected.
- *Serological tests. Saccharomyces cerevisiae* antibody is usually present while pANCA antibody is negative. The reverse is true in UC but the clinical value of these tests is limited.

Stool cultures

These should always be performed on presentation if diarrhoea is present.

Radiology and imaging

A *barium follow-through* examination or *CT scan* with oral contrast should always be performed in patients suspected of having CD. The findings include an asymmetrical alteration in the mucosal pattern with deep ulceration, and areas of narrowing or stricturing. Although commonly confined to the terminal ileum (Fig. 6.34), other areas of the small bowel can be involved and skip lesions with normal bowel are seen between affected sites.

Imaging of the small bowel may also be performed by *magnetic resonance enteroclysis*.

Colonoscopy is performed if colonic involvement is suspected except in patients presenting with severe acute disease. The findings vary from mild patchy superficial (aphthoid) ulceration to more widespread larger and deeper ulcers producing a cobblestone appearance (see Fig. 6.33).

In patients presenting acutely with colonic symptoms, plain abdominal X-ray, ultrasound or CT is used to outline the colon as for UC (see p. 291).

High-resolution ultrasound and CT scanning are both helpful techniques in defining thickness of the bowel wall and mesentery as well as intra-abdominal and para-intestinal abscesses.

Endoanal ultrasound and MRI are used to evaluate perianal disease.

Radionuclide scans with gallium-labelled polymorphs or indium- or technetium-labelled leucocytes are used in some centres to identify small intestinal and colonic disease and to localize extraintestinal abscesses.

Disease activity

This can be assessed using simple parameters such as Hb, white cell count, inflammatory markers (raised ESR, CRP and platelet count) and serum albumin. Formal clinical activity indices (e.g. CD activity index) are used in research studies. Calprotectin is a calcium-binding protein and accounts for 60% of cytosolic protein of neutrophils. Faecal calprotectin has the potential as a simple cheap non-invasive marker of disease activity in IBD and may be of value in predicting response to and failure of treatment.

Medical management of Crohn's disease

(Box 6.5)

The aim of management is to induce and then maintain a remission. Patients with mild symptoms may require only symptomatic treatment. Cigarette smoking should be stopped. Diarrhoea can be controlled with loperamide, codeine phosphate or co-phenotrope. Diarrhoea in long-standing inactive disease may be due to bile acid malabsorption (see p. 307) and should be treated with colestyramine. Anaemia, if due to vitamin B_{12}, folic acid or iron deficiency, should be treated with the appropriate haematinics. Anaemia in more active disease is usually normochromic and normocytic (anaemia of chronic disease, p. 399) and will usually improve as the patient gets better. Occasionally erythropoietin is required (see p. 632). Patients with active (moderate/severe) attacks may have to be admitted to hospital. Patients with moderate to severe total Crohn's colitis are treated as for UC (see p. 292).

Box 6.5 Options for medical treatment of Crohn's disease

Induction of remission

- Oral or i.v. glucocorticosteroids
- Enteral nutrition
- Oral glucocorticosteroids + azathioprine or 6 mercaptopurine (6MP)

Maintenance of remission

- Aminosalicylates (colonic disease)
- Azathioprine, 6MP, mycophenolate mofetil

Treatment of glucocorticosteroid/immunosuppressive therapy-resistant disease

- Methotrexate
- Intravenous ciclosporin
- Infliximab (TNF-α antibody)
- Adalimumab
- Certolizumab
- New biological agents (Box 6.6)

Perianal disease

- Ciprofloxacin and metronidazole

Induction of remission

Glucocorticosteroids are commonly used to induce remission in moderate and severe attacks of CD (oral prednisolone 30–60 mg/day). In patients with ileocaecal, but not colonic, CD, slow-release formulations of budesonide are as efficacious as oral prednisolone. Budesonide has high topical potency and because of its extensive hepatic inactivation has low systemic availability, which induces less suppression of endogenous cortisol and reduces frequency and intensity of steroidal side-effects. Overall remission/response rates vary from 60% to 90% depending on type, site and extent of disease.

Enteral nutrition is an underutilized means of inducing remission in moderate and severe attacks of CD, and efficacy is independent of nutritional status. If enteral diets with a low fat (1.3% of total calories) and a low linoleic acid content are administered as the sole source of nutrition for 28 days, rates of induction of remission are similar to those obtained with steroids. Relapse rates are high, however, particularly in those with colonic involvement.

In unresponsive patients, remission is sometimes induced and maintained by raising the dose of azathioprine to levels that make patients leucopenic.

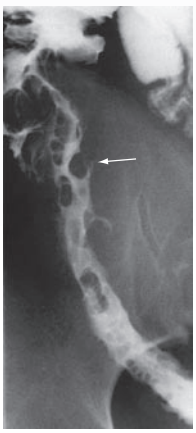


Fig. 6.34 **Terminal ileum on small bowel follow-through in Crohn's disease**, showing narrowing and ulceration of the terminal ileum (arrow). Note the presence of deep 'rose thorn' type ulcers and 'cobblestoning' in the terminal ileum.

Maintenance of remission

Flare ups commonly occur after steroid dosages are tapered and/or enteral nutrition stopped, and alternative treatment strategies have to be introduced. For example, steroid dosage may have to be temporarily increased, to re-induce remission. Remission in patients with Crohn's colitis, but not in those with small intestinal involvement, can be treated with aminosalicylates. Remission in other patients can be maintained with *azathioprine* (AZA, 2.5 mg/kg/day), *6-mercaptopurine* (6MP, 1.5 mg/kg/day), *methotrexate* (25 mg i.m./week) and *mycophenolate mofetil* (1 g/day) (Box 6.5). These drugs have steroid sparing properties. Long-term treatment with these drugs is necessary as the rate of relapse on discontinuation is high (70%).

AZA is the most widely used drug to maintain remission. Careful early monitoring is required as hypersensitivity reactions can occur. The key enzyme involved in AZA and 6MP metabolism is thiopurine methyl transferase (TPMT). This enzyme has a significant genetic variation and deficiencies can result in high circulating levels of drugs with increased risk of bone marrow depression. Assays of TPMT activity are now available and should be performed before treatment. TPMT deficiency is not the only cause of bone marrow depression so 3-monthly blood counts should be performed on all patients.

In patients in whom remissions cannot be induced or maintained with corticosteroid/immunosuppressive therapy, treatment with a biological agent will be indicated.

Biological agents

Infliximab, a chimeric anti TNF-α monoclonal antibody, is the most widely used biological agent. Successful at inducing remission in corticosteroid/immunosuppressive resistant patients, remission can be maintained (5 mg/kg i.v. weeks 0, 2, 6, and 8-weekly until week 46). Further courses of treatment can be given to reinduce and maintain a remission. Up to 60% of patients will form antibodies against infliximab which can shorten duration of response and predispose to an infusion reaction. Formation of antibodies against infliximab can be reduced by prior treatment with AZA and incidence of infusion reactions reduced by prior i.v. infusion of 200 mg hydrocortisone. Other adverse effects include opportunistic infections. The incidence of neoplasms and lymphoma is not increased on treatment.

Two other TNF-α antibody inhibitors have now been widely used.

Adalimumab is a fully human anti-TNF monoclonal antibody that is administered subcutaneously and is effective in patients who have failed to respond to infliximab as well as those who have not received it.

Certolizumab pegol is a pegylated Fab′ fragment of humanized anti TNF-α monoclonal antibody. It can also be administered subcutaneously and short-term and longer-term efficacy has been demonstrated.

New anti TNF-α antibody therapies include CDP571, etanercept and onercept. Box 6.6 summarizes some of the other novel biological agents that are being trialled for the treatment of CD.

Surgical management of Crohn's disease

Approximately 80% of patients will require an operation at some time during the course of their disease. Nevertheless, surgery should be avoided if possible and only minimal resections undertaken, as recurrence (15% per year) is almost inevitable. Patients undergoing their second surgery for CD should be treated with AZA or 6MP to reduce the chance of recurrence. The evidence base is not strong enough to support this recommendation for patients undergoing their first surgery. The indications for surgery are:

- failure of medical therapy, with acute or chronic symptoms producing ill-health
- complications (e.g. toxic dilatation, obstruction, perforation, abscesses, enterocutaneous fistula)
- failure to grow in children despite medical treatment.

In patients with small bowel disease, some strictures can be widened (stricturoplasty), whereas others require resection and end-to-end anastomosis.

When colonic CD involves the entire colon and the rectum is spared or minimally involved a subtotal colectomy and ileorectal anastomosis may be performed. An eventual recurrence rate of 60–70% in the ileum, rectum or both is to be expected; however, two-thirds of these patients retain a functional rectum for 10 years. If the whole colon and rectum are involved a panproctocolectomy with an end ileostomy is the standard operation. In this operation the colon and rectum are removed and the ileum is brought out through an opening in the right iliac fossa and attached to the skin. The patient wears an ileostomy bag, which is stuck on to the skin over the ileostomy spout. The bag needs to be emptied once or twice daily, so this is compatible with a near-normal lifestyle. Stoma care therapists are readily available with help and advice. CD patients are not suitable for a pouch operation (see p. 292) as recurrence in the pouch is high.

Problems associated with ileostomies include:

- mechanical problems
- dehydration, particularly in hot climates
- psychosexual problems
- infertility in men (due to rectal surgery)
- recurrence of CD.

Box 6.6 Some biological agents under evaluation for treatment of Crohn's disease

Compound (generic name)	Therapeutic target	Compound class
Visilizumab	T cells	Humanized mAb
Abatacept	T cells, dendritic cells macrophages	Fusion protein
CNT 01275/ ABT-874 (J 695)	Interleukin 12/23p40 Interleukin 12	Humanized mAb Humanized mAb
Atlizumab (MRA)	Interleukin 6	Humanized mAb
Natilizumab	α4-integrin	Human mAb
Alicaforsen (ISIS-2302) enema	Intracellular adhesion molecule 1	Antisense oligonucleotide
Sagramostim	PMN-macrophage	Growth factors

After Baumgart DC, Sandborn WJ. Inflammatory bowel disease: clinical aspects and established and evolving therapies. *Lancet* 2007; **369**: 1641–1657, with permission from Elsevier.

Ulcerative colitis

Clinical features

The major symptom in UC is diarrhoea with blood and mucus, sometimes accompanied by lower abdominal dis-

comfort. General features include malaise, lethargy and anorexia with weight loss. Aphthous ulceration in the mouth is seen. The disease can be mild, moderate or severe, and in most patients runs a course of remissions and exacerbations. Ten per cent of patients have persistent chronic symptoms, while some patients may have only a single attack.

When the disease is confined to the rectum (proctitis), blood mixed with the stool, urgency and tenesmus are common. There are normally few constitutional symptoms, but patients are nevertheless greatly inconvenienced by the frequency of defecation.

In an acute attack of UC, patients have bloody diarrhoea, passing up to 10–20 liquid stools per day. Diarrhoea also occurs at night, with urgency and incontinence that is severely disabling for the patient. Occasionally blood and mucus alone are passed.

Severe attack. The definition of a severe attack is given in Table 6.13. The patient is often very ill and needs urgent treatment in hospital.

Toxic megacolon is a serious complication. The plain abdominal X-ray shows a dilated thin-walled colon with a diameter of >6 cm; it is gas filled and contains mucosal islands (Fig. 6.35). It is a particularly dangerous stage of advanced disease with impending perforation and a high mortality (15–25%). Urgent surgery is required in all patients in whom toxic dilatation has not resolved within 48 hours with intensive therapy as above. The differential diagnosis includes an infectious colitis, e.g. *C. difficile* and CMV, particularly in patients with HIV infections.

Table 6.13 Definition and management of a severe attack of ulcerative colitis

(a) Definition	
Stool frequency	>6 stools per day with blood +++
Fever	>37.5°C
Tachycardia	>90 per minute
ESR	>30 mm per hour
Anaemia	<10 g/dL haemoglobin
Albumin	<30 g/L

(b) Management

Admit to hospital
Assess i.v. fluids
Monitor daily:
- stool frequency
- abdominal X-ray
- FBC, CRP
- albumin

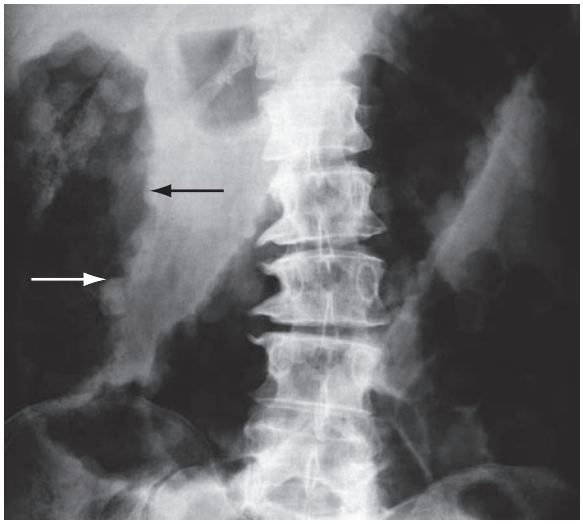


Fig. 6.35 **Plain abdominal X-ray showing toxic dilatation.** The arrows indicate mucosal islands.

Examination

In general there are no specific signs in UC. The abdomen may be slightly distended or tender to palpation. The anus is usually normal. Rectal examination will show the presence of blood. Rigid sigmoidoscopy is usually abnormal, showing an inflamed, bleeding, friable mucosa. Very occasionally rectal sparing occurs, in which case sigmoidoscopy will be normal.

Investigations

Blood tests

- In moderate to severe attacks an iron deficiency anaemia is commonly present and the white cell and platelet counts are raised.
- The ESR and CRP are often raised; liver biochemistry may be abnormal, with hypoalbuminaemia occurring in severe disease.
- pANCA may be positive. This is contrary to CD, where pANCA is usually negative (see p. 289).

Stool cultures

These should always be performed to exclude infective causes of colitis.

Imaging

A plain abdominal X-ray with an abdominal ultrasound are the key investigations in moderate to severe attacks. The extent of disease can be judged by the air distribution in the colon and the presence of colonic dilatation can be noted. Thickening of the colonic wall can be detected on ultrasound as can the degree of hyperaemia in the colonic wall, present in severe disease. The presence of free fluid within the abdominal cavity can also be assessed.

Colonoscopy

A colonoscopy should not be performed in severe attacks of disease for fear of perforation. In more longstanding and chronic disease it is useful in defining extent and activity of disease, and excluding the onset of dysplasia and carcinoma in patients with longstanding disease (see p. 293) of 10 years' duration or more.

Radionuclide scans

These can be used to assess colonic inflammation (see p. 245).

Medical management of ulcerative colitis (UC)

Wherever possible, patients with UC and CD should be managed in patient-focused inflammatory bowel disease clinics. All patients with UC should be treated with an aminosalicylate. The active moiety of these drugs is 5-aminosalicylic acid (5-ASA). 5-ASA is absorbed in the small intestine (and may be nephrotoxic), so the design of the various aminosalicylate preparations is based on the binding of 5-ASA by an azo bond to sulfapyridine

(sulfasalazine), 4-aminobenzoyl-β-alanine (balsalazide) or to 5-ASA itself (olsalazine), coating with a pH-sensitive polymer (Asacol) or packaging of 5-ASA in microspheres (Pentasa).

The azo bonds are broken down by colonic bacteria to release 5-ASA within the colon. The pH-dependent forms are designed to release 5-ASA in the terminal ileum. Luminal pH profiles in patients with inflammatory bowel disease are abnormal and in some patients capsules of 5-ASA coated with pH-sensitive polymer may pass through into the faeces intact. 5-ASA is released from microspheres throughout the small intestine and colon. The mode of action of 5-ASA in inflammatory bowel disease is unknown, but the aminosalicylates have been shown to be effective in inducing remission in mild to moderately active disease and maintaining remission in all forms of disease. Sulfasalazine is being used less frequently because of its wider side-effect profile from the sulphonamide component.

Proctitis

Oral aminosalicylates plus a local rectal steroid preparation (10% hydrocortisone foam; prednisolone 20 mg enemas or foam) are the first-line treatment. Mesalazine enemas and budesonide enemas can be tried. Some cases of proctitis can be 'resistant' to treatment. In these, oral corticosteroids alone or in combination with azathioprine are used and in rare cases short-chain fatty acid enemas may help.

Left-sided proctocolitis

Oral aminosalicylates plus local rectal steroid preparations may be effective but in moderate to severe attacks oral prednisolone will be required. If patients do not respond within 2 weeks they should be admitted to hospital.

Total colitis (moderate to severe attacks)

(Table 6.13)

Patients should be admitted to hospital and treated initially with hydrocortisone 100 mg i.v. 6-hourly with oral aminosalicylates. Full investigations (see above) should be performed initially and full supportive therapy administered (i.v. fluids, nutritional support via the enteral (not parenteral) route if required).

Patients who have been previously admitted within 2–3 years with moderate to severe attacks of total colitis can be started on azathioprine if not already on this treatment, as it takes time to work. The clinical status of patients should subsequently be monitored (fever, tachycardia, abdominal signs) and daily FBC, ESR, CRP, electrolytes and urea, tests of liver function, including serum albumin, straight abdominal X-ray and stool weights should be performed. Success or failure of medical treatment of a severe attack of UC must be judged by an experienced gastroenterologist.

A persistent fever, tachycardia, falling Hb, rising white cell count, falling potassium, falling albumin and persistently raised stool weights (>500 g/day) with loose blood-stained stool are all signs that the patient is not responding to treatment and that *surgery* may be indicated.

A concept of *rescue therapy* to avoid colectomy has now been developed for patients with poor prognostic signs and a CRP > 45 mg/L after 3 days of i.v. hydrocortisone. They should be treated, in addition to hydrocortisone, with either i.v. ciclosporin 2 mg/kg/day or possibly TNF-α antibody therapy. Other rescue agents under consideration include visilizumab and *leucocytapheresis* (selective removal of white cells from blood and re-infusion of red cells and leukocyte-poor plasma). Delaying colectomy in very sick patients can increase mortality.

In patients responding to i.v. hydrocortisone treatment, oral prednisolone therapy should be substituted and doses slowly tailed off (5–10 mg weekly). Maintenance of remission is with aminosalicylates. In patients in whom it is not possible to reduce the dose of prednisolone without flare-up, azathioprine is used.

Surgical management of ulcerative colitis

While the treatment of UC remains primarily medical, surgery continues to have a central role because it may be life-saving, is curative and eliminates the long-term risk of cancer. The main indication for surgery is for a severe attack which fails to respond to medical therapy. Other indications are listed in Box 6.7.

In acute disease, subtotal colectomy with end ileostomy and preservation of the rectum is the operation of choice. At a later date a number of surgical options are available. These include proctectomy with a permanent ileostomy. To avoid a permanent ileostomy, ileorectal anastomosis can be performed; annual biopsies of the rectal mucosa must be carried out to detect dysplasia, a histological change that precedes the development of a rectal stump carcinoma. With an ileo-anal anastomosis (Fig. 6.36), a pouch of ileum is formed that acts as a reservoir. The pouch is anastomosed to the anus at the dentate line following endoanal excision of the mucosa of the distal rectum and anal canal. Continence is usually achieved. A third of patients, however, will experience '*pouchitis*' in which there is inflammation of the pouch

Box 6.7 Indications for surgery in ulcerative colitis

Fulminant acute attack

- Failure of medical treatment (3 days)
- Toxic dilatation
- Haemorrhage
- Perforation

Chronic disease

- Incomplete response to medical treatment
- Excessive steroid requirement
- Non-compliance with medication
- Risk of cancer

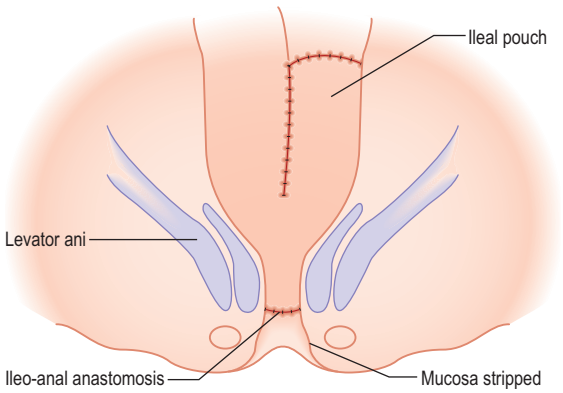


Fig. 6.36 **An ileo-anal pouch for ulcerative colitis.**

mucosa with clinical symptoms of diarrhoea, bleeding, fever and at times exacerbation of extracolonic manifestations (Fig. 6.37). The incidence of pouchitis is twice as high in patients with primary sclerosing cholangitis. Two-thirds of pouchitis cases will recur either as acute relapsing or chronic unremitting forms. Treatment is not always satisfactory and includes topical and oral 5-ASA, corticosteroids, metronidazole and ciprofloxacin. Probiotics (four strains of lactobacillus, three strains of bifidobacterium and one of streptococcus) have been shown to prevent the onset of pouchitis and to maintain remission in pouchitis patients. Probiotics have also been used as maintenance therapy in UC (Box 6.8). Finally, short-chain fatty acid enemas and alicaforsen (a selective inhibitor of intercellular adhesion molecule 1, ICAM-1, expression) enemas have shown promise in the treatment of pouchitis.

Course and prognosis

A third of patients with distal inflammatory proctitis due to UC will develop more proximal disease, with 5–10% developing total colitis. A third of patients with UC will have a single attack and the others will have a relapsing course. A third of patients with UC will undergo colectomy within 20 years of diagnosis.

Cancer in inflammatory bowel disease

Patients with UC and Crohn's colitis have an increased incidence of developing colon cancer and patients with CD of the small intestine a small increase in incidence of small bowel carcinoma. The following screening strategies are suggested, although evidence for some is lacking. A screening colonoscopy with multiple biopsies should be done every 2 years in patients with total UC and Crohn's colitis of 10 years' standing and annually after 20 years. Patients with left-sided UC and those with less than total Crohn's colitis should have a screening colonoscopy at 15 years and biannually thereafter. Patients with UC and primary sclerosing cholangitis are particularly at risk of developing colon cancer and should have a yearly screening colonoscopy.

Pregnancy and inflammatory bowel disease

Women with inactive IBD have normal fertility. Fertility, however, may be reduced in those with active disease, and patients with active disease are twice as likely to suffer spontaneous abortion than those with inactive disease.

The rate of relapse of UC in pregnant patients is similar to non-pregnant patients and is often due to inappropriate discontinuation of maintenance therapy. The risk of a flare up in the puerperal period is enhanced in patients who have a flare up in the first trimester. Patients with CD like those with UC do not have an increased risk of flare up during pregnancy. Relapse (if it does occur) is, however, more likely during the first trimester.

Aminosalicylates, steroids and azathioprine are safe at the time of conception and during pregnancy. The sulfapyridine moiety of sulfasalazine impairs spermatogenesis, so the partners of women trying to conceive should be treated with an alternative aminosalicylate. There is no good evidence that male patients with IBD should stop either AZA or 6MP.

Mortality in inflammatory bowel disease

Population-based studies demonstrate a mortality in UC similar to that in the general population. The two exceptions are patients with severe colitis who have a slightly higher mortality in the first year after diagnosis and patients aged

Box 6.8 Probiotic use in inflammatory bowel disease

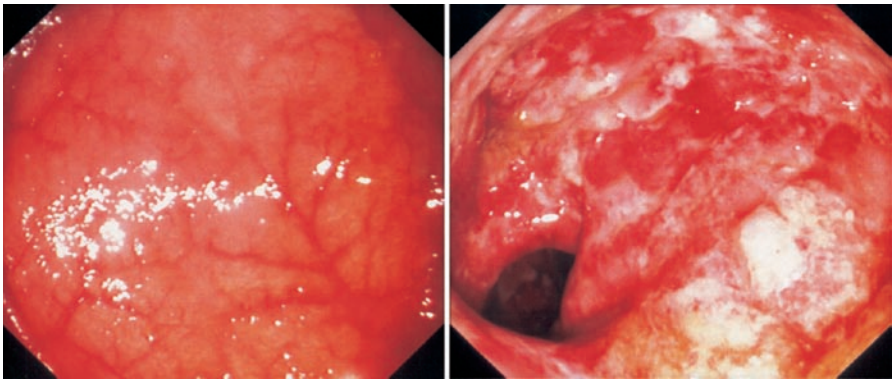
Probiotics are live microorganisms which when ingested can modify the composition of enteric microflora. Commonly used probiotics are lactobacilli, bifidobacteria, non-pathogenic *E. coli*.

Rationale for use

- Germ-free animals are susceptible to gastrointestinal infections and inflammation
- Pseudomembranous colitis due to *C. difficile* follows antibiotic therapy
- In vitro isolates of normal bacteria inhibit growth of pathogenic bacteria
- Regulatory signals between bacterial flora and intestinal epithelial cells maintain mucosal integrity

Use

- Pouchitis (see p. 292) – probiotics can prevent onset and maintain a remission, possibly by increasing tissue levels of IL-10 and reducing TH-1 cytokine production as well as normalizing colonic bacterial function
- May be useful in maintaining remissions in ulcerative colitis



(a) (b)

Fig. 6.37 **Pouchitis.** Fibreoptic sigmoidoscopic appearances. **(a)** Mild changes. **(b)** Severe haemorrhagic ulceration.

over 60 at the time of diagnosis. Although currently it is unclear whether there is a slightly higher overall mortality in patients with CD or not, those with extensive jejunal and ileal disease and those with gastric and duodenal disease have been shown to have a higher relative mortality.

Microscopic inflammatory colitis

Patients with this group of disorders present with chronic or fluctuating watery diarrhoea. Although the macroscopic features on colonoscopy are normal, the histopathological findings on biopsy are abnormal. There are three distinct forms of microscopic inflammatory colitis:

- *Microscopic ulcerative colitis*. There is a chronic inflammatory cell infiltrate in the lamina propria, with deformed crypt architecture, and goblet cell depletion with or without crypt abscesses. Treatment is as for UC; many patients respond to treatment with aminosalicylates alone.
- *Microscopic lymphocytic colitis*. There is surface epithelial injury, prominent lymphocytic infiltration in the surface epithelium and increased lamina propria mononuclear cells. It affects males and females equally and is associated with a high prevalence of antibiotic use.
- *Microscopic collagenous colitis*. There is a thickened subepithelial collagen layer (>10 μm) adjacent to the basal membrane with increased infiltration of the lamina propria with lymphocytes and plasma cells and surface epithelial cell damage. It is predominantly a disorder of middle-aged or elderly females, and is associated with a variety of autoimmune disorders (arthritis, thyroid disease, CREST syndrome (see p. 546) and primary biliary cirrhosis). The prevalence of collagenous colitis has been shown to be $15.7/10^5$ population with an annual incidence of $1.8/10^5$ population. There are a number of reports linking drugs to the development of collagenous colitis, e.g. NSAIDs, simvastatin, H_2-receptor antagonists. The incidence of both microscopic lymphocytic and collagenous colitis is increased in patients with coeliac disease and this must be excluded in these patients.

There are no controlled clinical trials of treatment in microscopic lymphocytic or collagenous colitis. Treatment is usually with aminosalicylates, budesonide, bismuth-containing preparations (tripotassium dicitratobismuthate, Pepto-Bismol), prednisolone and azathioprine in that order. A small number of patients with microscopic lymphocytic and collagenous colitis have co-existing bile acid malabsorption and as such can respond to colestyramine. In general, patients with microscopic collagenous colitis are easier to treat than those with the lymphocytic form of the disease.

FURTHER READING

Achkar JP, Duerr R. The expanding universe of inflammatory bowel disease genetics. *Current Opinion in Gastroenterology* 2008; **24**(4): 429–434.

Baumgart DC, Carding SR. Inflammatory bowel disease: Cause and immunobiology. *Lancet* 2007; **369**: 1627–1640.

Baumgart DC, Sandborn WJ. Inflammatory bowel disease: clinical aspects and established and evolving therapies. *Lancet* 2007; **369**: 1641–1657.

Carter MJ, Lobo AJ, Travis SP. Guidelines for the management of inflammatory bowel disease in adults. *Gut* 2004; **53**(suppl V): 1–16.

Dubois PC, van Heel DA. New susceptibility genes for ulcerative colitis. *Nature Genetics* 2008; **40**: 686–688.

Fisher SA, Tremelling M, Anderson CA et al. Genetic determinants of ulcerative colitis include the ECM1 locus and five loci implicated in Crohn's disease. *Nature Genetics* 2008; **40**: 710–712.

Heetun ZSS, Byurnes C, Neary P et al. Review article: reproduction in patients with inflammatory bowel disease. *Alimentary Pharmacology and Therapeutics* 2007; **26**: 513–533.

Peyrin-Biroulet L, Desreumaux P, Sandborn W et al. Crohn's disease; beyond antagonists of tumour necrosis factor. *Lancet* 2008; **372**: 67–81.

Travis S. Review article: saving the colon in severe colitis – the case for medical therapy. *Alimentary Pharmacology and Therapeutics* 2006; **24**(suppl 3): 68–73.

THE COLON AND RECTUM

Structure

The large intestine starts at the caecum, on the posterior medial wall of which is the appendix. The colon is made up of ascending, transverse, descending and sigmoid parts, which join the rectum at the rectosigmoid junction.

The muscle wall consists of an inner circular layer and an outer longitudinal layer. The outer layer is incomplete, coming together to form the taenia coli, which produce the haustral pattern seen in the normal colon.

The mucosa of the colon is lined with epithelial cells with crypts but no villi, so that the surface is flat. The mucosa is full of goblet cells. A variety of cells, mainly lymphocytes and macrophages, are found in the lamina propria.

The blood supply to the colon is from the superior and inferior mesenteric vessels. Generally there are good anastomotic channels, but the caecum and splenic flexure are areas where ischaemia can occur. The colon is innervated mainly by the enteric nervous system with input from the parasympathetic and sympathetic pathways. Spinal afferent neurones from the dorsal root ganglia innervate the entire colon.

The rectum is about 12 cm long. Its interior is divided by three crescentic circular muscles producing shelf-like folds. These are the rectal valves that can be seen at sigmoidoscopy. The anal canal has an internal and an external sphincter.

Physiology of the colon

The main roles of the colon are the absorption of water and electrolytes (Table 6.14) and the propulsion of contents from the caecum to the anorectal region. Approximately 1.5–2 L of fluid pass the ileocaecal valve each day. Absorption is stimulated by short-chain fatty acids which are produced predominantly in the right colon by the anaerobic metabolism of dietary fibre by bacterial polysaccharidase enzyme systems. Colonic contents are mixed, aiding absorption by non-propagative segmenting muscular contractions. High-amplitude propagative colonic contractions cause propulsion. Peristalsis is induced by the release of serotonin (5-HT) from neuroendocrine cells in response to luminal distension.

Table 6.14 Input and output of water and electrolytes in the gastrointestinal tract over 24 hours

	Water (mL)	Sodium (mmol)	Potassium (mmol)
Input			
Diet	1500	150	80
GI secretions	7500	1000	40
Totals	9000	1150	120
Output			
Faeces	150	5	12
Ileostomy (adapted)	500–1000	60–120	4

Serotonin activates the HT_4 receptors, which in turn results in the activation of sensory (calcitonin gene-related peptide (CGRP)) neurones. Normal colonic transit time is 24–48 hours with normal stool weights of up to 250 g/day.

Physiology of defecation

The role of the rectum and anus in defecation is complex. The rectum is normally empty. Stool is propelled into the rectum by propagated colonic contractions. Sensation of fullness, a desire to defecate and urgency to defecate are experienced with increasing volumes of rectal content (threshold 100 mL). The sensations are associated with rectal contraction and a relaxation of the internal anal sphincter, both of which serve to push the stool down into the proximal anal canal. This increases the defecatory urge, which can only be suppressed by vigorous contraction of the external sphincter and puborectalis. If conditions are appropriate for defecation the subject sits or squats, contracts the diaphragm and abdominal muscles and relaxes the pelvic floor muscles, including the puborectalis, and the anal sphincter muscles with the result that stool is expelled.

CONSTIPATION

'Constipation' is a very common symptom, particularly in women and the elderly. It is often more of a perception than a real entity. A consensus definition used in research (The Rome III criteria) defines constipation as having two or more of the following for at least 12 weeks: infrequent passage of stools (<3/week), straining > 25% of time, passage of hard stools, incomplete evacuation and sensation of anorectal blockage. According to these definitions 'constipation' affects more than 1 in 5 of the population.

Many symptoms are attributed by patients to constipation and include headaches, malaise, nausea and a bad taste in the mouth. Other symptoms include abdominal bloating and/or discomfort (undistinguishable from the irritable bowel syndrome) as well as local and perianal pain. The causes of constipation are shown in Table 6.15.

Table 6.15 Causes of constipation

General
Pregnancy
Inadequate fibre intake
Immobility

Metabolic/endocrine
Diabetes mellitus
Hypercalcaemia
Hypothyroidism
Porphyria

Functional
Irritable bowel syndrome
Idiopathic slow transit

Drugs
Opiates
Antimuscarinics
Calcium-channel blockers, e.g. verapamil
Antidepressants, e.g. tricyclics
Iron

Neurological
Spinal cord lesions
Parkinson's disease

Psychological
Depression
Anorexia nervosa
Repressed urge to defecate

Gastrointestinal disease
Intestinal obstruction and pseudo-obstruction
Colonic disease, e.g. carcinoma, diverticular disease
Aganglionosis, e.g. Hirschprung's disease, Chagas' disease
Painful anal conditions, e.g. anal fissure

Defecatory disorders
Rectal prolapse, mucosal prolapse intussusception and solitary rectal ulcer syndrome
Large rectocele
Pelvic floor dyssynergia/anismus
Megarectum

Assessment of constipation

This relies on the history. When there has been a recent change in bowel habit in association with other symptoms (e.g. rectal bleeding) a barium enema or colonoscopy is indicated. A barium enema should always be preceded by a rectal examination and rigid sigmoidoscopy to exclude anorectal lesions that can otherwise be missed. By these means, gastrointestinal causes such as colorectal cancer and narrowed segments due to diverticular disease (Table 6.15) can be excluded.

Constipation can be classified into three broad categories but there is much overlap:

- normal transit through the colon (59%)
- defecatory disorders (25%)
- slow transit (13%).

Defecatory disorders with slow transit can occur together (3%).

Normal-transit constipation

In normal-transit constipation, stool traverses the colon at a normal rate, the stool frequency is normal and yet patients believe they are constipated. This is likely to be due to perceived difficulties of evacuation or the passage of hard stools. Patients may complain of abdominal pain or bloating.

Normal-transit constipation can be distinguished from slow-transit constipation by undertaking marker studies of colonic transit. Capsules containing 20 radio-opaque shapes are swallowed on days 1, 2 and 3 and an abdominal X-ray

obtained 120 hours after ingestion of the first capsule. Each capsule contains shapes of different configuration and the presence of more than 4 shapes from the first capsule, 6 from the second and 12 from the third denotes moderate to severe slow transit (Fig. 6.38).

Slow-transit constipation

Slow-transit constipation occurs predominantly in young women who have infrequent bowel movements (usually less than once a week). The condition often starts at puberty and the symptoms are usually an infrequent urge to defecate, bloating, abdominal pain and discomfort, which can make the condition difficult to distinguish from constipation-predominant irritable bowel syndrome. Some patients with severe slow-transit constipation have delayed emptying of the proximal colon and others a failure of 'meal-stimulated' colonic motility. Histopathological abnormalities have been demonstrated in the colons of some patients with severe slow-transit constipation, and some patients have coexisting disorders of small intestinal motility, consistent with a diagnosis of chronic idiopathic intestinal pseudo-obstruction (see p. 317).

Defecatory disorders

A 'paradoxical' contraction rather than the normal relaxation of the puborectalis and external anal sphincter and associated muscles during straining may prevent evacuation (pelvic floor dyssynergia, anismus). These are mainly due to dysfunction of the anal sphincter and pelvic floor. An anterior rectocele is a common problem where there is a weakness of the rectovaginal septum, resulting in protuberance of the anterior wall of the rectum with trapping of stool if the diameter is >3 cm. In some patients the mucosa of the anterior rectal wall prolapses downwards during straining (see p. 300) impeding the passage of stool, whilst in others there may be a higher mucosal intussusception.

In some patients the rectum can become unduly sensitive to the presence of small volumes of stool, resulting in the urge to pass frequent amounts of small-volume stool and the sensation of incomplete evacuation.

The defecatory disorders can often be characterized by performing evacuation proctography (Fig. 6.39) and tests of anorectal physiology.

Treatment

Any underlying cause should be treated. In patients with normal and slow-transit constipation the main focus should be directed to increasing the fibre content of the diet in conjunction with increasing fluid intake. Fibre intake should be increased by dietary means rather than by prescribing commercially available fibre sources in order to avoid substrate inducibility of colonic bacterial polysaccharidase enzyme systems which can lead in turn to excessive flatus production. These patients should therefore be referred to a dietician.

The use of laxatives should be restricted to severe cases. Types of laxatives available are listed in Box 6.9. Osmotic laxatives act by increasing colonic inflow of fluid and electrolytes; this acts not only to soften the stool but to stimulate colonic contractility. Magnesium sulphate 5–10 g dissolved in a glass of hot water should be taken before breakfast; it works in 2–4 hours. The polyethylene glycols (Macrogols) have the advantage over the synthetic disaccharide lactulose in that they are not fermented anaerobically in the colon to gas which can distend the colon to cause pain. The osmotic laxatives are preferred to the stimulatory laxatives, which act by stimulating colonic contractility and by causing intestinal secretion. The use of irritant suppositories can be helpful in some patients with defecatory disorders. The use of enemas should be restricted to the management of elderly, infirm and immobile patients and those with neurological disorders.

Patients with defecatory disorders should be referred to a specialist centre as surgery may be indicated, for example, for anterior rectocele or internal anal mucosal intussusception. Anterior mucosal prolapse can be treated by injection, and those with pelvic floor dyssynergia (anismus) can benefit from biofeedback therapy.

FURTHER READING

Lembo A, Camilleri M. Chronic constipation. *New England Journal of Medicine* 2003; **349**: 1360–1368.

Fig. 6.38 **Slow-transit constipation.** Straight abdominal X-ray taken on day 6 after ingestion of capsules each containing 20 radio-opaque shapes, which were administered daily on days 1, 2 and 3. All the markers are retained, confirming the diagnosis of severe slow-transit constipation.

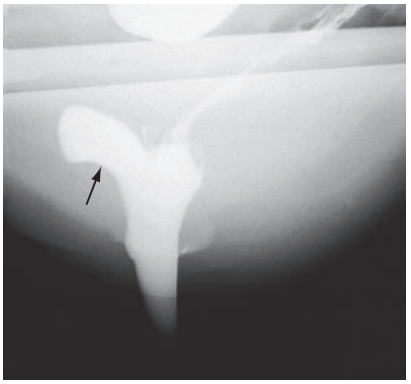


Fig. 6.39 **Evacuating proctogram** showing the presence of an anterior rectocele (arrowed).

Box 6.9 Laxatives and enemas

Bulking-forming laxatives

Dietary fibre
Wheat bran
Methylcellulose
Mucilaginous gums – sterculia
Mucilaginous seeds and seed coats, e.g. ispaghula husk

Stimulant laxatives (stimulate motility and intestinal secretion)

Phenolphthalein
Bisacodyl
Anthraquinones – senna and dantron (only for the terminally ill)
Docusate sodium
Methylnaltrexone (for opioid induced constipation)
Lubiprostone

Osmotic laxatives

Magnesium sulphate
Lactulose
Macrogols

Suppositories

Bisacodyl
Glycerol

Enemas

Arachis oil
Docusate sodium
Hypertonic phosphate
Sodium citrate

MISCELLANEOUS COLONIC CONDITIONS

Megacolon

The term 'megacolon' is used to describe a number of congenital and acquired conditions in which the colon is dilated. In many instances it is secondary to chronic constipation and in some parts of the world Chagas' disease is a common cause.

All young patients with megacolon should have Hirschsprung's disease excluded. In this disease, which presents in the first years of life, an aganglionic segment of the rectum (megarectum) gives rise to constipation and subacute obstruction. Occasionally Hirschsprung's disease affecting only a short segment of the rectum can be missed in childhood. A preliminary rectal biopsy is performed and stained with special stains for ganglion cells in the submucosal plexus. In doubtful cases full-thickness biopsy, under anaesthesia, should be obtained. A frozen section is stained for acetylcholinesterase, which is elevated in Hirschsprung's disease. Manometric studies show failure of relaxation of the internal sphincter, which is diagnostic of Hirschsprung's disease. This disease can be successfully treated surgically.

Treatment of other causes of a megacolon is similar to that of slow-transit constipation, but saline washouts and manual removal of faeces are sometimes required.

Faecal incontinence

Seven per cent of the healthy population over the age of 65 experience incontinence at least weekly. Incontinence occurs when the intrarectal pressure exceeds the intra-anal pressure and is classified as minor (inability to control flatus or liquid stool, causing soiling) or major (frequent and inadvertent evacuation of stool of normal consistency). The common causes of incontinence are shown in Table 6.16. Obstetric injury is a common cause and sphincter defects have been found in up to 30% of primiparous women. A faecal incontinence rate of over 25% has been described and this gets more severe with increasing age. Endoanal ultrasonography with enhancing per vaginal views is the investigation of choice in the assessment of anal sphincter damage (Fig. 6.40). Neurophysiological investigation of pudendal nerve function, anal sensation and anal sphincter function may be required to elicit the cause of the problem. In the future, static and dynamic pelvic MRI may become the investigation of choice in faecal incontinence.

Initial *management* of minor incontinence is bowel habit regulation. Loperamide is the most potent antidiarrhoeal agent which also increases internal sphincter tone.

Biofeedback is effective in some patients with faecal incontinence associated with impaired function of the puborectalis muscle and the external anal sphincter. Sacral spinal nerve stimulation has been shown to be effective in the treatment of patients with a functionally deficient but morphologically intact external anal sphincter.

Surgery may be required for anal sphincter trauma and should be carried out in specialist centres.

Ischaemic disease of the colon (ischaemic colitis)

Occlusion of branches of the superior mesenteric artery (SMA) or inferior mesenteric artery (IMA), often in the older age group, commonly presents with sudden onset of abdom-

FURTHER READING

Wald A. Faecal incontinence in adults. *New England Journal of Medicine* 2007; **356**: 1648–1655.

Table 6.16 Aetiology of faecal incontinence

Congenital
e.g. imperforate anus

Anal sphincter dysfunction
Structural damage:
- Surgery – anorectal, vaginal hysterectomy
- Obstetric injury during childbirth
- Trauma
- Radiation
- Perianal Crohn's disease

Pudendal nerve damage:
- Childbirth

Perineal descent:
- Prolonged straining at stool

Rectal prolapse

Faecal impaction with overflow diarrhoea

Severe diarrhoea
e.g. ulcerative colitis, functional diarrhoea, irritable bowel syndrome

Neurological and psychological disorders
Spinal trauma (S2–S4)
Spina bifida
Stroke
Multiple sclerosis
Diabetes mellitus (with autonomic involvement)
Dementia
Psychological illness

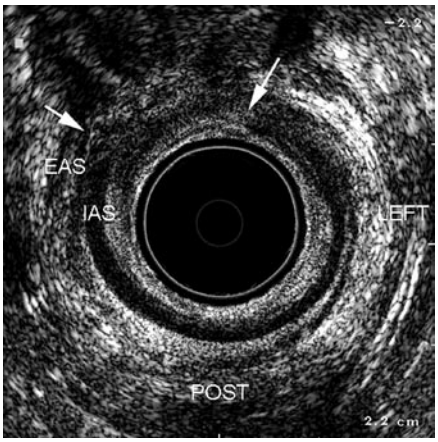


Fig. 6.40 **Endoanal ultrasound scan,** axial mid canal image, showing a large tear between 10 and 1 o'clock (arrows) following vaginal delivery, involving the external (EAS) and internal anal sphincters (IAS) and resulting in faecal incontinence. Courtesy of Professor Clive Bartram, Princess Grace Hospital, London.

inal pain and the passage of bright red blood per rectum, with or without diarrhoea. There may be signs of shock and evidence of underlying cardiovascular disease. The majority of cases affect the splenic flexure and left colon. This condition has also been described in women taking the contraceptive pill and in patients with thrombophilia (see p. 442) and small- or medium-vessel vasculitis (see p. 550).

On *examination* the abdomen is distended and tender. A straight abdominal X-ray often shows thumb-printing (a characteristic sign of ischaemic disease) at the site of the splenic flexure.

The *differential diagnosis* is of other causes of acute colitis. A rigid sigmoidoscopy (normal mucosal appearances often with presence of blood) and *a gentle instant enema* is preferred to colonoscopy (to avoid perforation) in cases where the diagnosis is in doubt. A barium enema should be performed when the patient has fully recovered, to exclude the formation of a stricture at the site of disease. Patients without evidence of underlying cardiovascular disease should be screened for thrombophilia and vasculitis.

Treatment

Most patients settle on symptomatic treatment. A few develop gangrene and perforation and require urgent surgery.

Pneumatosis cystoides intestinalis

This is a rare condition in which multiple gas-filled cysts are found in the submucosa of the intestine, chiefly the colon. The cause is unknown but some cases are associated with chronic obstructive pulmonary disease. Patients are usually asymptomatic, but abdominal pain and diarrhoea do occur and occasionally the cysts rupture to produce a pneumoperitoneum. This condition is diagnosed on X-ray of the abdomen, barium enema or sigmoidoscopy when cysts are seen.

Treatment is often unnecessary but continuous oxygen therapy will help to disperse the largely nitrogen-containing cysts. Metronidazole may help.

DIVERTICULAR DISEASE

Diverticula are frequently found in the colon and occur in 50% of patients over the age of 50 years. They are most frequent in the sigmoid, but can be present over the whole colon.

The term *diverticulosis* indicates the presence of diverticula; *diverticulitis* implies that these diverticula are inflamed. It is perhaps better to use the more general term diverticular disease, as it is often difficult to be sure whether the diverticula are inflamed. The precise mechanism of diverticula formation is not known. There is thickening of the muscle layer and, because of high intraluminal pressures, pouches of mucosa extrude through the muscular wall through weakened areas near blood vessels to form diverticula. An alternative explanation is cholinergic denervation with increasing age which leads to hypersensitivity and increased uncoordinated muscular contraction. Diverticular disease seems to be related to the low-fibre diet eaten in developed countries.

Diverticulitis occurs when faeces obstruct the neck of the diverticulum causing stagnation and allowing bacteria to multiply and produce inflammation. This can then lead to bowel perforation (peridiverticulitis), abscess formation, fistulae into adjacent organs, or even generalized peritonitis.

Clinical features and management

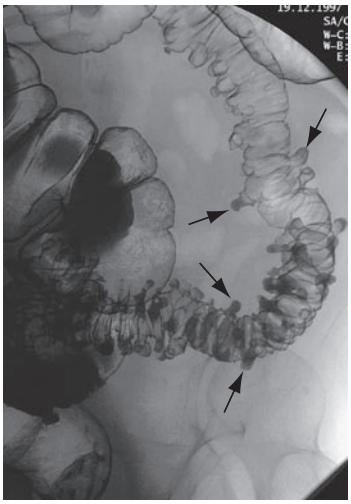
Diverticular disease is asymptomatic in 95% of cases and is usually discovered incidentally on a barium enema examination. No treatment other than dietary advice is required in those patients. In symptomatic patients intermittent left iliac fossa pain or discomfort and an erratic bowel habit commonly occur. In severe disease luminal narrowing can occur in the sigmoid colon, giving rise to severer pain and constipation. In the absence of clinical signs of acute diverticulitis a colonoscopy or 'virtual colonoscopy' (see p. 244) is the investigation of choice. Barium enema (Fig. 6.41) combined with flexible sigmoidoscopy is also used. Technically it is sometimes difficult to obtain adequate views of the sigmoid region in diverticular disease and if this is the case a fibreoptic sigmoidoscopy may be required. Treatment of uncomplicated symptomatic disease is with a well-balanced (soluble and insoluble) fibre diet (20 g/day) with smooth muscle relaxants if required.

Acute diverticulitis

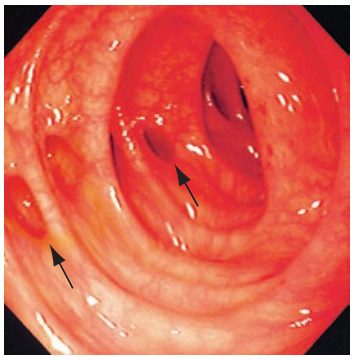
This most commonly affects diverticula in the sigmoid colon. It presents with severe pain in the left iliac fossa, often accompanied by fever and constipation. These symptoms and signs are similar to appendicitis but on the left side. On examination the patient is often febrile with a tachycardia. *Abdominal examination* shows tenderness, guarding and rigidity on the left side of the abdomen. A palpable tender mass is sometimes felt in the left iliac fossa.

Investigations

- **Blood tests**. A polymorphonuclear leucocytosis is often present. The ESR and CRP are raised.
- **Spiral CT of the lower abdomen** (Fig. 6.42) will show colonic wall thickening, diverticula and often pericolic collections and abscesses. There is usually a streaky increased density extending into the immediate pericolic fat with thickening of the pelvic fascial planes. These



(a)



(b)

Fig. 6.41 **Diverticular disease. (a) Double-contrast enema showing diverticula (arrows).** The barium is black on these films. **(b) Diverticula (arrows) seen on colonoscopy.**

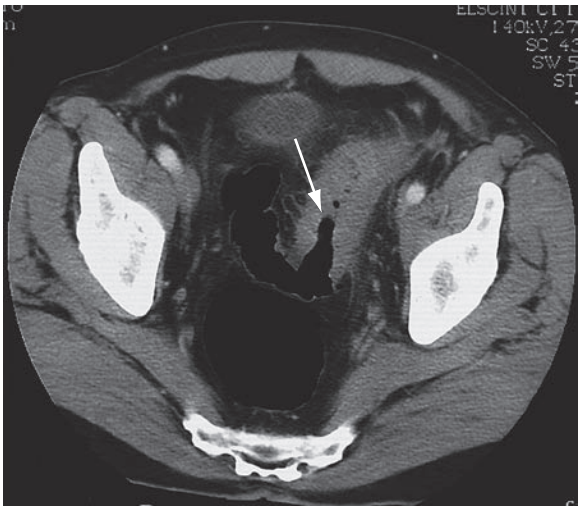


Fig. 6.42 **Spiral CT of lower abdomen, showing acute diverticulitis (arrow).** The bowel wall is thickened and there is loss of clarity of the pericolic fat. A narrow segment of bowel is seen to the left of the diseased segment.

findings are diagnostic of acute diverticulitis (95% sensitivity and specificity) and differ from those of malignant disease. Sigmoidoscopy and colonoscopy are not performed during an acute attack.

Ultrasound examination is often more readily available and is cheaper. It can demonstrate thickened bowel and large pericolic collections, but is less sensitive than CT.

Treatment

Acute attacks can be treated on an outpatient basis using a cephalosporin and metronidazole. Patients who are more ill will require admission for bowel rest, intravenous fluids and i.v. antibiotic therapy (e.g. gentamicin, or a cephalosporin) and metronidazole.

Complications of diverticular disease

- Perforation, which usually, but not always, occurs in association with acute diverticulitis, can lead to formation of a paracolic or pelvic abscess or generalized peritonitis. Surgery may be required.
- Fistula formation into the bladder, causing dysuria or pneumaturia, or into the vagina, causing discharge.
- Intestinal obstruction (see p. 316) usually after repeated episodes of acute diverticulitis.
- Bleeding, which is sometimes massive. In most cases the bleeding stops and the cause of the bleeding can be established by colonoscopy and sometimes angiography. In rare cases emergency segmental colectomy is required.
- Mucosal inflammation in areas of diverticula occurs, giving the appearance of a segmental colitis at endoscopy.

FURTHER READING

Jacobs DO. Diverticulitis. *New England Journal of Medicine* 2007; **357**: 2057–2066.

Stouman N. Diverticular disease. *Lancet* 2004; **363**: 631–639.

ANORECTAL DISORDERS

Pruritus ani

Pruritus ani, or an itchy bottom, is common. Perianal excoriation results from scratching. Usually the condition results from seepage from haemorrhoids or overactivity of sweat glands. *Treatment* consists of salt baths, keeping the area dry with powder; the use of all creams should be avoided. Secondary causes include threadworm (*Enterobius vermicularis*) infestation, fungal infections (e.g. candidiasis) and perianal eczema, which should be treated appropriately.

Haemorrhoids

Haemorrhoids (primary, internal, second degree, prolapsing, third degree prolapsed) usually cause rectal bleeding, discomfort and pruritus ani. Patients may notice red blood on their toilet paper and blood on the outside of their stools. They are the most common cause of rectal bleeding (Fig. 6.24). Diagnosis is made by inspection, rectal examination and proctoscopy. If symptoms are minor no treatment is required; depending on severity of symptoms, treatment is with injection of sclerosant, rubber band ligation or surgery.

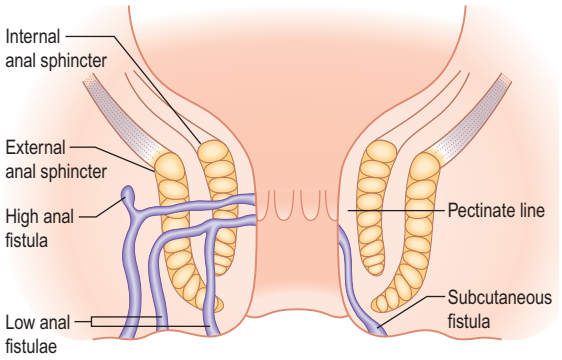


Fig. 6.43 **Common sites of anal fistulae.** Note subcutaneous fistulae do not traverse the sphincters, whereas low and high fistulae do.

Anal fissures

An anal fissure is a tear in the sensitive skin-lined lower anal canal distal to the dentate line which produces pain on defecation. It can be an isolated primary problem in young to middle-aged adults or occur in association with Crohn's disease or ulcerative colitis, in which case perianal abscesses and anal fistulae can complicate the fissure. Diagnosis can usually be made on the history alone and the fissures can be seen usually posteriorly. Rectal examination is often not possible because of pain and sphincter spasm. The spasm not only causes pain but impairs wound healing. In severe cases proctoscopy and sigmoidoscopy should be performed under anaesthesia to exclude other anorectal disease. Initial treatment is with local anaesthetic gel and stool softeners. Use of 0.4% glyceryl trinitrate and 2% diltiazem ointments are of benefit. Botulinum toxin is used in chronic fissures but lateral subcutaneous internal sphincterotomy is the best treatment for severe cases.

Fistula in ano

Subcutaneous, low and high anal fistulae are the commonest types (Fig. 6.43). Anorectal fistulae are rarer forms. The fistulae usually present as abscesses and heal after the abscess is incised. In other cases a small discharging sinus may be noted by the patient. Endoanal ultrasonography, magnetic resonance and/or examination under anaesthetic is usually required to define the primary and any secondary tracks and detect any associated disease. Management is usually surgical with approximately 90% of fistulae being laid open or excised.

Anorectal abscesses

Anorectal abscesses are a common cause of admission to hospital. They are two to three times commoner in men, particularly in men who have sex with men and who indulge in penetrative anal sex. They may be the first manifestation of Crohn's disease, ulcerative colitis and tuberculosis. Perianal and ischiorectal abscess, the commonest forms, present with painful, tender swellings and discharge. Examination under anaesthetic, MRI and endoanal ultrasound are used for evaluation. Treatment is with surgical incision and drainage with antibiotics.

Rectal prolapse, intussusception and solitary rectal ulcer syndrome (SRUS)

All these conditions are thought to be related, with rectal prolapse being the unifying pathology. Some patients with SRUS do not have prolapse but strain excessively and ulcerate the anterior rectal wall, which is forced into the anus during futile straining efforts.

Rectal prolapse starts as an *intussusception* of the upper rectum which passes into the lower part of the viscus, and progression results in its protrusion through the anal canal to emerge as an external prolapse. Constipation and chronic straining are the likely causes. In addition, the presence of an intussuscepting rectum will mimic presence of stool in the rectum and lead to straining efforts to evacuate the prolapse. In some patients repeated straining leads to traumatic ulceration of the mucosa and formation of SRUS. Patients commonly present with slight bleeding and mucus on defecation, tenesmus and sensation of anal obstruction.

SRUS is commonly on the anterior wall of the rectum within 13 cm of the anal verge and this is sometimes difficult to distinguish from cancer and Crohn's disease during endoscopic examination. SRUS has typical histological features of non-specific inflammatory changes with bands of smooth muscle extending into the lamina propria.

Asymptomatic SRUS should not be treated. Symptomatic patients should be advised to stop straining and measures taken to soften the stool. If rectal prolapse can be demonstrated during defecation, this should be repaired; in severe cases surgical treatment by resection rectopexy may be indicated. Surgical treatment for complete rectal prolapse is also required.

COLONIC TUMOURS

Colon polyps and polyposis syndromes

A colonic polyp is an abnormal growth of tissue projecting from the mucous membrane of the colon. They may be millimetres to several centimetres in diameter and may be single or multiple, pedunculated, sessile or 'flat' (Fig. 6.44a,b,c).

Many histological types of polyps are found in the colon (Table 6.17). Adenomas are the precursor lesions in most cases of colon cancer.

Classification of colorectal polyps (Table 6.17)

Neoplastic polyps

Sporadic adenomas An adenoma is a benign tumour of columnar cells or glandular tissue. They may have tubular, tubulovillous or villous morphology. The vast majority of adenomas are not inherited and are termed 'sporadic'. Although many sporadic adenomas do not become malignant in the patient's lifetime, they have a tendency to progress to cancer via increasing grades of dysplasia due to progressive accumulation of genetic changes. Factors favouring malignant transformation in colorectal polyps and the relation between adenoma size and likelihood of cancer are shown in Box 6.10.

The progression from benign polyp to cancer is shown in Figure 6.45 on page 303.

The likelihood of an adenoma being present increases with age; they are rare before the age of 30 years. By the age of 60–70, 5% of asymptomatic subjects will have a polyp of 1 cm or more, or cancer with no symptoms, and up to

Table 6.17 Classification of colorectal polyps and polyposis syndromes

Histology	Polyposis syndrome	Defective gene	Inheritance	CRC risk
Hyperplastic	Hyperplastic polyposis	Unknown		Yes
Hamartoma	Juvenile polyposis	*MADH4 or BMPR1A*	AD	10–70%
	Peutz–Jeghers syndrome	*STK11*	AD	yes
	Cowden syndrome	*PTEN*	AD	10%?
	Lhermitte–Duclos disorder	*PTEN*	AD	
	Bannayan–Riley–Ruvalcaba syndrome	*PTEN*	AD	
Inflammatory	None	None		No
Lymphoid	Benign lymphoid polyposis	Unknown		No
Adenoma	FAP	APC	AD	100%
	AFAP			yes
	Gardner			yes
	Turcot		AR	yes
	MYH-AP			
Adenoma	HNPCC (Lynch type I or II)	Mismatch repair genes (MSH-2, MLH-1)	AD	70–80%

AD, autosomal dominant; AR, autosomal recessive; AFAP, attenuated FAP; FAP, familial adenomatous polyposis; HNPCC, hereditary non-polyposis colorectal cancer; MLH-1, MutL homolog 1; MSH-2, MutS homolog 2; MYH-AP, MUT Y homolog-associated polyposis; PTEN, phosphatase and tensin homolog.

Ask the authors
www.studentconsult.com

Box 6.10 Factors affecting risk of malignant change in an adenoma

	Higher risk	Lower risk
Size	>1.5 cm	<1 cm
Type	Sessile or flat	Pedunculated
Histology	Severe dysplasia	Mild dysplasia
	Villous architecture	Tubular architecture
	Squamous metaplasia	
Number	Multiple polyps	Single polyp

50% will have at least one small < 1 cm adenoma. Removal of polyps at colonoscopy and subsequent surveillance reduces the risk of development of colon cancer by approximately 80%. It is thought that the remaining 20% are either newly formed, or missed, or difficult to detect, e.g. a flat adenoma. Techniques such as chromoscopy using dye spray or narrow band imaging are now being used to assist in their detection (flat adenomas account for approximately 12% of all adenomas).

Polyps in the rectum and sigmoid often present with rectal bleeding. More proximal lesions rarely produce symptoms and most are diagnosed on barium enema, CT colography or on colonoscopy performed for screening for other reasons. Large villous adenomas can present with profuse diarrhoea with mucus and hypokalaemia.

Once a polyp has been found it is almost always possible to remove it endoscopically. Guidelines (British Society of Gastroenterology (BSG) 2002) recommend a repeat colonoscopy

- at 5 years if 1 or 2 adenomas < 1 cm are found,
- at 3 years if there are 3–4 small adenomas or at least one > 1 cm, and
- at 1 year if there are 5 or more small adenomas or there are 3 or more, at least one of which is >1 cm.

If any doubt exists about the completeness of excision of any polyp then an earlier repeat examination is suggested.

Inherited polyposis syndromes

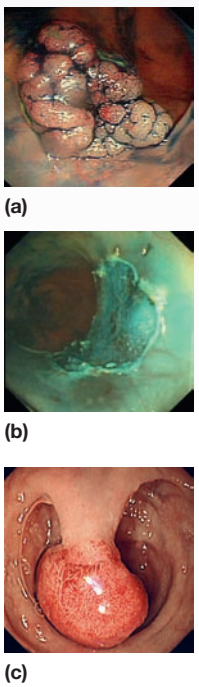
About 5% of colorectal cancers have a well-defined single gene basis.

Familial adenomatous polyposis (FAP) is an autosomal dominant condition arising from germline mutations of the APC gene located on chromosome 5q21-q22. More than 825 different mutations have been identified. Penetrance is virtually 100%. It is characterized by the presence of hundreds to thousands of colorectal adenomas and duodenal adenomas. The mean age of adenoma development is 16 years; the average age for developing colorectal cancer is 39 years. Tracing and screening of relatives is essential, usually after 12 years of age, and affected individuals should be offered a prophylactic colectomy, often before the age of 20. Surgical options include colectomy and ileorectal anastomosis which requires life-long surveillance of the rectal stump, or a restorative proctocolectomy or pouch procedure with complete removal of rectal mucosa.

Cystic gland polyps, predominantly in the proximal stomach, and duodenal adenomas are frequently found in FAP, as well as other extraintestinal lesions such as osteomas, epidermoid cysts and desmoid tumours. The duodenal adenomas may progress to cancer and are the commonest cause of death in colectomized patients with FAP. Congenital hypertrophy of the retinal pigment epithelium (CHRPE) occurs in many families with FAP. Other cancers in FAP include thyroid, pancreatic and hepatoblastomas.

APC gene mutations can be found in about 80% of families with FAP. Once the mutation has been identified in an index case, other family members can be tested for the mutation and screening can then be directed at mutation carriers. If a mutation cannot be found in a known FAP case, all family members should undergo clinical screening with regular colonoscopy.

Attenuated FAP may be missed as it presents later (44 years average age) and has fewer polyps (<100), which tend to occur more on the right side of the colon than on the left. It may be indistinguishable from sporadic cases but the gene mutation is in the APC germ line.



(a)

(b)

(c)

Fig. 6.44
Colonoscopic appearance of **(a) a sessile flat polyp with Indigo carmine; (b) colon after resection of a sessile polyp; (c) a large pedunculated polyp.**

MYH-associated polyposis. MYH (MUT Y Homolog) -associated polyposis is an autosomal recessive inherited syndrome of multiple colorectal adenomas and cancer. MYH is a base-excision-repair gene that corrects oxidative DNA damage. MYH-AP may account for 7–8% of families with the FAP phenotype in whom APC mutations cannot be found. Subjects with multiple adenomas or an FAP phenotype without APC mutations and with a family history compatible with a recessive pattern of inheritance should be tested for MYH-AP.

Hereditary non-polyposis colon cancer (HNPCC; Lynch syndrome). HNPCC is called 'non-polyposis' to distinguish it from FAP, though polyps are formed in the colon and may progress rapidly to colon cancer. It affects 1:5000 people, causing 3–10% of colorectal cancer cases.

The disease is caused by a mutation in one of the DNA mismatch repair genes, usually *hMSH2* or *hMLH1* but others (*hMSH6*, *PMS1* and *PMS2*) are involved. Mismatch repair genes are responsible for maintaining the stability of DNA during replication. Inheritance is autosomal dominant. The defect in function of the mismatch repair mechanism causes naturally occurring highly repeated short DNA sequences known as microsatellites to be shorter or longer than normal, a phenomenon called microsatellite instability (MSI).

Onset of cancer is earlier than in sporadic cases, at age 40–50 or younger. Tumours have a predilection for the right colon, in contrast to sporadic cases. In contrast to FAP, the lifetime risk of colon cancer (penetrance of the gene) in mutation carriers is 70–80%. Other cancers are also more common in HNPCC: stomach, small intestine, bladder, skin, brain and hepatobiliary system. Female patients are at risk for endometrial and ovarian cancer.

The diagnosis is made from the family history of colon cancer at a young age and the presence of associated cancers in the family. These are formalized in the various editions of the Amsterdam and the Bethesda criteria (Table 6.18).

Turcot's syndrome consists of FAP or hereditary non-polyposis colon cancer with brain tumours.

Gardner's syndrome has in addition to FAP desmoid tumours, osteomas of the skull and other lesions.

Non-neoplastic polyps

Hamartomatous polyps are commonly large and stalked. The inherited syndromes show autosomal dominant inheritance and include:

- *Juvenile polyps* which occur mainly in children and teenagers and are found mainly in the colon and histologically show mucus retention cysts. Most are sporadic, but a syndrome of juvenile polyposis is defined as: >3–5 juvenile colonic polyps, juvenile polyps throughout the GI tract, or any number of polyps with a family history. It is an autosomal dominant condition and the relevant gene has been identified (Table 6.17). The polyps are a cause of bleeding and intussusception in the first decade of life. There is also an increased risk of colonic cancer (relative risk (RR) of 34), and surveillance and removal of polyps must be undertaken.
- *Peutz–Jeghers syndrome* which is considered on page 285.
- *PTEN hamartoma-tumour syndrome* (PHTS) which includes Cowden syndrome, Bannayan–Riley–Ruvalcaba syndrome and all syndromes caused by germline PTEN mutations. Cowden (multiple hamartoma) syndrome is associated with characteristic skin stigmata and intestinal polyps regarded as hamartomas but with a mixture of cell types. These patients have an increased risk of various extraintestinal malignancies (thyroid, breast, uterine and ovarian). These syndromes are uncommon and together account for <1% of colon cancer cases.
- *Hyperplastic (metaplastic) polyps.* These are frequently found in the rectum and sigmoid colon. These pale, sessile mucosal nodules usually measure < 5 mm and are normally without significant malignant potential. However, *metaplastic polyposis* is defined as the presence of more than 10 colonic metaplastic polyps, some of which are large. These phenotypes are rare but appear to exhibit an increased risk of colon cancer.

Colorectal carcinoma

Colorectal cancer (CRC) is the third most common cancer world-wide and the second most common cause of cancer death in the UK.

Each year over 30000 new cases are diagnosed in England and Wales (68% colon, 32% rectal cancer) and it is registered as the cause of death in about half this number. The prevalence rate per 100000 (at all ages) is 53.5 for men and 36.7 for women. The incidence increases with age; the average age at diagnosis is 60–65 years. Approximately 20% of patients in the UK have distant metastases at diagnosis. The disease is much more common in westernized countries than in Asia or Africa.

Factors related to risk of colorectal cancer are shown in Table 6.19.

Genetics of colorectal cancer

Most colorectal cancers develop as a result of a stepwise progression from normal mucosa to adenoma to invasive

Table 6.18 Diagnostic criteria for hereditary non-polyposis colon cancer (HNPCC)

Modified Amsterdam Criteria
- One individual diagnosed with colorectal cancer (or extracolonic HNPCC-associated tumours) before age 50 years
- Two affected generations
- Three affected relatives, one a first-degree relative of the other two
- FAP should be excluded
- Tumours should be verified by pathological examination

Bethesda Guidelines
- Colorectal cancer (CRC) diagnosed in patient who is younger than 50 years
- Presence of synchronous, metachronous CRC, or other HNPCC-associated tumours, irrespective of age
- CRC with the MSI-H histology diagnosed in a patient who is younger than 60 years
- CRC diagnosed in one or more first degree relatives with an HNPCC-related tumour, with one of the cancers being diagnosed under the age 50 years
- CRC diagnosed in two or more first or second degree relatives with HNPCC-related tumours, irrespective of age

MSI-H, microsatellite instability – high.

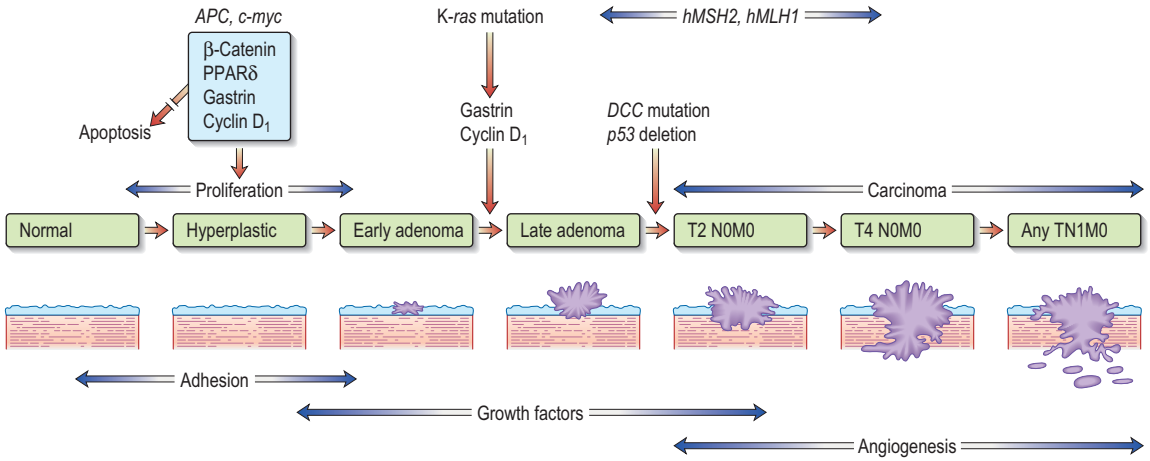


Fig. 6.45 **Genetic model for colorectal tumorigenesis, showing the progression from adenoma to carcinoma.** The stages are shown at which mutations occur in the genes: *APC* (adenomatous polyposis coli), K-*ras*, *DCC*, *p53*, *hMSH2* and *hMLH1*.

Table 6.19 Risk factors in colorectal cancer

Increased risk
- Increasing age
- Low dietary fibre; fibre increases faecal bulk and reduces transit time
- Animal fat (saturated) and red meat consumption
- Sugar consumption
- Colorectal polyps
- Family history of colon cancer
- Chronic inflammatory bowel disease
- Personal history of breast cancer
- Physical inactivity
- Diabetes mellitus
- Obesity (body and abdominal)
- Smoking
- Alcohol excess
- Acromegaly
- Abdominal radiotherapy
- Ureterosigmoidostomy

Decreased risk
- Vegetable, garlic, milk, calcium consumption
- Increased fibre intake in the diet
- Exercise (colon only)
- Aspirin and other NSAIDs
- Combined oestrogen/progesterone hormone-replacement therapy

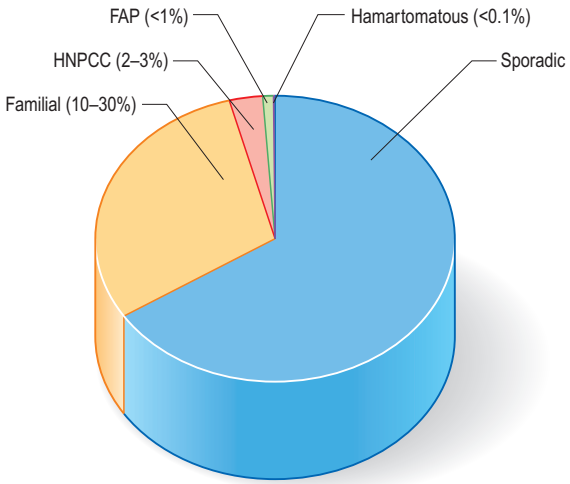


Fig. 6.46 **Percentages of colon cancer according to family risk.** HNPCC, hereditary non-polyposis colorectal cancer; FAP, familial adenomatous polyposis.

cancer. This progression is controlled by the accumulation of abnormalities in a number of critical growth-regulating genes. These include *APC* mutation and loss, K-*ras* mutation, *Smad2/4* loss, and *TP53* mutation and loss, and altered DNA methylation with progression to carcinoma. CDK 8 has recently been found to regulate gene expression in the proliferation of colorectal cancer, and it also regulates the WNT/β-catenin signalling pathway involved in many colon cancers. Microsatellite instability (MSI) and chromosomal instability (CIN) are frequently detected in colon cancers. A third group, MACS (Microsatellite And Chromosomal Stable) is also recognized. MSI has been described (see under HNPCC, p. 302). CIN indicates loss of heterozygosity (LOH) in a number of cancer-related genes, though the underlying mechanisms are not well understood. About 15% of sporadic colorectal cancers show MSI and 50% exhibit LOH.

Cancer families

A family history of CRC confers an increased risk to relatives. Family history is, next to age, the most common risk factor for CRC. FAP (Fig. 6.46) is the best-recognized syndrome predisposing to colorectal cancer but represents less than 1% of all colorectal cancers. Hereditary non-polyposis colorectal cancer (HNPCC) accounts for 3–10% of familial cancer (see p. 302).

Additionally, some colon cancers arise, at least in part, from an inherited predisposition, so-called familial risk (Table 6.20). Estimates of their frequency range from 10% to 30% of all CRC but the genes involved have yet to be identified. The risk of CRC can be estimated from a family history matched with empirical risk tables so that appropriate advice regarding screening can be offered.

Most colorectal cancers are, however, sporadic and occur in individuals without a strong family history. Their distribution is shown in Figure 6.47.

Pathology

CRC, which is usually a polypoid mass with ulceration, spreads by direct infiltration through the bowel wall. It involves lymphatics and blood vessels with subsequent spread, most commonly to the liver and lung. Synchronous

Table 6.20	Lifetime risk of colorectal cancer in first-degree relatives of a patient with colorectal cancer
Population risk	1 in 50
One first-degree relative affected (any age)	1 in 17
One first-degree and one second-degree relative affected	1 in 12
One first-degree relative affected (age < 45)	1 in 10
Two first-degree relatives affected	1 in 6
Autosomal dominant pedigree	1 in 2

Houlston RS, Murday V, Harocopos C et al. Screening and genetic counselling for relatives of patients with colorectal cancer in a family cancer clinic. *British Medical Journal* 1990; **301**: 366–368.

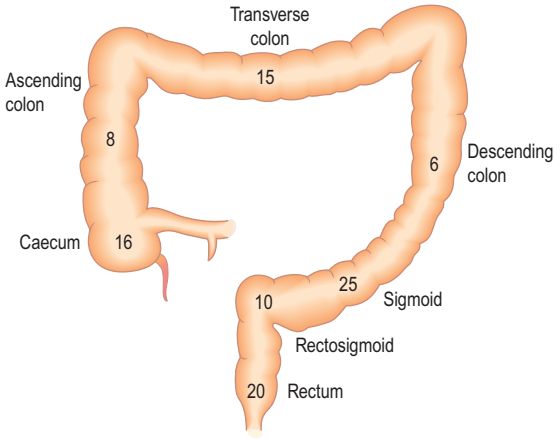


Fig. 6.47 **Distribution of sporadic colorectal cancer.** Only 60% are in the range of flexible sigmoidoscopy.

tumours are present in 2% of cases. Histology is adenocarcinoma with variably differentiated glandular epithelium with mucin production. 'Signet ring' cells in which mucin displaces the nucleus to the side of the cell are relatively uncommon and generally have a poor prognosis.

Clinical features

Symptoms suggestive of colorectal cancer include change in bowel habit with looser and more frequent stools, rectal bleeding, tenesmus and symptoms of anaemia. Looser and more frequent stools, with or without abdominal pain, are common symptoms of left-sided colonic lesions. Rectal and sigmoid cancers often bleed, blood being mixed in with the stool. Presentation with constipation with hard stools is **not** a risk factor for colon cancer. A rectal or abdominal mass may be palpable. Cancers arising in the caecum and right colon are often asymptomatic until they present as an iron deficiency anaemia. Cancer may present with intestinal obstruction.

Patients aged over 35–40 years presenting with new large bowel symptoms should be investigated. Digital examination of the rectum is essential and sigmoidoscopy should be performed in all cases. Hepatomegaly may be found if there are large liver metastases.

Investigation

- *Colonoscopy* (Fig. 6.48) is the gold standard for investigation and allows biopsy for histology. *Biopsy* of the tumour is mandatory, usually at endoscopy.
- *Double-contrast barium enema* can visualize the large bowel but is now being superseded by CT colonography.
- *Endoanal ultrasound and pelvic MRI* are used for staging rectal cancer.
- *Chest, abdominal and pelvic CT* scanning to evaluate tumour size, local spread and liver and lung metastases.
- *PET scanning* is useful for detecting occult metastases and for evaluation of suspicious lesions found on CT or MR.
- *MR* is also useful for evaluating suspicious lesions found on CT or US, especially in the liver.
- *Serum carcinoembryonic antigen (CEA)* is of little use for primary diagnosis, but may be useful for follow-up; rising levels suggest recurrence. It should not be performed as a screening test.
- *Faecal occult blood tests* have been used for mass screening but are of no value in hospital or general practice.

Treatment

Treatment should be undertaken by multidisciplinary teams working in specialist units. About 80% of patients with colorectal cancer undergo surgery, though fewer than half of these survive more than 5 years. The operative procedure depends on the cancer site. Long-term survival relates to the stage of the primary tumour and the presence of metastatic disease (Table 6.21). There has been a gradual move from using Dukes' classification to using the TNM classification system (see p. 456). Long-term survival is only likely when the cancer is completely removed by surgery with adequate clearance margins and regional lymph node clearance.

- *Total mesorectal excision (TME)* is used to carefully remove the entire package of mesorectal tissue surrounding the cancer. A low rectal anastomosis is then performed. Abdomino-perineal excision which requires a permanent colostomy is reserved for very low tumours within 5 cm of the anal margin. TME combined with preoperative radiotherapy reduces local recurrence rates in rectal cancer to around 8% and improves survival. Pre- or postoperative chemotherapy reduces local recurrence rates but had no effect on survival in a recent study.
- *Colon cancer. A segmental resection* and restorative anastomosis with removal of the draining lymph nodes as far as the root of the mesentery is used for cancer elsewhere in the colon. Surgery in patients with obstruction carries greater morbidity and mortality. Where technically possible, preoperative decompression by endoscopic stenting with a mesh-metal stent relieves obstruction so surgery can be elective rather than emergency, and is probably associated with a decrease in morbidity and mortality.
- *Local transanal surgery* is very occasionally used for early superficial rectal cancers.
- *Surgical or ablative treatment of liver and lung metastases* prolongs life where treatment is technically feasible and the patient is fit enough to undergo the treatment.
- *Radiotherapy is not helpful for colonic cancers* proximal to the rectum because of difficulties delivering a

Fig. 6.48 **Colonoscopic appearance of a carcinoma in the ascending colon** – a large irregular ulcer.

Table 6.21 Staging and survival of colorectal cancers

TNM classification			Modified Dukes' classification	5-year survival (%)
Stage I (N0, M0)	Tumours invade submucosa	T1	A	90
	Tumours invade muscularis propria	T2		
Stage IIA (N0, M0)	Tumours invade into subserosa	T3	B	70
IIB	Tumours invade directly into other organs	T4		65
Stage III (M0)	T1, T2 + 1–3 regional lymph nodes involved	N1	C	60
IIIB	T3, T4 + 1–3 regional lymph nodes involved	N1		35
IIIC	Any T + 4 or more regional lymph nodes	N2		25
Stage IV	Any T, any N + distant metastases	M1	D	7

sufficient dose to the tumour without excess toxicity to adjacent structures, particularly the small bowel.

- *Adjuvant postoperative chemotherapy* improves disease-free survival and overall survival in stage III (Dukes' C) colon cancer (see p. 490). Those with Duke's B tumours with advanced features such as vascular invasion may also benefit.

Treatment of advanced colorectal cancer is discussed on page 490.

Follow-up

All patients who have surgery should have a total colonoscopy performed before surgery to look for additional lesions. If total colonoscopy cannot be achieved before surgery, a second 'clearance' colonoscopy within 6 months of surgery is essential. Patients with stage II or III disease should be followed up with regular colonoscopy and CEA measurements; rising levels of CEA suggest recurrence. CT scanning to detect operable liver metastases should be performed for up to 5 years post surgery.

Prevention and screening

- *Diet.* A low-fat, high-fibre diet for the prevention of sporadic colorectal cancer and endoscopic screening is recommended for at-risk patients with a strong family history and for inherited syndromes (e.g. FAP, HNPCC).
- *NSAIDs or aspirin* may play a role in prevention.
- *Faecal occult blood (FOB) tests* have been studied as a screening test for colorectal cancer. Several large randomized studies have demonstrated a reduction in cancer-related mortality of 15–33%. The disadvantage of screening with FOB is its relatively low sensitivity, which means many negative colonoscopies. In FOBT screen-positive patients in the UK National Bowel Cancer Screening Programme (NHS BCSP), about 10% have cancer, 40% have adenomas and the colon is normal in 50%.
- *Colonoscopy* is the gold-standard technique for the examination of the colon and rectum and is the investigation of choice for high-risk patients. Universal screening strategies have been recommended in the USA, but the shortage of skilled endoscopists, the expense, the need for full bowel preparation and the small risk of perforation make colonoscopy impractical as a population screening tool at present.
- *CT colonography* ('virtual colonoscopy') (see Fig. 6.3) is being increasingly used.
- Genetic testing and stool DNA tests also contribute to screening programmes.

FURTHER READING

Johnson CD, Chen M-H, Toledano AY. Accuracy of CT colonography for detection of large adenomas and cancers. *New England Journal of Medicine* 2008; **359**: 1207–1217.

Meyerhardt JA, Mayer RJ. Systemic therapy for colorectal cancer. *New England Journal of Medicine* 2005; **352**: 476–487.

Weitz J, Koch M, Debus J et al. Colorectal cancer. *Lancet* 2005; **365**: 153–165.

DIARRHOEA

Diarrhoea is a common clinical problem and there is no uniformly accepted definition of diarrhoea. Organic causes (stool weights > 250 g per day) have to be distinguished from 'functional causes', and is the first step in the assessment of the history. Sudden onset of bowel frequency associated with crampy abdominal pains, and a fever will point to an infective cause; bowel frequency with loose blood-stained stools to an inflammatory basis; and the passage of pale offensive stools that float, often accompanied by loss of appetite and weight loss, to steatorrhoea. Nocturnal bowel frequency and urgency usually point to an organic cause. Passage of frequent small-volume stools (often formed) points to a functional cause (see Functional gastrointestinal disorders, p. 309).

Mechanisms

Osmotic diarrhoea

The gut mucosa acts as a semi-permeable membrane and fluid enters the bowel if there are large quantities of non-absorbed hypertonic substances in the lumen. This occurs because:

- the patient has ingested a non-absorbable substance (e.g. a purgative such as magnesium sulphate or magnesium-containing antacid)
- the patient has generalized malabsorption so that high concentrations of solute (e.g. glucose) remain in the lumen
- the patient has a specific absorptive defect (e.g. disaccharidase deficiency or glucose–galactose malabsorption).

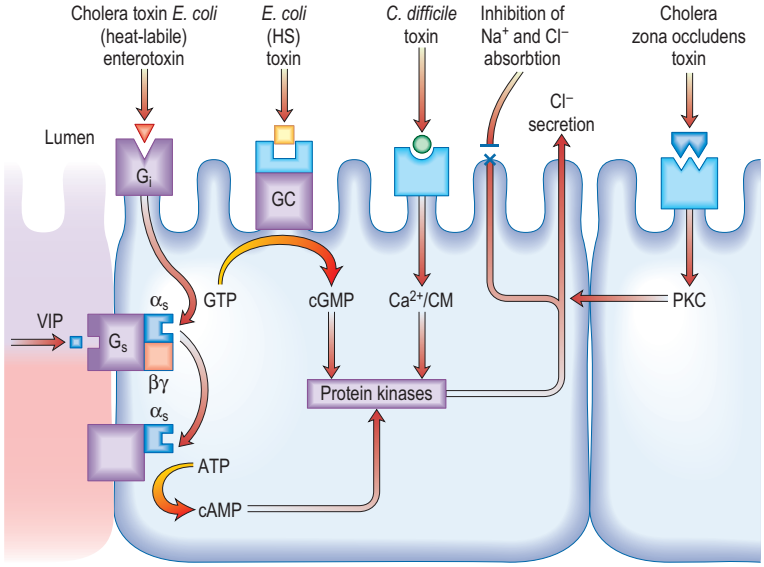
The volume of diarrhoea produced by these mechanisms is reduced by the absorption of fluid by the ileum and colon. The diarrhoea stops when the patient stops eating or the malabsorptive substance is discontinued.

Secretory diarrhoea

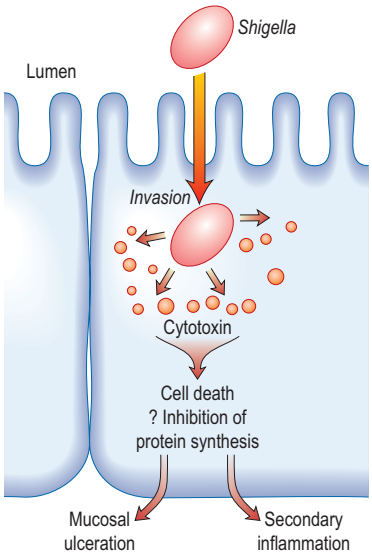
In this disorder, there is both active intestinal secretion of fluid and electrolytes as well as decreased absorption. The mechanism of intestinal secretion is shown in Figure 6.49a.

Common causes of secretory diarrhoea are:

- enterotoxins (e.g. cholera, *E. coli* – thermolabile or thermostable toxin, *C. difficile*)
- hormones (e.g. vasoactive intestinal peptide in the Verner–Morrison syndrome, p. 385)
- bile salts (in the colon) following ileal resection
- fatty acids (in the colon) following ileal resection
- some laxatives (e.g. docusate sodium).



(a) Enterocyte



(b) Colonic mucosal cell

Fig. 6.49 **Mechanisms of diarrhoea. (a) Small intestinal secretion of water and electrolytes.**

- *Cholera toxin* binds to its receptor (monosialoganglioside G_i) via fimbria (toxin co-regulated pilus) on its β subunit. This activates the α_s subunit (of the G_s protein), which in turn dissociates and activates cyclic AMP (cAMP). The increase in cAMP activates intermediates (e.g. protein kinase and Ca^{2+}) which then act on the apical membrane causing Cl^- secretion (with water) and inhibition of Na^+ and Cl^- absorption.
- *E. coli*. Heat-labile *E. coli* enterotoxin shares a receptor with cholera toxin. Heat-stable (HS) *E. coli* toxin binds to its own receptor and activates guanylate cyclase (cGMP), producing the same effect on secretion.
- *C. difficile* activates the protein kinases via Ca^{2+}/calmodulin (Ca^{2+}/CM).
- *Zona occludens toxin*. This toxin is the product of the *ZOT* gene, which is the gene required for the *CTX* gene which encodes for cholera toxin. It has enterotoxic activity, producing secretion.

  Cholera and *E. coli* cause these effects without invasion of the cell.

  G, G protein consisting of subunits α, β, γ; i, inhibitory; s, stimulatory; ATP, adenosine triphosphate; GTP, guanosine triphosphate; GMP, guanosine monophosphate; GC, guanyl cyclase; PKC, protein kinase C; VIP, vasoactive intestinal polypeptide.

(b) Colonic mucosal cell. This demonstrates one of the mechanisms by which an invasive pathogen (e.g. *Shigella*) acts. Following penetration, the pathogens generate cytotoxins which lead to mucosal ulceration and cell death.

Inflammatory diarrhoea (mucosal destruction)

Diarrhoea occurs because of damage to the intestinal mucosal cell so that there is a loss of fluid and blood (Fig. 6.49b). In addition, there is defective absorption of fluid and electrolytes. Common causes are infective conditions (e.g. dysentery due to *Shigella*), and inflammatory conditions (e.g. ulcerative colitis and Crohn's disease).

Abnormal motility

Diabetic, post-vagotomy and hyperthyroid diarrhoea are all due to abnormal motility of the upper gut. In many of these cases the volume and weight of the stool is not all that high, but frequency of defecation occurs; this therefore may not be true diarrhoea.

Causes of diarrhoea are shown in Table 6.22. It should be noted that the irritable bowel syndrome, colorectal cancer, diverticular disease and faecal impaction with overflow in the elderly do not cause 'true' organic diarrhoea (i.e. >250 g/day), even though the patients may complain of diarrhoea.

Table 6.22 Causes of diarrhoea

Infective causes
Bacterial, e.g.
- *Campylobacter jejuni*
- *Salmonella* spp.
- *Shigella*
- *Escherichia coli* (p. 130)
- Staphylococcal enterocolitis
- *Bacillus cereus*
- *Clostridium perfringens, botulinum, difficile*
- gastrointestinal tuberculosis

Viral, e.g.
- rotavirus

Fungal, e.g.
- histoplasmosis

Parasitic, e.g.
- amoebic dysentery (*Entamoeba histolytica*)
- schistosomiasis
- *Giardia intestinalis*

Non-infective causes of diarrhoea
Inflammatory bowel disease
Radiation proctitis or colitis
Behçet's disease
Diverticular disease
Ischaemic colitis
Gastrointestinal lymphoma
Carcinoma of the colon (change in bowel habit)
Malabsorption
Gut resection
Bile acid malabsorption
Drugs – many, including
- laxatives
- metformin
- anticancer drugs
- statins

Faecal impaction with overflow
Irritable bowel syndrome and functional diarrhoea

Endocrine
Zollinger–Ellison syndrome
Vipoma
Somatostatinoma
Glucagonoma
Carcinoid syndrome
Thyrotoxicosis
Medullary carcinoma of thyroid
Diabetic autonomic neuropathy

Factitious diarrhoea (4%)
Purgative abuse
Dilutional diarrhoea

World-wide, infection and infestation are a major problem and these are discussed under the causative organisms in Chapter 4.

Acute diarrhoea (excluding cholera, discussed on p. 143)

Diarrhoea of sudden onset is very common, often short-lived and requires no investigation or treatment. This type of diarrhoea is seen after dietary indiscretions, but diarrhoea due to viral agents also lasts 24–48 hours (see p. 113). The causes of other infective diarrhoeas are shown on page 129. Travellers' diarrhoea, which affects people travelling outside their own countries, particularly to developing countries, usually lasts 2–5 days; it is discussed on page 131. Clinical features associated with the acute diarrhoeas include fever, abdominal pain and vomiting. If the diarrhoea is particularly severe, dehydration can be a problem; the very young and very old are at special risk from this. Investigations are necessary if the diarrhoea has lasted more than 1 week. Stools (up to three) should be sent immediately to the laboratory for culture and examination for ova, cysts and parasites. If the diagnosis has still not been made, a sigmoidoscopy and rectal biopsy should be performed and imaging should be considered.

Oral fluid and electrolyte replacement is often necessary. Special oral rehydration solutions (e.g. sodium chloride and glucose powder) are available for use in severe episodes of diarrhoea, particularly in infants. Antidiarrhoeal drugs are thought to impair the clearance of any pathogen from the bowel but may be necessary for short-term relief (e.g. codeine phosphate 30 mg four times daily, or loperamide 2 mg three times daily). Antibiotics are sometimes given (see p. 132) depending on the organism.

More online
www.studentconsult.com

Chronic diarrhoea

This always needs investigation. All patients should have a sigmoidoscopy and rectal biopsy. The flow diagram in Figure 6.50 is illustrative; whether the large or the small bowel is investigated first will depend on the clinical story of, for example, bloody diarrhoea or steatorrhoea. The investigations and treatment are described in detail under the individual diseases.

When difficulties exist in distinguishing between functional and organic causes of diarrhoea, hospital admission for a formal 72-hour assessment of stool weights is helpful and will also assist in the diagnosis of factitious causes of diarrhoea.

Antibiotic-associated diarrhoea (pseudomembranous colitis) (see p. 131)

Pseudomembranous colitis may develop following the use of any antibiotic. Diarrhoea occurs in the first few days after taking the antibiotic or even up to 6 weeks after stopping the drug. The causative agent is *Clostridium difficile* (see p. 131).

Bile acid malabsorption

Bile acid malabsorption is an underdiagnosed cause of chronic diarrhoea and many patients with this disorder are assumed to have irritable bowel syndrome. Bile acid diarrhoea occurs when the terminal ileum fails to reabsorb bile

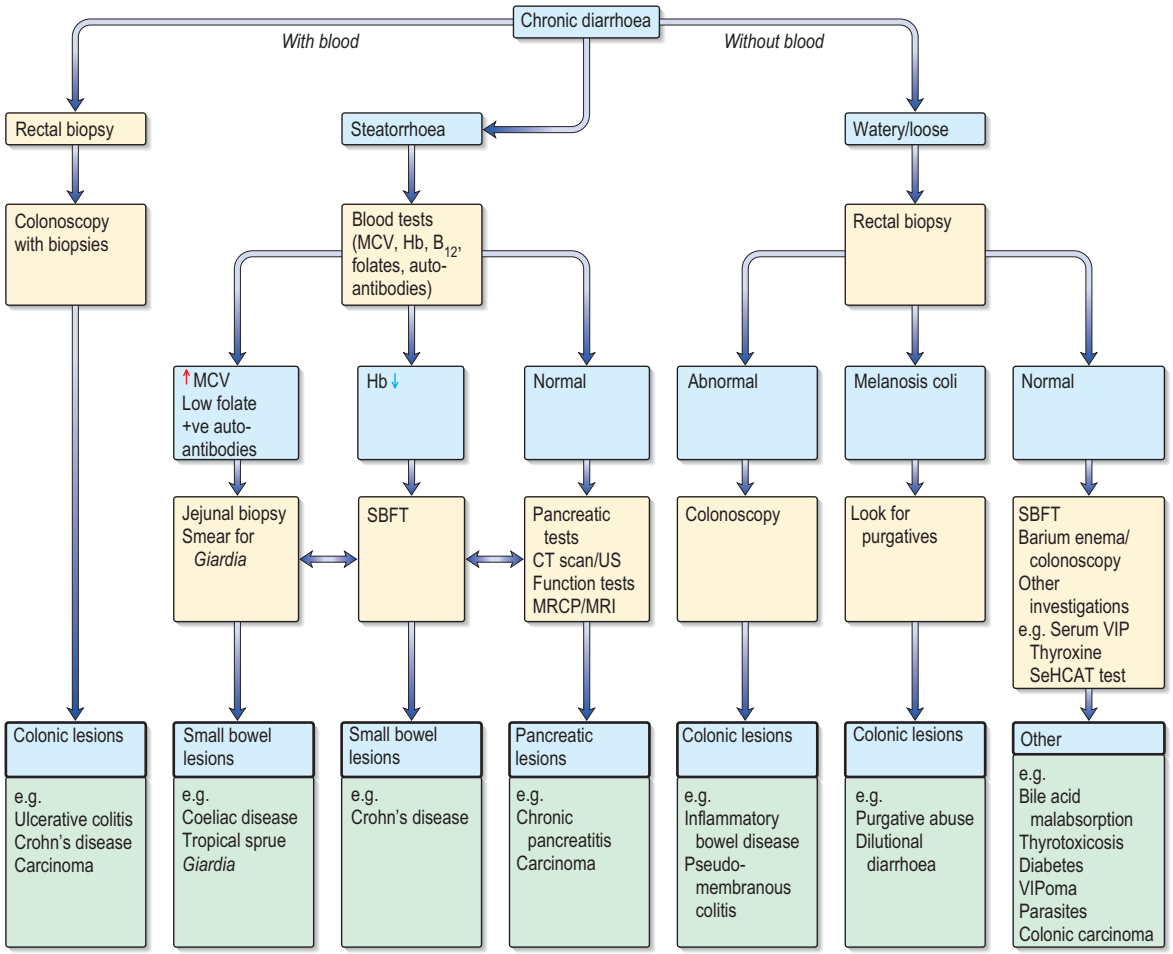


Fig. 6.50 **Flow diagram for the investigation of chronic diarrhoea.** NB: All patients should have had stool cultures. SBFT, small bowel follow-through; VIP, vasoactive intestinal polypeptide; MRCP, magnetic resonance cholangiopancreatography; SeHCAT, ^{75}Se-homochoyl taurine.

acids. Bile acids (particularly the dihydroxy bile acids – deoxycholate and chenodeoxycholate) when present in increased concentrations in the colon lead to diarrhoea by reducing absorption of water and electrolytes and, at higher concentrations, inducing secretion as well as increasing colonic motility. A variety of causes of bile acid malabsorption are recognized (Table 6.23).

Bile acid malabsorption should be considered not only in patients with chronic diarrhoea of unknown cause but also in patients with diarrhoea and associated disease who are not responding to standard therapy (e.g. patients with terminal ileal Crohn's disease, microscopic inflammatory colitis).

Diagnosis is made using the SeHCAT test in which a radiolabelled bile acid analogue is administered and percentage retention at 7 days calculated (<19% retention abnormal). Treatment is with colestyramine, a resin which binds and inactivates the action of bile acids in the colon. The best results of treatment are obtained in patients with a SeHCAT retention of <5%.

Table 6.23 **Causes of bile acid malabsorption**

Ileal resection
Ileal disease, e.g. active or inactive Crohn's disease
Idiopathic or primary bile acid malabsorption (structurally normal ileum)
Postinfective gastroenteritis
Associated with:
- rapid small bowel transit, e.g. functional diarrhoea
- post-cholecystectomy diarrhoea
- diabetic diarrhoea
- post-vagotomy diarrhoea
- chronic pancreatitis
- cystic fibrosis
- coeliac disease
- microscopic inflammatory colitis
- drugs (e.g. colchicine, metformin)

Factitious diarrhoea

Factitious diarrhoea accounts for up to 4% of new patients with diarrhoea attending gastroenterology clinics.

Purgative abuse

This is most commonly seen in females who surreptitiously take high-dose purgatives and are often extensively investigated for chronic diarrhoea. The diarrhoea is usually of high volume (>1 L daily) and patients may have a low serum potassium. Sigmoidoscopy may show pigmented mucosa, a condition known as melanosis coli. Histologically the rectal biopsy shows pigment-laden macrophages in patients taking

Table 6.24 Gastrointestinal problems in patients with AIDS

Symptoms	Causes
Mouth/oesophagus	
Dysphagia	Herpes simplex virus (HSV)
Candidiasis	Cytomegalovirus (CMV)
Retrosternal discomfort	
Oral ulceration	
Small bowel/colon	
Chronic diarrhoea	Parasites:
Steatorrhoea	*Entamoeba histolytica*
Weight loss	*Giardia intestinalis*
	Cryptosporidium
	Blastocystis hominis
	Isospora belli
	Microsporidia
	Cyclospora cayetanensis
	Viruses:
	CMV/HSV, adenovirus
	Bacteria:
	Salmonella
	Campylobacter
	Shigella
	Mycobacterium avium-intracellulare
	Non-infective enteropathy – cause unknown
Rectum/colon	
Bloody diarrhoea	Bacterial infection (e.g. *Shigella*)
Any site	
Weight loss	Neoplasia:
Diarrhoea	Kaposi's sarcoma
	Lymphoma
	Squamous carcinoma
	Infection – disseminated, e.g. *Mycobacterium avium-intracellulare*
	HAART therapy

an anthraquinone purgative (e.g. senna). Melanosis coli is also seen in people regularly taking purgatives in normal doses.

In advanced cases a barium enema may show a dilated colon and loss of haustral pattern.

Phenolphthalein laxatives can be detected by pouring an alkali (e.g. sodium hydroxide) on the stools, which then turn pink; a magnesium-containing purgative will give a high faecal magnesium content. Anthraquinones can also be measured in the urine. If the diagnosis is suspected, a locker or bed search (while the patient is out of the ward) is occasionally necessary. Management is difficult as most patients deny purgative ingestion. Purgative abuse often occurs in association with eating disorders and all patients needs psychiatric help. It is sometimes safer not to confront the patient with their diagnosis.

Dilutional diarrhoea

In this condition raised stool weights occur as a consequence of patients deliberately diluting their stool with urine or tap water. The diagnosis is made by measuring stool osmolality and electrolyte concentrations in order to calculate the faecal osmolar gap. Measurement of stool creatinine content is helpful in excluding dilution of stools with urine, and early admission to hospital for formal assessment of stool weights may avoid unnecessary invasive investigations being carried out.

Diarrhoea in patients with HIV infection

Chronic diarrhoea is a common symptom in HIV infection, but HIV's role in the pathogenesis of diarrhoea is unclear. *Cryptosporidium* (see p. 142) is the pathogen most commonly isolated. *Isospora belli* and microsporidia have also been found.

The cause of the diarrhoea is often not found and treatment is symptomatic. Table 6.24 shows the conditions affecting the gastrointestinal tract in patients with AIDS.

FUNCTIONAL GASTROINTESTINAL DISORDERS

There is a large group of gastrointestinal disorders that are termed 'functional' because symptoms occur in the absence of any demonstrable abnormalities in the digestion and absorption of nutrients, fluid and electrolytes and no structural abnormality can be identified in the gastrointestinal tract.

Table 6.25 lists some of the symptoms that are suggestive of a functional gastrointestinal disorder. Modern classification systems are based on the premise that for each disorder there is a symptom cluster that 'breeds true' across clinical and population groups. There is inevitably overlap, with some symptoms being common to more than one disorder.

Table 6.26 lists the common functional gastrointestinal disorders as defined by Rome III criteria. These conditions are extremely common world-wide, making up to 80% of patients seen in the gastroenterology clinic.

Table 6.25 Chronic gastrointestinal symptoms suggestive of a functional gastrointestinal disorder (FGID)

Nausea alone
Vomiting alone
Belching
Chest pain unrelated to exercise
Postprandial fullness
Abdominal bloating
Abdominal discomfort/pain (right or left iliac fossa)
Passage of mucus per rectum
Frequent bowel actions with urgency first thing in morning

FURTHER READING

Thapar N, Sanderson IR. Diarrhoea in children: an interface between developing and developed countries. *Lancet* 2004; **363**: 641–653.

Thomas PD, Forbes A, Green J et al. Guidelines for the investigation of chronic diarrhoea, 2nd edition. *Gut* 2003; **52**(suppl): 917–1074.

Table 6.26 Modified Rome III functional gastrointestinal disorders

A Functional oesophageal disorders
- Heartburn
- Chest pain of presumed oesophageal origin
- Dysphagia
- Globus

B Functional gastroduodenal disorders
- Non-ulcer dyspepsia
- Belching disorders
- Nausea and vomiting disorders
- Rumination syndrome in adults

C Functional bowel disorders
- Irritable bowel syndrome
- Functional bloating
- Functional constipation
- Functional diarrhoea
- Unspecified functional bowel disorder

D Functional abdominal pain syndrome

E Functional gallbladder and sphincter of Oddi (SO) disorders

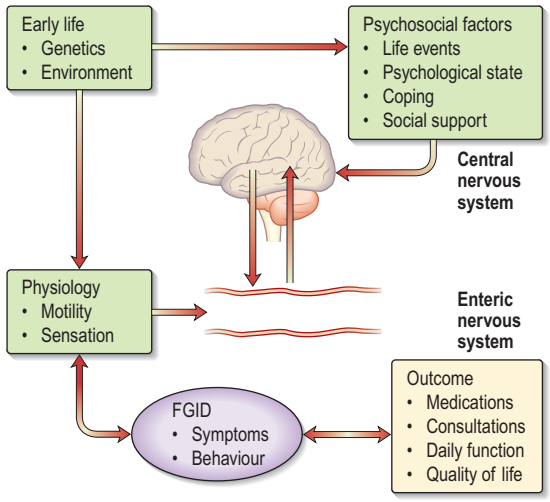


Fig. 6.51 **A biopsychosocial conceptualization of the pathogenesis and clinical expression of functional gastrointestinal disorders (FGIDs),** showing how genetic, environmental and psychological factors may interact to cause dysregulation of brain–gut function. Modified from Drossman DA. The functional gastrointestinal disorders and the Rome III process. *Gastroenterology* 2006; **130**: 1377–1390, with permission from Elsevier.

Pathophysiology and brain–gut interactions

People with functional gastrointestinal disorders (FGID) are characterized by having a greater gastrointestinal motility response to life events than normal subjects. There is, however, a poor association between measured gastrointestinal motility changes and symptoms in many of the FGIDs. Patients with FGID have been shown to have abnormalities in visceral sensation and have a lower pain threshold when tested with balloon distension (visceral hyperalgesia). Visceral hypersensitivity possibly relates to:

- altered receptor sensitivity at the viscus itself
- increased excitability of the spinal cord dorsal horn neurones
- altered central modulation of sensations.

Symptoms are likely to be generated as a consequence of disturbed gastrointestinal motility that leads to distension with visceral hyperalgesia accentuating the pain. In a group of patients who developed IBS symptoms following severe acute enteric infection, inflammatory changes were demonstrated in the enteric mucosa with serotonin-producing enteroendocrine cell hyperplasia, the excessive release of serotonin being thought by some to contribute to symptom development.

The brain–gut axis describes a combination of intestinal motor, sensory and CNS activities (Fig. 6.51). Thus extrinsic (e.g. vision, smell) and intrinsic (e.g. emotion, thought) information can affect gastrointestinal sensation because of the neural connections from higher centres. Conversely, viscerotropic events can affect central pain perception, mood and behaviour.

Psychological stress can exacerbate gastrointestinal symptoms, and psychological disturbances are more common in patients with FGIDs. These disturbances alter attitude to illness, promote healthcare seeking, and often lead to a poor clinical outcome. They have psychosocial consequences with poor quality of life at home and work. Early in life, genetic and environmental influences (e.g. family attitudes towards bowel training, verbal or sexual abuse, exposure to an infection) may affect psychosocial development (susceptibility to life stress, psychological state, coping skills, development of social support) or the development of gut dysfunction (abnormal motility or visceral hypersensitivity).

FGID should be regarded as a dysregulation of brain–gut function.

FUNCTIONAL OESOPHAGEAL DISORDERS

The criteria for diagnosis rest mainly on compatible symptoms. However, *pathological gastro-oesophageal* reflux and other causes may need full investigation, particularly in the elderly with a short history.

Globus

This presents as:

- Persistent or intermittent sensation of a lump or foreign body in the throat.
- Occurrence of the sensation between meals.
- The absence of dysphagia and pain on swallowing (odynophagia).

Treatment is with explanation and reassurance and a trial of antireflux therapy. Antidepressants may be tried.

Functional chest pain, of presumed oesophageal origin

This is characterized by episodes of mainly midline chest pain, not burning in nature, that are potentially of oesophageal origin and which occur in the absence of a cardiological cause, gastro-oesophageal reflux and achalasia.

More than half of patients will respond to high-dose acid-suppression therapy in the first week; some will respond to nitrates and calcium-channel blockers.

A history of psychiatric disorder is found in more than 60% of patients and antidepressant therapy, e.g. amitripty-

line or the selective serotonin reuptake inhibitor citalopram, have been shown to be effective.

FUNCTIONAL GASTRODUODENAL DISORDERS

Functional dyspepsia

This is the second most common functional gastrointestinal disorder (after irritable bowel syndrome). Patients can present with a spectrum of symptoms including upper abdominal pain/discomfort, fullness, early satiety, bloating and nausea.

These patients have no structural abnormality as an explanation for their symptoms.

Functional dyspepsia subgroups

Two subgroups based on the predominant (or most bothersome) single symptoms are suggested:

- *Epigastric pain syndrome* with pain centred in the upper abdomen as the predominant (most bothersome) symptom.
- *Postprandial distress syndrome* with an unpleasant or troublesome non-painful sensation (discomfort) centred in the upper abdomen being the predominant symptom. This sensation may be associated with upper abdominal fullness, early satiety, bloating and nausea.

There is considerable overlap between these two groups.

Investigations

Many young patients (<50) require no investigation. Older patients or those with alarm symptoms require endoscopy. Gastroscopy often shows gastritis but whether this is the cause of the symptoms is doubtful.

Treatment

The range of therapies prescribed for functional dyspepsia reflects the uncertain pathogenesis and the lack of satisfactory treatment options. Management is further confounded by high placebo response rates (20–60%). A proportion of patients will respond satisfactorily to reassurance, explanations and lifestyle changes. Proton pump inhibitors and prokinetic agents may be used for patients with epigastric pain syndrome and postprandial distress syndrome respectively. Reducing intake of fat, coffee, alcohol and cigarette smoking may help.

H. pylori eradication therapy has been shown to be effective in some patients with functional dyspepsia.

Aerophagia

Aerophagia refers to a repetitive pattern of swallowing or ingesting air and belching. It is usually an unconscious act unrelated to meals. Usually no investigation is required. Explanation that the symptoms are due to swallowed air and reassurance are necessary, as is treatment of associated psychiatric disease.

Functional vomiting

Functional vomiting is a rare condition in clinical practice, although chronic nausea is a frequent accompaniment in all functional gastrointestinal disorders. Clinically functional vomiting is characterized by:

- frequent episodes of vomiting, occurring on at least 1 day a week
- absence of criteria for an eating disorder, rumination or major psychiatric disease
- absence of self-induced and medication-induced vomiting
- absence of abnormalities in the gut or central nervous system and metabolic disease to explain the recurrent vomiting.

Investigation is often not required but always exclude non-GI disorders (see Table 6.1).

Treatment is with anti-nausea drugs and antidepressants; behavioural therapy and psychotherapy are helpful. Dietary changes occasionally help.

Rumination syndrome

This presents as:

- Persistent or recurrent effortless regurgitation of recently ingested food into the mouth with subsequent re-mastication and re-swallowing. Regurgitation is not preceded by retching.
- Absence of nausea and vomiting, abdominal discomfort, heartburn.
- Cessation of the process when the regurgitated material becomes acidic. The regurgitant contains recognizable food with a pleasant taste.

Central factors contribute significantly to the occurrence of rumination, and the disorder is common in individuals with learning difficulties. A range of medical, behavioural and nutritional approaches have been attempted in these patients, with varying success.

FUNCTIONAL BOWEL DISORDERS

Irritable bowel syndrome (IBS)

IBS is the commonest FGID. In western populations, up to one in five people report symptoms consistent with IBS. Only approximately 50% will consult their doctors and of these up to 30% will be referred by their GP to a hospital specialist. Up to 40% of all patients seen in specialist gastroenterology clinics will have IBS. Estimates in the UK put the annual cost of IBS to healthcare resources as £45.6 million; in the US the cost is higher at $8 billion. In the UK approximately a quarter of IBS patients lose time off work for periods ranging from 7 to 13 days each year.

The factors that determine whether an IBS sufferer in the community seeks medical advice include the demonstration that consulters have higher illness attitude scores and higher anxiety and depression scores than non-consulters. Consulters perceive that their symptoms are severer than non-consulters, and consulting behaviour may be determined by the number of presenting symptoms. Female consulters outnumber male consulters by a factor of 2–3. Reasons for this include the fact that anxiety and depression scores are higher in women than in men and the gut may be more sensitive to various stimuli in women. It is likely that men and women perceive internal events in the abdomen differently

and that women may be more focused on these events. Food and eating are of more special psychological significance for women, as evidenced by a much higher incidence of eating disorders in women. The whole pelvic region carries a more specific significance for women, being associated not only with defecation, urination and sexuality but additionally with menstruation, pregnancy and childbirth. Finally, women in western society in general seem more willing than men to seek medical attention for a whole variety of disorders.

IBS – a multisystem disorder

IBS patients suffer from a number of non-intestinal symptoms (Table 6.27). The non-intestinal symptoms of IBS can be more intrusive than the classical features of IBS. IBS coexists with chronic fatigue syndrome (see p. 1192), fibromyalgia (see p. 516) and temporomandibular joint dysfunction.

The biopsychosocial conceptualization of the pathogenesis and clinical expression of FGIDs (Fig. 6.51) is particularly relevant to IBS and Box 6.11 lists some common factors that have been shown to trigger IBS symptoms. Infectious diarrhoea precedes the onset of IBS symptoms in 7–30% of patients. Whether this is a factor for all patients or just a small subgroup remains controversial. Risk factors in these patients have been shown to include female gender, severity and duration of diarrhoea, pre-existing life events and high hypochondriacal anxiety and neurotic scores at the time of the initial illness. Symptoms of anxiety and depression are more common in IBS patients and stress or adverse life events often precede the onset of chronic bowel symptoms.

Diagnostic criteria (Rome III 2006)

These criteria state that in the preceding 3 months there should be at least 3 days per month of recurrent abdominal pain or discomfort associated with two or more of the following:

1 improvement with defecation
2 onset associated with a change in frequency of stool
3 onset associated with a change in form (appearance) of stool.

These are useful for comparative studies.

Subgroups of IBS patients can be identified according to the criteria listed in Table 6.28.

The decision as to whether to investigate and if so what choice of investigations is required should be based on clinical judgement. Pointers to the need for thorough investigation are the presence of the above symptoms in association with rectal bleeding, nocturnal pain, fever and weight loss and a clinical suspicion of organic diarrhoea (stool wt > 250 g/day).

Table 6.27 Non-gastrointestinal features of irritable bowel syndrome

Gynaecological symptoms
Painful periods (dysmenorrhoea)
Pain following sexual intercourse (dyspareunia)
Premenstrual tension

Urinary symptoms
Frequency
Urgency
Passing urine at night (nocturia)
Incomplete emptying of bladder

Other symptoms
Back pain
Headaches
Bad breath, unpleasant taste in the mouth
Poor sleeping
Fatigue

Box 6.11 Some factors that can trigger onset of irritable bowel symptoms

- Affective disorders, e.g. depression, anxiety
- Psychological stress and trauma
- Gastrointestinal infection
- Antibiotic therapy
- Sexual, physical, verbal abuse
- Pelvic surgery
- Eating disorders

Table 6.28 Subtyping irritable bowel syndrome by predominant stool pattern

IBS with constipation (IBS-C)	Hard lumpy stools ≥ 25% and loose (mushy) or watery stools < 25% of bowel movements
IBS with diarrhoea (IBS-D)	Loose (mushy) or watery stools ≥ 25% and hard or lumpy stools < 25% of bowel movements
Mixed IBS (IBS-M)	Hard or lumpy stools ≥ 25% and loose (mushy) or watery stools ≥ 25% of bowel movements
Unsubtyped IBS	Insufficient abnormality of stool consistency to meet criteria for IBS-C, D or M

Treatment

Current strategies for treatment of IBS are based on the biopsychosocial conceptualization of IBS (Fig. 6.51) with targeting of central and end-organ therapies (Box 6.12). End-organ and central approaches to treatment should not be mutually exclusive and can be used in sequence and in combinations. The initial promise of the emerging receptor active drugs (HT_3 receptor antagonists for diarrhoea-predominant IBS, HT_4 receptor agonists for constipation-predominant IBS) have not been realized as to date the new drugs have either not received regulatory approval or have been voluntarily withdrawn by the manufacturers. Kappa opioid agonists for use in patients in whom visceral hyperalgesia plays a predominant role in the pathogenesis of their symptoms may still become available.

Probiotics and prebiotics

Probiotics are live or attenuated bacteria or bacterial products that confer a significant health benefit to the host. Two recent studies have shown a beneficial effect of *Bifidobacterium infantis* 35624 on IBS symptoms and in one the symptomatic response was associated with normalization of the ratio of an anti-inflammatory to a proinflammatory cytokine suggesting an immune modulating role for this organism in IBS.

Problems with quality control and formulation of probiotics are currently restricting the clinical availability of some probiotics.

Prebiotics are non-digestible food supplements that are fermented by host bacteria thereby altering the microbiota of

Box 6.12 Approaches to management of the irritable bowel syndrome (IBS)

End organ treatment	Action
Explore dietary triggers	Refer to dietician
High-fibre diet ± fibre supplements for constipation	Refer to dietician ± prescribe ispaghula husk
Antidiarrhoeal drugs for bowel frequency	Loperamide Codeine phosphate Co-phenotrope
Smooth muscle relaxants for pain	Mebeverine hydrochloride Dicycloverine hydrochloride Peppermint oil
Central treatment	**Action**
Explain physiology and symptoms	At consultation (leaflets with diagrams help)
Psychotherapy	Refer to clinical psychologist (see p. 1192)
Hypnotherapy	
Cognitive behavioural therapy	Refer to psychiatrist
Antidepressants	Clomipramine in functional diarrhoea; tricyclic group, e.g. amitriptyline in diarrhoea-predominant IBS; selective serotonin reuptake inhibitors in constipation-predominant IBS, e.g. paroxetine

the host often by stimulating the growth of healthy bacteria. A trans-galacto oligosaccharide prebiotic has now been shown to be bifidobacteria enhancing in IBS patients and to alleviate symptoms. This group of compounds may have considerable potential as therapeutic agents in IBS.

Pain/gas/bloat syndrome/midgut dysmotility

There exist a group of patients with functional bowel disease whose abdominal pain and other clinical features are likely to occur as a consequence of disordered motility and visceral sensation that predominantly affects the small intestine or midgut. The symptom-based diagnostic criteria are abdominal pain, often exacerbated by eating and not relieved by opening the bowels and not associated with the passage of more frequent or looser stools than normal and not associated with constipation. Other symptoms include abdominal distension (bloating). Abdominal distension that is not restricted to the upper abdomen occurs, as well as postprandial fullness, nausea and, on occasions, anorexia and weight loss.

Some patients with pain/gas/bloat syndrome have particularly severe and chronic symptoms, that may also be nocturnal. A subgroup of these have been shown to have manometric features consistent with a diagnosis of chronic idiopathic intestinal pseudo-obstruction (CIIP), and specifically of an enteric neuropathy. Full-thickness small intestinal biopsies have confirmed the diagnosis in some patients, while in others a deficiency of α actin staining in the inner circular layer of smooth muscle has been demonstrated. More appropriately these patients should be considered to have a gastrointestinal neuromuscular disorder (GINMD) of the gut. About 10% of these patients are subsequently found to have an underlying autoimmune overlap disorder (see p. 548).

Treatment of patients with pain/gas/bloat syndrome is not easy; and in some, pain can be chronic and severe. Narcotics should always be avoided. Central and end-organ targeted treatment approaches should be combined, e.g. selective serotonin reuptake inhibitor paroxetine combined with a prokinetic agent domperidone or smooth muscle relaxant, e.g. mebeverine. Treatment of patients with neuromuscular disorders of the gut requires a multidisciplinary approach, with emphasis on management of pain, psychological state and nutrition. Patients with underlying autoimmune inflammatory mixed connective tissue disorders may benefit from primary treatment of these.

Functional diarrhoea

In this form of functional bowel disease, symptoms occur in the absence of abdominal pain and commonly are:

- The passage of several stools in rapid succession usually first thing in the morning. No further bowel action may occur that day or defecation only after meals.
- The first stool of the day is usually formed, the later ones mushy, looser or watery.
- Urgency of defecation.
- Anxiety, uncertainty about bowel function with restriction of movement (e.g. travelling).
- Exhaustion after the 'morning rush'.

Chronic diarrhoea without pain is caused by many diseases indistinguishable by history from functional diarrhoea. Features atypical for a functional disorder (e.g. large-volume stools, rectal bleeding, nutritional deficiency and weight loss) call for more extensive studies of intestinal structure and function. In cases where it proves difficult to distinguish between functional and organic causes of diarrhoea, patients should be admitted to hospital for a formal 3-day analysis of stool weights and faecal fat estimation, and a purgative screen together with stool osmolality and creatinine contents to exclude factitious causes of diarrhoea (see p. 308). Outpatient analysis of stool weights is unreliable as brain–gut dysrhythmia may result in increased stool weights in the normal home environment.

Treatment of functional diarrhoea is with loperamide often combined with a tricyclic antidepressant prescribed at night (e.g. clomipramine 10–30 mg).

FURTHER READING

Castle MZD, Silk DBA, Libby GW. Review article: the rationale for antidepressant therapy in functional gastrointestinal disorders. *Alimentary Pharmacology and Therapeutics* 2004; **19**: 969–979.

Cogliandro RF, De Giorgio R, Barbara G et al. Chronic intestinal pseudo-obstruction. *Best Practice and Research Clinical Gastroenterology* 2007; **21**: 657–669.

Drossman DA. The functional gastrointestinal disorders and the Rome III process. *Gastroenterology* 2006; **130**: 1377–1390.

Knowles CH, Silk DB, Darzi A et al. Deranged smooth muscle α-actin as a biomarker of intestinal pseudo-obstruction: a controlled multinational case series. *Gut* 2004; **53**: 1583–1589.

Mayer EA, Bradesi S, Chang L et al. Functional gastrointestinal disorders: from animal models to drug development. *Gut* 2008; **57**: 384–404.

Spiller R. Clinical update: irritable bowel syndrome. *Lancet* 2007; **369**: 1586–1588.

Spiller R, Aziz Q, Creed F et al. Guidelines on the irritable bowel syndrome: mechanisms and practical management. *Gut* 2007; **56**: 1770–1798.

THE ACUTE ABDOMEN

This section deals with the acute abdominal conditions that cause the patient to be hospitalized within a few hours of the onset of pain (Table 6.29). The diagnosis when made quickly reduces morbidity and mortality. Although a specific diagnosis should be attempted, the immediate problem in management is to decide whether an 'acute abdomen' exists and whether surgery is required.

History

This should include previous operations, any gynaecological problems and whether any concurrent medical condition is present.

Pain

The onset, site, type and subsequent course of the pain should be determined as accurately as possible. In general, the pain of an acute abdomen can either be constant (usually owing to inflammation) or colicky because of a blocked 'tube'. The inflammatory nature of a constant pain will be supported by a raised temperature, tachycardia and/or a raised white cell count. If these are normal, then other causes (e.g. musculoskeletal, aortic aneurysm) or rare causes (e.g. porphyria) should be considered. Colicky pain can be due to an obstruction of the gut, biliary system, urogenital system or the uterus. These will probably initially require conservative management along with analgesics. If a colicky pain becomes a constant pain, then inflammation of the organ may have supervened (e.g. strangulated hernia, ascending cholangitis or salpingitis).

A *sudden onset* of pain suggests:

- a perforation (e.g. of a duodenal ulcer)
- a rupture (e.g. of an aneurysm)
- torsion (e.g. of an ovarian cyst)
- acute pancreatitis.

Back pain suggests:

- pancreatitis
- rupture of an aortic aneurysm
- renal tract disease.

Inflammatory conditions (e.g. appendicitis) produce a more gradual onset of pain. With peritonitis the pain is continuous and may be made worse by movement.

Table 6.29 Common causes of acute abdominal pain

Diagnosis	No. of patients
Non-specific abdominal pain	466
Acute appendicitis	449
Renal colic	61
Gynaecological disorders	44
Intestinal obstruction	32
Urinary tract infection	30
Gall bladder disease	12
Perforated ulcer/dyspepsia	10
Diverticular disease	6
Other diagnoses	58
No diagnosis established	39

Data drawn from a series of 1204 patients reported by Dixon JM, Elton RA, Rainey JB et al. Rectal examination in patients with pain in the right lower quadrant of the abdomen. *British Medical Journal* 1991; **302**: 386–388.

Vomiting

Vomiting may accompany any acute abdominal pain but, if persistent, it suggests an obstructive lesion of the gut. The character of the vomit should be asked – does it contain blood, bile or small bowel contents?

Other symptoms

Any change in bowel habit or of urinary frequency should be documented and, in females, a gynaecological history should be taken.

Physical examination

The general condition of the person should be noted. Does the patient look ill? Is he or she shocked? Large volumes of fluid may be lost from the vascular compartment into the peritoneal cavity or into the lumen of the bowel, giving rise to hypovolaemia, i.e. a pale cold skin, a weak rapid pulse and hypotension.

The abdomen

- **Inspection**. Look for the presence of scars, distension or masses.
- **Palpation**. The abdomen should be examined gently for sites of tenderness and the presence or absence of guarding. Guarding is involuntary spasm of the abdominal wall and it indicates peritonitis. This can be localized to one area or it may be generalized, involving the whole abdomen.
- **Bowel sounds**. Increased high-pitch tinkling bowel sounds indicate fluid obstruction; this occurs because of fluid movement within the large dilated bowel lumen. Absent bowel sounds suggest peritoneal involvement. In an obstructed patient, absent bowel sounds suggest strangulation, ischaemia or ileus. It is essential that the hernial orifices be examined if intestinal obstruction is suspected.

Vaginal and rectal examination

Vaginal examination can be very helpful, particularly in diagnosing gynaecological causes of an acute abdomen (e.g. a ruptured ectopic pregnancy). Rectal examination is less helpful as localized tenderness may be due to any cause; it may show blood on the finger stall.

Sigmoidoscopy

If diarrhoea is present, sigmoidoscopy is indicated to aid exclusion of infective, inflammatory and ischaemic causes of acute pain. A specimen of stool should be taken for stool culture for bacterial pathogens (e.g. *Campylobacter*, *Salmonella*, *Shigella*) when diarrhoea is present – stool should also be tested for *Clostridium difficile* toxin if antibiotic therapy or hospital precedes onset of diarrhoea and acute abdominal pain (see p. 131).

Other observations

- *Mouth*. The tongue is furred in some cases and a fetor is present.
- *Temperature*. Fever is more common in acute inflammatory processes.
- *Urine*. Examine for:
  - blood – suggests urinary tract infection or renal colic
  - glucose and ketones – ketoacidosis can present with acute pain
  - protein and white cells – to exclude acute pyelonephritis.
- Think of medical causes (Table 6.30).

Investigations

- **Blood count**. A raised white cell count occurs in inflammatory conditions.
- **Serum amylase**. High levels (more than five times normal) indicate acute pancreatitis. Raised levels below this can occur in any acute abdomen and should not be considered diagnostic of pancreatitis.
- **Serum electrolytes**. These are not particularly helpful for diagnosis but useful for general evaluation of the patient.
- **Pregnancy**. A urine dipstick is used with women of child-bearing age.
- **X-rays**. A chest X-ray is useful to detect air under the diaphragm owing to a perforation. Dilated loops of bowel or fluid levels are suggestive of obstruction (erect and supine abdominal X-ray).
- **Ultrasound**. This is useful in the diagnosis of acute cholangitis, cholecystitis and aortic aneurysm, and in expert hands is reliable in the diagnosis of acute appendicitis. Gynaecological and other pelvic causes of pain can be detected.
- **CT scan**. Spiral CT is the most accurate investigation in most acute emergencies. It should be used more often to avoid unnecessary laparotomies.
- **Laparoscopy**. This has gained increasing importance as a diagnostic tool prior to proceeding with surgery, particularly in men and women over the age of 50 years. In addition, therapeutic manoeuvres, such as appendicectomy, can be performed.

Table 6.30 Medical causes of acute abdomen

Referred pain
Pneumonia
Myocardial infarction

Functional gastrointestinal disorders

Renal causes
Pelviureteric colic
Acute pyelonephritis

Metabolic causes
Diabetes mellitus
Acute intermittent porphyria
Lead poisoning
Familial Mediterranean fever

Haematological causes
Haemophilia and other bleeding disorders
Henoch–Schönlein purpura
Sickle cell crisis
Polycythaemia vera
Paroxysmal nocturnal haemoglobinaemia

Vasculitis
Embolic

Acute appendicitis

This is the most common surgical emergency. It affects all age groups. Appendicitis should always be considered in the differential diagnosis if the appendix has not been removed.

Acute appendicitis mostly occurs when the lumen of the appendix becomes obstructed with a faecolith; however, in some cases there is only generalized acute inflammation. If the appendix is not removed at this stage, gangrene occurs with perforation, leading to a localized abscess or to generalized peritonitis.

Clinical features and management

Most patients present with abdominal pain; in many it starts vaguely in the centre of the abdomen, becoming localized to the right iliac fossa in the first few hours. Nausea, vomiting, anorexia and occasional diarrhoea can occur.

Examination of the abdomen usually reveals tenderness in the right iliac fossa, with guarding due to the localized peritonitis. There may be a tender mass in the right iliac fossa. Although raised white cell counts, ESR and CRP are helpful, other laboratory tests can be unhelpful. An ultrasound scan can detect an inflamed appendix and can also indicate an appendix mass or other localized lesion. CT is highly sensitive and specific, and reduces the incidence of removing the 'normal' appendix. With the use of these investigations the incidence of 'normal' appendix histology has fallen to 15–20%.

Differential diagnosis

- Non-specific mesenteric lymphadenitis – may mimic appendicitis.
- Acute terminal ileitis due to Crohn's disease or *Yersinia* infection.
- Gynaecological causes.
- Inflamed Meckel's diverticulum.
- Functional bowel disease.

Treatment

The appendix is removed by open surgery or laparoscopically. If an appendix mass is present, the patient is usually treated conservatively with intravenous fluids and antibiotics. The pain subsides over a few days and the mass usually disappears over a few weeks. Interval appendicectomy is recommended at a later date to prevent further acute episodes.

Gynaecological causes

Ruptured ectopic pregnancy. The fallopian tube is the commonest extrauterine site of implantation. Delayed diagnosis is the major cause of morbidity. Most patients will present

Ask the authors
www.studentconsult.com

with recurrent low abdominal pain associated with vaginal bleeding. Diagnosis is usually made with abdominal and transvaginal ultrasound. Most patients can be managed by laparoscopic salpingostomy or salpingectomy.

Ovarian:

- Rupture of 'functional' ovarian cysts in the middle of the cycle (Mittelschmerz).
- Torsion or rupture of ovarian cysts.

Acute salpingitis. Most cases are associated with sexually transmitted infection. Patients commonly present with bilateral low abdominal pain, a fever and vaginal discharge. In the Fitz–Hughes–Curtis syndrome the *Chlamydia* infection tracks up the right paracolic gutter to cause a perihepatitis. Patients can present with acute right hypochondrial pain, fever and mildly abnormal liver biochemistry.

Acute peritonitis

Localized peritonitis

There is virtually always some degree of localized peritonitis with all acute inflammatory conditions of the gastrointestinal tract (e.g. acute appendicitis, acute cholecystitis). Pain and tenderness are largely features of this localized peritonitis. The treatment is for the underlying disease.

Generalized peritonitis

This is a serious condition resulting from irritation of the peritoneum owing to infection (e.g. perforated appendix), or from chemical irritation due to leakage of intestinal contents (e.g. perforated ulcer). In the latter case, superadded infection gradually occurs; *E. coli* and *Bacteroides* are the most common organisms.

The peritoneal cavity becomes acutely inflamed, with production of an inflammatory exudate that spreads throughout the peritoneum, leading to intestinal dilatation and paralytic ileus.

Clinical features and management

In perforation, the onset is sudden with acute severe abdominal pain, followed by general collapse and shock. The patient may improve temporarily, only to become worse later as generalized toxaemia occurs.

When the peritonitis is secondary to inflammatory disease, the onset is less rapid with the initial features being those of the underlying disease.

Investigations should always include an erect chest. X-ray is used to detect free air under the diaphragm, and a serum amylase to diagnose acute pancreatitis, which is treated conservatively. Imaging with ultrasound and/or CT should always be performed for diagnosis.

Peritonitis is treated surgically after adequate resuscitation with the re-establishment of a good urinary output. This includes insertion of a nasogastric tube, intravenous fluids and antibiotics. Surgery has a twofold objective:

- peritoneal lavage of the abdominal cavity
- specific treatment of the underlying condition.

Complications

Any delay in treatment of peritonitis produces more profound toxaemia and septicaemia which may lead to development of multiorgan failure (see p. 909). Local abscess formation can occur and should be suspected if a patient continues to remain unwell postoperatively with a swinging fever, high white cell count and continuing pain. Abscesses are commonly pelvic or subphrenic and can be localized and drained using ultrasound and CT scanning techniques.

Intestinal obstruction

Most intestinal obstruction is due to a mechanical block. Sometimes the bowel does not function, leading to a paralytic ileus. This occurs temporarily after most abdominal operations and with peritonitis. Some causes of intestinal obstruction are shown in Table 6.31.

Obstruction of the bowel leads to bowel distension above the block, with increased secretion of fluid into the distended bowel. Bacterial contamination occurs in the distended stagnant bowel. In strangulation the blood supply is impeded, leading to gangrene, perforation and peritonitis unless urgent treatment of the condition is undertaken.

Clinical features

The patient complains of abdominal colic, vomiting and constipation without passage of wind. In upper gut obstruction the vomiting is profuse but in lower gut obstruction it may be absent.

Examination of the abdomen reveals distension with increased bowel sounds. Marked tenderness suggests strangulation, and urgent surgery is necessary. Examination of the hernial orifices and rectum must be performed. X-ray of the abdomen reveals distended loops of bowel proximal to the obstruction. Fluid levels are seen in small bowel obstruction on an erect film. In large bowel obstruction, the caecum and ascending colon are distended. An instant water-soluble gastrografin enema without air insufflation may help to demonstrate the site of the obstruction. CT can localize the lesion accurately and is the investigation of choice.

Management

Initial management is by resuscitation with intravenous fluids (mainly isotonic saline with potassium) and decompression. Many cases will settle on conservative management, but an increasing temperature, raised pulse rate, increasing pain and a rising white cell count require urgent scanning and possible exploratory laparotomy.

Laparotomy with removal of the obstruction will be necessary in some cases of small bowel obstruction. If the bowel is gangrenous owing to strangulation, gut resection will be required. A few patients (e.g. those with Crohn's disease) may have recurrent episodes of incomplete intestinal obstruc-

Table 6.31 Some causes of intestinal obstruction

Small intestinal obstruction
Adhesions (80% in adults)
Hernias
Crohn's disease
Intussusception
Obstruction due to extrinsic involvement by cancer
Colonic obstruction
Carcinoma of the colon
Sigmoid volvulus
Diverticular disease

tion that can be managed conservatively. In large bowel obstruction due to malignancy, colorectal stents are being used, followed by elective surgery. In critically ill patients, a defunctioning colostomy may be required. Volvulus of the sigmoid colon can be managed by the passage of a flexible sigmoidoscope or a rectal tube to un-kink the bowel, but recurrent volvulus may require sigmoid resection.

Acute colonic pseudo-obstruction

It is now recognized that a clinical picture mimicking mechanical obstruction may develop in patients who do not have a mechanical cause. In more than 80% of cases it complicates other clinical conditions, for example:

- intra-abdominal trauma, pelvic spinal and femoral fractures
- postoperatively (abdominal, pelvic, cardiothoracic, orthopaedic, neurosurgical)
- intra-abdominal sepsis
- pneumonia
- metabolic (e.g. electrolyte disturbances, malnutrition, diabetes mellitus, Parkinson's disease)
- drugs – opiates (particularly after orthopaedic surgery), antidepressants, antiparkinsonian drugs.

Patients present with rapid and progressive abdominal distension and pain. X-ray shows a gas-filled large bowel. Management is of the underlying problem (e.g. withdraw opiate analgesia) together with a trial of i.v. neostigmine therapy (Box 6.13). Patients should be monitored carefully and consideration should be given to surgery if the diameter of the caecum exceeds 14 cm.

Small intestinal pseudo-obstruction is a rare chronic condition that can occur in association with systemic sclerosis, systemic lupus erythematosus (SLE), Sjögren's syndrome, thyroid disease, amyloidosis and paraneoplastic syndromes. Primary myopathic and neuropathic forms also exist, with the former sometimes being familial. There are other patients with clinical and manometric features of small intestinal pseudo-obstruction who have normal full-thickness biopsies of smooth muscle but α actin deficiency in the inner circular layer of the smooth muscle. Myopathic forms can present with attacks of non-mechanical obstruction and/or functional small intestinal failure with dilated non-propulsive intestines and coexisting bacterial overgrowth. These patients are managed in specialist centres with facilities to manage home total parenteral nutrition (TPN). The other patients, including those with enteric neuropathies, often present with a long history of abdominal pain and other intractable midgut symptoms. Many have extraintestinal symptoms, and a multidisciplinary approach to management is required, including, in a number, access to facilities to provide needle catheter jejunostomy enteral feeding as well as home TPN.

Box 6.13 Treatment of acute colonic pseudo-obstruction

- Neostigmine 2.0 mg i.v. over 3–5 min in presence of doctor with ECG monitor
- 0.3–1 mg atropine if symptomatically bradycardic. Nurse the patient supine for 60 min
- Monitor abdominal circumference and the diameter of the caecum, ascending, transverse and descending colon on straight abdominal X-ray

THE PERITONEUM

The peritoneal cavity is a closed sac lined by mesothelial cells, which produce surfactant that acts as a lubricant within the peritoneal cavity. The cavity contains less than 100 mL of serous fluid containing less than 30 g/L of protein.

The mesothelial cells lining the diaphragm have gaps that allow communication between the peritoneum and the diaphragmatic lymphatics. Approximately one-third of fluid drains through these lymphatics, the remainder through the parietal peritoneum. These mechanisms allow particulate matter to be removed rapidly from the peritoneal cavity.

Complement activation is an early defence mechanism followed rapidly by upregulation of the peritoneal mesothelial cells and migration of polymorphonuclear neutrophils and macrophages into the peritoneum.

Mast cells release potent mediators of inflammation and interact with T cells to generate an immune response.

The peritoneal-associated lymphoid tissue includes the omental milky spots, the lymphocytes within the peritoneal cavity and the draining lymph nodes. B cells with a unique CD5+ are common. This defence system plays a major role in localizing peritoneal infection. Some conditions that can affect the peritoneum are shown in Table 6.32.

Peritonitis can be acute or chronic, as seen in tuberculosis. Most cases of infective peritonitis are secondary to gastrointestinal disease, but it occurs occasionally without intra-abdominal sepsis in ascites due to liver disease. Very rarely, fungal and parasitic infections can also cause primary peritonitis (e.g. amoebiasis, candidiasis). Peritonitis is discussed further on page 316.

The peritoneum can be involved by secondary malignant deposits, and the most common cause of ascites in a young to middle-aged woman is an ovarian carcinoma.

A subphrenic abscess is usually secondary to infection in the abdomen and is characterized by fever, malaise, pain in the right or left hypochondrium and shoulder-tip pain. An erect chest X-ray shows gas under the diaphragm, impaired movement of the diaphragm on screening and a pleural effusion. Ultrasound is usually diagnostic. Percutaneous catheter

Table 6.32 Disease of the peritoneum

Infective (bacterial) peritonitis
Secondary to gut disease, e.g.
appendicitis
perforation of any organ
Chronic peritoneal dialysis
Spontaneous, usually in ascites with liver disease
Tuberculosis
Neoplasia
Secondary deposits (e.g. from ovary, stomach)
Primary mesothelioma
Vasculitis
Rheumatic autoimmune disease
Polyserositis (e.g. familial Mediterranean fever)

FURTHER READING

Silk DBA. Chronic idiopathic intestinal pseudo-obstruction: the need for a multidisciplinary approach to management. *Proceedings of the Nutrition Society* 2004; **63**: 1–8.

drainage inserted under CT or ultrasound guidance and antibiotics is highly successful therapy.

Ascites is associated with all diseases of the peritoneum. The fluid that collects is an exudate with a high protein content. It is also seen in liver disease. The mechanism, causes and investigation of ascites are discussed on page 351.

Retroperitoneal fibrosis (periaortitis)

This is a rare condition in which there is a marked fibrosis over the posterior abdominal wall and retroperitoneum. It is described on p. 616.

Tuberculous peritonitis

This is the second most common form of abdominal TB.

Three subgroups can be identified: wet, dry and fibrous.

- In patients with the wet type, ascitic fluid should be examined for protein concentration (>20 g/L) and tubercle bacilli (rarely found).
- In the dry form, patients present with subacute intestinal obstruction which is due to tuberculous small bowel adhesions.
- In the fibrous form, patients present with abdominal pain, distension and ill-defined irregular tender abdominal masses.

The diagnosis of peritoneal TB can be supported by findings on ultrasound or CT screening (mesenteric thickening and lymph node enlargement). A histological diagnosis is not always required before instituting treatment. In some patients careful laparoscopy (to avoid perforation) may have to be performed, and rarely laparotomy.

Treatment

Drug treatment is similar to that for pulmonary TB (see p. 866) and should be supervised by chest physicians who have experience in dealing with contacts.

FURTHER READING

Hall JC, Heel JM, Papadimitriou JM et al. The pathobiology of peritonitis. *Gastroenterology* 1998; **114**: 188–196.

CHAPTER BIBLIOGRAPHY

Feldman M, Scharschmidt BF, Sleisenger MH. *Sleisenger and Fordtran's Gastrointestinal and Liver Disease*, 8th edn. Philadelphia: WB Saunders, 2006.

SIGNIFICANT WEBSITES

http://www.bsg.org.uk
British Society of Gastroenterology

http://www.coeliac.co.uk
Coeliac disease

http://www.digestivedisorders.org.uk/leaflets/ibs.html
Irritable bowel syndrome

http://www.graylab.ac.uk/cancernet/100025.html
Gastric cancer

http://www.jr2.ox.ac.uk/bandolier/band27/b27-7.htmlband38/b38-4.html
Gastric ulcer and GORD

http://www.nacc.org.uk
UK National Association for colitis and Crohn's disease

7 Liver, biliary tract and pancreatic disease

THE LIVER

Structure of the liver and biliary system 319
Functions of the liver 320
Investigations 323
Symptoms of liver disease 328
Signs of liver disease 328
Jaundice 329
Hepatitis 332
Viral hepatitis 333
Hepatitis A 333
Hepatitis B 335
Hepatitis D 339
Hepatitis C 340
Hepatitis E 341
Hepatitis non-A–E 342
Acute hepatitis due to other infectious agents 342
Fulminant hepatic failure (FHF) 342
Autoimmune hepatitis 343
Drug-induced chronic hepatitis 344
Chronic hepatitis of unknown cause 344
Non-alcoholic fatty liver disease 344
Cirrhosis 345
Liver transplantation 347
Complications and effects of cirrhosis 348
Types of cirrhosis 354
Alcoholic liver disease 358
Budd–Chiari syndrome 359
Veno-occlusive disease 360
Fibropolycystic diseases 360
Liver abscess 360
Other infections of the liver 361
Liver disease in pregnancy 362
Liver tumours 363
Primary malignant tumours 363
Benign tumours 364
Miscellaneous conditions of the liver 364
Drugs and the liver 365

GALL BLADDER AND BILIARY SYSTEM

Gallstones 367
Management of gall bladder stones 369
Common bile duct stones 370
Complications of gallstones 372
Miscellaneous conditions of the biliary tract 372
Gall bladder 372
Extrahepatic biliary tract 372
Tumours of the biliary tract 373

THE PANCREAS

Structure and function 375
Investigation of the pancreas 376
Pancreatitis 377
Carcinoma of the pancreas 382
Cystic tumours of the pancreas 384
Neuroendocrine tumours of the pancreas 384

In many countries alcohol is the major cause of liver disease, followed by hepatitis C virus infection. Hepatitis B virus is still a significant factor but widespread vaccination will reduce its prevalence. Health education and the improvement in public health should help to stop the spread of viral infections.

Imaging techniques enable the liver, biliary tree and pancreas to be visualized with precision, resulting in earlier diagnosis. Therapeutic endoscopy, laparoscopic and minimally invasive surgery are now widely available for biliary tract and pancreatic disease. Finally, liver transplantation is established therapy for both acute and chronic liver disease.

THE LIVER

STRUCTURE OF THE LIVER AND BILIARY SYSTEM

The liver

The liver is the largest internal organ in the body (1.2–1.5 kg) and is situated in the right hypochondrium. Functionally, it is divided into right and left lobes by the middle hepatic vein. The right lobe is larger and contains the caudate and quadrate lobes. The liver is further subdivided into a total of eight segments (Fig. 7.1) by divisions of the right, middle and left hepatic veins. Each segment receives its own portal pedicle, permitting individual segment resection at surgery.

The blood supply to the liver constitutes 25% of the resting cardiac output and is via two main vessels:

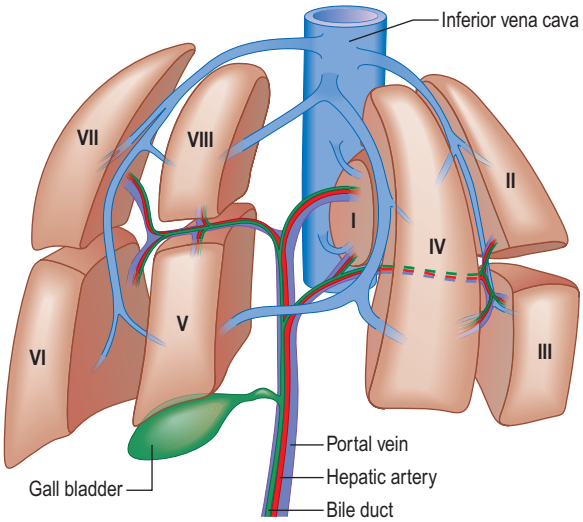


Fig. 7.1 **Segmental anatomy of the liver showing the eight hepatic segments.** I, caudate lobe; II–IV the left hemiliver; V–VIII the right hemiliver.

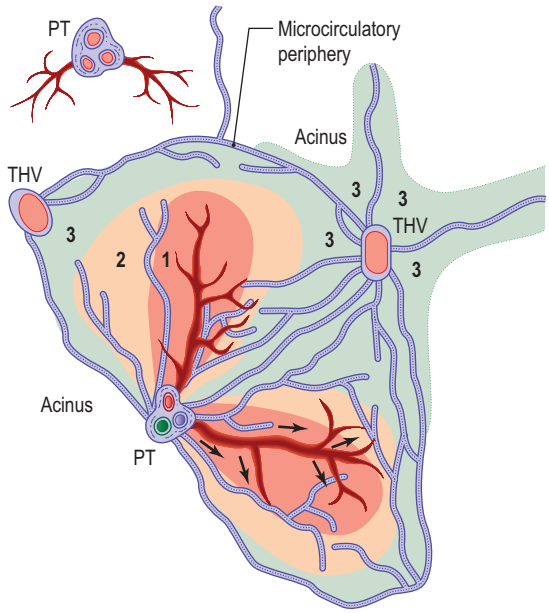


Fig. 7.2 **Diagram of an acinus.** *Zones 1, 2 and 3* represent areas supplied by blood, with zone 1 being best oxygenated. *Zone 3* is supplied by blood remote from afferent vessels and is in the microcirculatory periphery of the acinus. The perivascular area (star shaped green area around THV) is formed by the most peripheral parts of zone 3 of several adjacent acini and is the least well oxygenated. THV, terminal hepatic venule; PT, portal triad.

- *The hepatic artery*, which is a branch of the coeliac axis, supplies 25% of the total blood flow. Autoregulation of blood flow by the hepatic artery ensures a constant total liver blood flow.
- *The portal vein* drains most of the gastrointestinal tract and the spleen. It supplies 75% of the blood flow. The normal portal pressure is 5–8 mmHg; flow increases after meals.

Both vessels enter the liver via the hilum (porta hepatis). The blood from these vessels is distributed to the segments and passes into the sinusoids via the portal tracts.

Blood leaves the sinusoids, entering branches of the hepatic vein which join into three main branches before entering the inferior vena cava.

The *caudate lobe* is an autonomous segment as it receives an independent blood supply from the portal vein and hepatic artery, and its hepatic vein drains directly into the inferior vena cava.

Lymph, formed mainly in the perisinusoidal space, is collected in lymphatics which are present in the portal tracts. These small lymphatics enter larger vessels which eventually drain into the hepatic ducts.

The acinus is the functional unit of the liver. This consists of parenchyma supplied by the smallest portal tracts containing portal vein radicles, hepatic arterioles and bile ductules (Fig. 7.2). The hepatocytes near this triad (zone 1) are well supplied with oxygenated blood and are more resistant to damage than the cells nearer the terminal hepatic (central) veins (zone 3).

The *sinusoids* lack a basement membrane and are loosely surrounded by specialist fenestrated endothelial cells and Kupffer cells (phagocytic cells). Sinusoids are separated by plates of liver cells (hepatocytes). The subendothelial space that lies between the sinusoids and hepatocytes is the space of Disse, which contains a matrix of basement membrane constituents and stellate cells. (See Fig. 7.22.)

Stellate cells store retinoids in their resting state and contain the intermediate filament, desmin. When activated (to myofibroblasts) they are contractile and probably regulate sinusoidal blood flow. Endothelin and nitric oxide play a major role in modulating stellate cell contractility. Stellate cells, after activation, produce extracellular matrix components, including collagen (see p. 345).

FURTHER READING

Sherlock S, Dooley J. Anatomy and function. In: *Diseases of the Liver and Biliary System,* 11th edn. Oxford: Blackwell Science, 2002.

The biliary system

Bile canaliculi form a network between the hepatocytes. These join to form thin bile ductules near the portal tract, which in turn enter the bile ducts in the portal tracts. These then combine to form the right and left hepatic ducts that leave each liver lobe. The hepatic ducts join at the porta hepatis to form the common hepatic duct. The cystic duct connects the gall bladder to the lower end of the common hepatic duct. The gall bladder lies under the right lobe of the liver and stores and concentrates hepatic bile; it has a capacity of approximately 50 mL. The common bile duct is formed by the combination of the cystic and hepatic ducts and is approximately 8 mm in diameter, narrowing at its distal end to pass into the duodenum. The common bile duct and pancreatic duct open into the second part of the duodenum through a common channel at the ampulla of Vater. The lower end of the common bile duct contains the muscular sphincter of Oddi, which contracts rhythmically and prevents bile from entering the duodenum in the fasting state.

FUNCTIONS OF THE LIVER

Protein metabolism (see also p. 209)

Synthesis and storage

The liver is the principal site of synthesis of all circulating proteins apart from γ-globulins, which are produced in the reticuloendothelial system. The liver receives amino acids

from the intestine and muscles and, by controlling the rate of gluconeogenesis and transamination, regulates levels in the plasma. Plasma contains 60–80 g/L of protein, mainly in the form of albumin, globulin and fibrinogen.

Albumin has a half-life of 16–24 days, and 10–12 g are synthesized daily. Its main functions are first to maintain the intravascular oncotic (colloid osmotic) pressure, and second to transport water-insoluble substances such as bilirubin, hormones, fatty acids and drugs. Reduced synthesis of albumin over prolonged periods produces hypoalbuminaemia and is seen in chronic liver disease and malnutrition. Hypoalbuminaemia is also found in hypercatabolic states (e.g. trauma with sepsis) and in diseases where there is an excessive loss (e.g. nephrotic syndrome, protein-losing enteropathy).

Transport or carrier proteins such as transferrin and caeruloplasmin, acute-phase and other proteins (e.g. α_1-antitrypsin and α-fetoprotein) are also produced in the liver.

The liver also synthesizes all factors involved in coagulation (apart from one-third of factor VIII) – that is, fibrinogen, prothrombin, factors V, VII, IX, X and XIII, proteins C and S and antithrombin (see Ch. 8) as well as components of the complement system. The liver stores large amounts of vitamins, particularly A, D and B_{12}, and lesser amounts of others (vitamin K and folate), and also minerals – iron in ferritin and haemosiderin and copper.

Degradation (nitrogen excretion)

Amino acids are degraded by transamination and oxidative deamination to produce ammonia, which is then converted to urea and excreted by the kidneys. This is a major pathway for the elimination of nitrogenous waste. Failure of this process occurs in severe liver disease.

Carbohydrate metabolism

Glucose homeostasis and the maintenance of the blood sugar is a major function of the liver. It stores approximately 80 g of glycogen. In the immediate fasting state, blood glucose is maintained either by glucose released from the breakdown of glycogen (glycogenolysis) or by newly synthesized glucose (gluconeogenesis). Sources for gluconeogenesis are lactate, pyruvate, amino acids from muscles (mainly alanine and glutamine) and glycerol from lipolysis of fat stores. In prolonged starvation, ketone bodies and fatty acids are used as alternative sources of fuel as the body tissues adapt to a lower glucose requirement (see Ch. 5).

Lipid metabolism

Fats are insoluble in water and are transported in the plasma as protein–lipid complexes (lipoproteins). These are discussed in detail on page 1061.

The liver has a major role in the metabolism of lipoproteins. It synthesizes very-low-density lipoproteins (VLDLs) and high-density lipoproteins (HDLs). HDLs are the substrate for lecithin–cholesterol acyltransferase (LCAT), which catalyses the conversion of free cholesterol to cholesterol ester (see below). Hepatic lipase removes triglyceride from intermediate-density lipoproteins (IDLs) to produce low-density lipoproteins (LDLs) which are degraded by the liver after uptake by specific cell-surface receptors (see Fig. 19.19).

Triglycerides are mainly of dietary origin but are also formed in the liver from circulating free fatty acids (FFAs) and glycerol and incorporated into VLDLs. Oxidation or de novo synthesis of FFA occurs in the liver, depending on the availability of dietary fat.

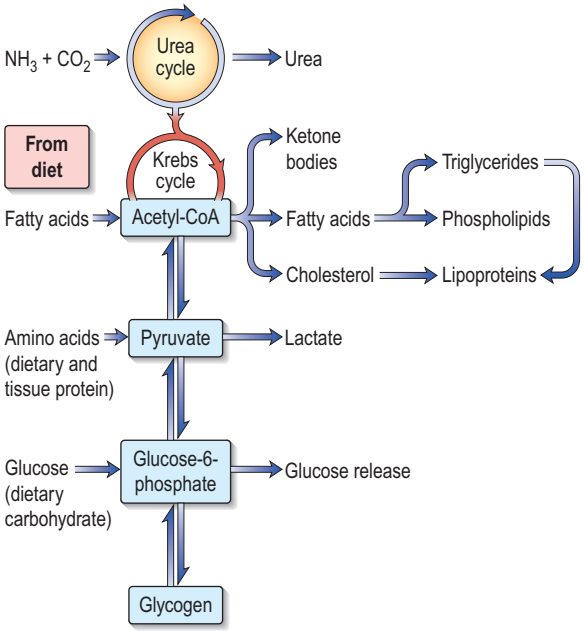


Fig. 7.3 **Interrelationships of protein, carbohydrate and lipid metabolism in the liver.**

Cholesterol may be of dietary origin but most is synthesized from acetyl-CoA mainly in the liver, intestine, adrenal cortex and skin. It occurs either as free cholesterol or is esterified with fatty acids; this reaction is catalysed by LCAT. This enzyme is reduced in severe liver disease, increasing the ratio of free cholesterol to ester, which alters membrane structures. One result of this is the red cell abnormalities (e.g. target cells) seen in chronic liver disease. Phospholipids (e.g. lecithin) are synthesized in the liver. The complex interrelationships between protein, carbohydrate and fat metabolism are shown in Figure 7.3.

Formation of bile

Bile secretion and bile acid metabolism

Bile consists of water, electrolytes, bile acids, cholesterol, phospholipids and conjugated bilirubin. Two processes are involved in bile secretion across the canalicular membrane of the hepatocyte – a *bile salt-dependent* and a *bile salt-independent* process – each contributing about 230 mL per day. The remainder of the bile (about 150 mL daily) is produced by the epithelial cells of the bile ductules.

Bile formation requires firstly the uptake of bile acids and other organic and inorganic ions across the basolateral (sinusoidal) membranes by multiple transport proteins (sodium taurocholate co-transporting polypeptide (NTCP) and sodium independent organic anion transporting polypeptide 2 (OATP2), Fig. 7.4). This process is driven by Na^+/K^+-ATPase in the basolateral membrane. Intracellular transport across the hepatocyte is partly through microtubules and partly by cytosol transport proteins.

Bile acids are also synthesized in hepatocytes from cholesterol. The rate-limiting step in their production is that catalysed mainly by cholesterol-7α-hydroxylase and the P450 enzymes (CYP7A1 and CYP8B1).

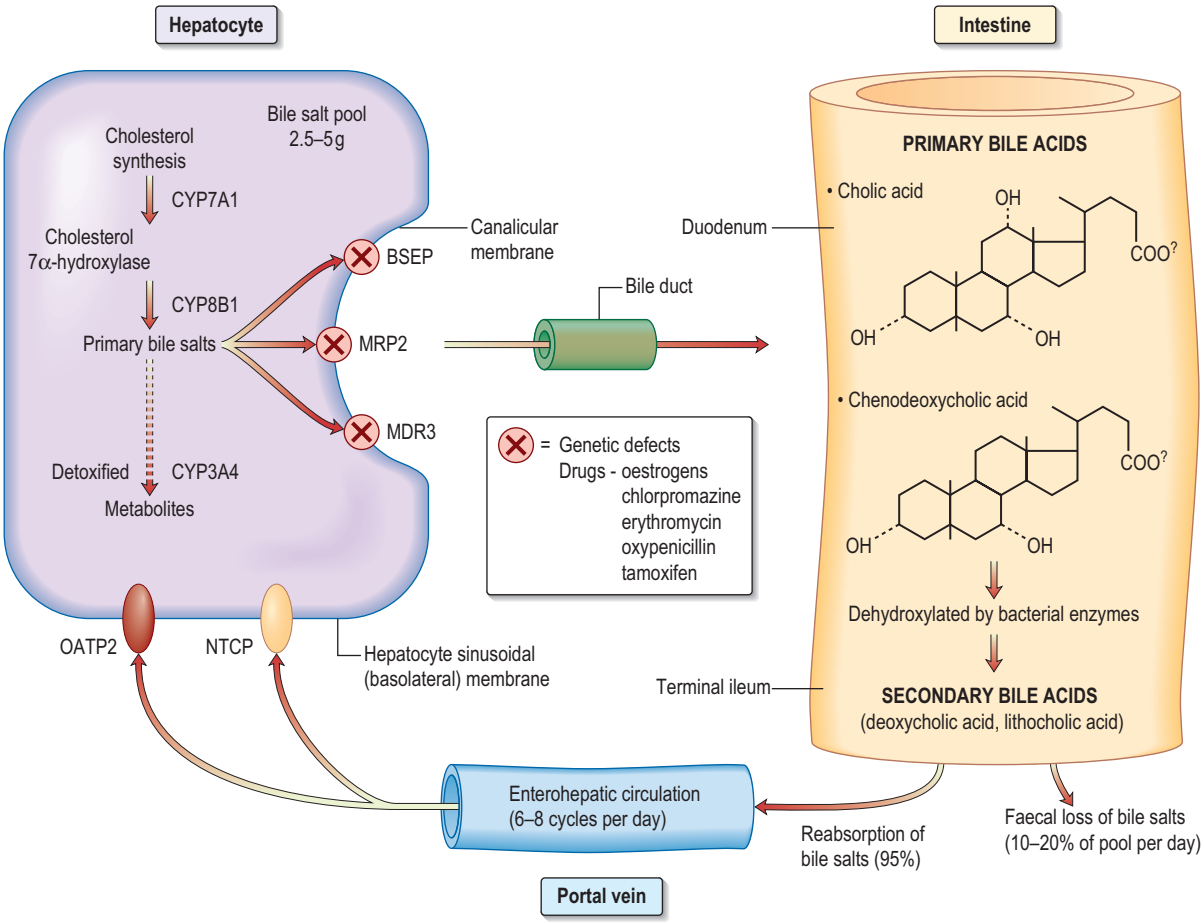


Fig. 7.4 **Cholesterol synthesis and its conversion to primary and secondary bile acids.** All bile acids are normally conjugated with glycine or taurine. BSEP, bile salt export pump; MRP2, multidrug-resistant protein 2; MDR3, multidrug-resistant 3; OATP2, Na^+ independent organic anion-transporting polypeptide 2; NTCP, Na^+ taurocholate co-transporting polypeptide.

The bile acid receptor, farnesol X, blocks bile acid formation from cholesterol and also regulates the transport proteins (NTCP, OATP2) that increase bile acid uptake by the liver.

The canalicular membrane contains additional transporters, mainly ATPase-dependent, which carry molecules into the biliary canaliculi against a concentration gradient. These canalicular multispecific organic anion transporters (multidrug-resistance protein 2 (MRP2)), multidrug resistance protein (MDR3) and the bile salt excretory pump (BSEP)) mediate transport of a broad range of compounds including bilirubin diglucuronide, glucuronidated and sulphated bile acids and other organic anions. Na^+ and water follow the passage of bile salts into the biliary canaliculus by diffusion across the tight junction between hepatocytes (a bile salt-dependent process). In the bile salt-independent process water flow is due to other osmotically active solutes such as glutathione and bicarbonate.

Secretion of a bicarbonate-rich solution is stimulated mainly by secretin and is inhibited by somatostatin. In this process several membrane proteins are involved, including the Cl^-/HCO_3^- exchanger and the cystic fibrosis transmembrane conductance regulator which controls Cl^- secretion, as well as water channels (aquaporins) in cholangiocyte membranes.

The bile acids are excreted into the bile and then pass via the common bile duct into the duodenum. The two primary bile acids – cholic acid and chenodeoxycholic acid (Fig. 7.4) – are conjugated with glycine or taurine (in a ratio of 3:1 in humans) and this process increases their solubility. Intestinal bacteria convert these acids into secondary bile acids, deoxycholic and lithocholic acid. Figure 7.5 shows the enterohepatic circulation of bile acids.

The average total bile flow is approximately 600 mL per day. In the fasted state half of the bile flows directly into the duodenum and half is diverted into the gall bladder. The mucosa of the gall bladder absorbs 80–90% of the water and electrolytes, but is impermeable to bile acids and cholesterol. Following a meal, cholecystokinin is secreted by the I cells of the duodenal mucosa and stimulates contraction of the gall bladder and relaxation of the sphincter of Oddi, so that bile enters the duodenum. An adequate bile flow is dependent on bile salts being returned to the liver by the enterohepatic circulation.

Bile acids act as detergents; their main function is lipid solubilization. Bile acid molecules have both a hydrophilic and a hydrophobic end. In aqueous solutions they aggregate to form micelles, with their hydrophobic (lipid-soluble) ends in the centre. Micelles are expanded by cholesterol and phospholipids (mainly lecithin), forming mixed micelles.

Bilirubin metabolism

Bilirubin is produced mainly from the breakdown of mature red cells in the Kupffer cells of the liver and in the reticulo-

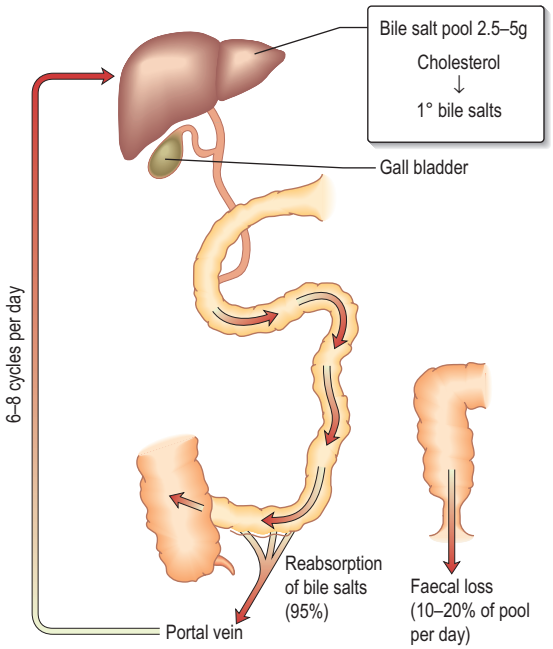


Fig. 7.5 **Recirculation of bile acids.** The bile salt pool is relatively small and the entire pool recycles six to eight times via the enterohepatic circulation. Synthesis of new bile acids compensates for faecal loss.

endothelial system; 15% of bilirubin comes from the catabolism of other haem-containing proteins, such as myoglobin, cytochromes and catalases.

Normally, 250–300 mg (425–510 mmol) of bilirubin are produced daily. The iron and globin are removed from the haem and are reused. Biliverdin is formed from the haem and this is reduced to form bilirubin. The bilirubin produced is unconjugated and water-insoluble, due to internal hydrogen bonding, and is transported to the liver attached to albumin. Bilirubin dissociates from albumin and is taken up by the hepatic cell membrane and transported to the endoplasmic reticulum by cytoplasmic proteins, where it is conjugated with glucuronic acid and excreted into bile. The microsomal enzyme, uridine diphosphoglucuronosyl transferase, catalyses the formation of bilirubin monoglucuronide and then diglucuronide. This conjugated bilirubin is water-soluble and is actively secreted into the bile canaliculi and excreted into the intestine within the bile (Fig. 8.5). It is not absorbed from the small intestine because of its large molecular size. In the terminal ileum, bacterial enzymes hydrolyse the molecule, releasing free bilirubin which is then reduced to urobilinogen. Some of this is excreted in the stools as stercobilinogen. The remainder is absorbed by the terminal ileum, passes to the liver via the enterohepatic circulation, and is re-excreted into the bile. Urobilinogen bound to albumin enters the circulation and is excreted in the urine via the kidneys. When hepatic excretion of conjugated bilirubin is impaired, a small amount of conjugated bilirubin is found strongly bound to serum albumin. It is not excreted by the kidneys and accounts for the continuing hyperbilirubinaemia for a short time after cholestasis has resolved.

Hormone and drug inactivation

The liver catabolizes hormones such as insulin, glucagon, oestrogens, growth hormone, glucocorticoids and parathyroid hormone. It is also the prime target organ for many hormones (e.g. insulin). It is the major site for the metabolism of drugs (see p. 365) and alcohol (see p. 238). Fat-soluble drugs are converted to water-soluble substances that facilitate their excretion in the bile or urine. Cholecalciferol is converted to 25-hydroxycholecalciferol.

Immunological function

The reticuloendothelial system of the liver contains many immunologically active cells. The liver acts as a 'sieve' for the bacterial and other antigens carried to it via the portal tract from the gastrointestinal tract. These antigens are phagocytosed and degraded by Kupffer cells, which are macrophages attached to the endothelium. Kupffer cells have specific membrane receptors for ligands and are activated by several factors, such as infection. They secrete interleukins, tumour necrosis factor (TNF), collagenase and lysosomal hydrolases. Antigens are degraded without the production of antibody, as there is very little lymphoid tissue. They are thus prevented from reaching other antibody-producing sites in the body and thereby prevent generalized adverse immunological reactions. The reticuloendothelial system plays a role in tissue repair, T and B lymphocyte interaction, and cytotoxic activity in disease processes. Following stimulation by, for example, endotoxin, the Kupffer cells release IL-6, IL-8 and TNF-α. These cytokines stimulate the sinusoidal cells, stellate cells and natural killer cells to release proinflammatory cytokines. The stimulated hepatocytes themselves express adhesion molecules and release IL-8, which is a potent neutrophil chemoattractant. Homing of mucosal lymphocytes (enterohepatic circulation) has been proposed. These exogenous leucocytes again release more cytokines – all damaging the function of the hepatocyte, including hepatocellular bile formation which leads to cholestasis. Cytokines also stimulate hepatic apoptosis.

INVESTIGATIONS

Investigative tests can be divided into:

- **Blood tests**
  - **(a)** Liver 'function' tests:
    - (i) serum albumin and bilirubin
    - (ii) prothrombin time
  - **(b)** Liver biochemistry:
    - (i) serum aspartate and alanine aminotransferases – reflecting hepatocellular damage
    - (ii) serum alkaline phosphatase, γ-glutamyl transpeptidase – reflecting cholestasis
    - (iii) total protein
  - **(c)** Viral markers
  - **(d)** Additional blood investigations; haematological, biochemical, immunological, markers of liver fibrosis and genetic analysis.
- **Urine tests** – for bilirubin and urobilinogen
- **Imaging techniques** – to define gross anatomy
- **Liver biopsy** – for histology.

Most routine 'liver function tests' sent to the laboratory will be processed by an automated multichannel analyser to produce serum levels of bilirubin, aminotransferases, alkaline phosphatase, γ-glutamyl transpeptidase (γ-GT) and total proteins. These routine tests are markers of liver damage, but not actual tests of 'function' per se. Subsequent investigations are often based on these tests.

Blood tests

Useful blood tests for certain liver diseases are shown in Table 7.1.

Liver function tests

Serum albumin

This is a marker of synthetic function and is a valuable guide to the severity of chronic liver disease. A falling serum albumin in liver disease is a bad prognostic sign. In acute liver disease initial albumin levels may be normal.

Prothrombin time (PT)

This is also a marker of synthetic function. Because of its short half-life, it is a sensitive indicator of both acute and chronic liver disease. Vitamin K deficiency should be excluded as the cause of a prolonged PT by giving an intravenous bolus (10 mg) of vitamin K. Vitamin K deficiency commonly occurs in biliary obstruction, as the low intestinal concentration of bile salts results in poor absorption of vitamin K.

Prothrombin times vary in different laboratories depending upon the thromboplastin used in the assay. The International Normalized Ratio (INR) is therefore used in many countries (see p. 434) but does not eliminate variation across laboratories.

Table 7.1 Useful blood tests for certain liver diseases

Test	Disease
Anti-mitochondrial antibody	Primary biliary cirrhosis
Anti-nuclear, smooth muscle (actin), liver/kidney microsomal antibody	Autoimmune hepatitis
Raised serum immunoglobulins:	
IgG	Autoimmune hepatitis
IgM	Primary biliary cirrhosis
Viral markers	Hepatitis A, B, C, D and others
α-Fetoprotein	Hepatocellular carcinoma
Serum iron, transferrin saturation, serum ferritin	Hereditary haemochromatosis
Serum and urinary copper, serum caeruloplasmin	Wilson's disease
α_1-Antitrypsin	Cirrhosis (± emphysema)
Anti-nuclear cytoplasmic antibodies	Primary sclerosing cholangitis
Markers of liver fibrosis	Non-alcoholic fatty liver disease Hepatitis C
Genetic analyses	e.g. *HFE* gene (hereditary haemochromatosis)

Liver biochemistry

Bilirubin

In the serum, bilirubin is normally almost all unconjugated. In liver disease, increased serum bilirubin is usually accompanied by other abnormalities in liver biochemistry. Determination of whether the bilirubin is conjugated or unconjugated is only necessary in congenital disorders of bilirubin metabolism (see below) or to exclude haemolysis.

Aminotransferases

These enzymes (often referred to as transaminases) are present in hepatocytes and leak into the blood with liver cell damage. Two enzymes are measured:

- *Aspartate aminotransferase* (AST) is primarily a mitochondrial enzyme (80%; 20% in cytoplasm) and is also present in heart, muscle, kidney and brain. High levels are seen in hepatic necrosis, myocardial infarction, muscle injury and congestive cardiac failure.
- *Alanine aminotransferase* (ALT) is a cytosol enzyme, more specific to the liver so that a rise only occurs with liver disease.

Alkaline phosphatase (ALP)

This is present in the canalicular and sinusoidal membranes of the liver, but is also present in many other tissues, such as bone, intestine and placenta. If necessary, its origin can be determined by electrophoretic separation of isoenzymes or bone-specific monoclonal antibodies. Alternatively, if there is also an abnormality of, for example, the γ-GT, the ALP can be presumed to come from the liver.

Serum ALP is raised in cholestasis from any cause, whether intrahepatic or extrahepatic disease. The synthesis of ALP is increased and this increase is detected in the blood. In cholestatic jaundice, levels may be four to six times the normal limit. Raised levels may also occur in conditions with infiltration of the liver (e.g. metastases) and in cirrhosis, frequently in the absence of jaundice. The highest serum levels due to liver disease (>1000 IU/L) are seen with hepatic metastases and primary biliary cirrhosis.

γ-Glutamyl transpeptidase

This is a microsomal enzyme that is present in many tissues as well as the liver. Its activity can be induced by such drugs as phenytoin and by alcohol. If the ALP is normal, a raised serum γ-GT is a good guide to alcohol intake and can be used as a screening test (see p. 1212). Mild elevation of the γ-GT is common even with a small alcohol consumption and does not necessarily indicate liver disease if the other liver biochemical tests are normal. In cholestasis the γ-GT rises in parallel with the ALP as it has a similar pathway of excretion. This is also true of the 5-nucleotidase, another microsomal enzyme that can be measured in blood.

Total proteins

This measurement, in itself, is of little value. Serum albumin is discussed above. The globulin fraction consists of many proteins that can be separated on electrophoresis. A raised globulin fraction, seen in liver disease, is usually due to increased circulating immunoglobulins and is polyclonal (see below).

Viral markers

Viruses are a major cause of liver disease. Virological studies have a key role in diagnosis; markers are available for most common viruses that cause hepatitis.

Additional blood investigations

Haematological

A full blood count is always performed. Anaemia may be present. The red cells are often macrocytic and can have abnormal shapes – target cells and spur cells – owing to membrane abnormalities. Vitamin B_{12} levels are normal or high, while folate levels are often low owing to poor dietary intake. Other changes are caused by the following:

- Bleeding produces a hypochromic, microcytic picture.
- Alcohol causes macrocytosis, sometimes with leucopenia and thrombocytopenia.
- Hypersplenism results in pancytopenia.
- Cholestasis can often produce abnormal-shaped cells and also deficiency of vitamin K.
- Haemolysis accompanies acute liver failure and jaundice.
- Aplastic anaemia is present in up to 2% of patients with acute viral hepatitis.
- A raised serum ferritin with transferrin saturation (>60%) is seen in hereditary haemochromatosis.

Biochemical

- *a_1-Antitrypsin*. A deficiency of this enzyme can produce cirrhosis.
- *α-Fetoprotein*. This is normally produced by the fetal liver. Its reappearance in increasing and high concentrations in the adult indicates hepatocellular carcinoma. Increased concentrations in pregnancy in the blood and amniotic fluid suggest neural tube defects of the fetus. Blood levels are also slightly raised with regenerative liver tissue in patients with hepatitis, chronic liver disease and also in teratomas.
- Serum and urinary copper and serum caeruloplasmin – for Wilson's disease (see p. 357).

Immunological tests

Serum immunoglobulins

Increased γ-globulins are thought to be due to reduced phagocytosis by sinusoidal and Kupffer cells of the antigens absorbed from the gut. These antigens then stimulate antibody production in the spleen, lymph nodes and lymphoid and plasma cell infiltrate in the portal tracts. In primary biliary cirrhosis, the predominant serum immunoglobulin that is raised is IgM, while in autoimmune hepatitis it is IgG.

Serum autoantibodies

- *Anti-mitochondrial antibody* (AMA) is found in the serum in over 95% of patients with primary biliary cirrhosis (p. 355). Many different AMA subtypes have been described, depending on their antigen specificity. AMA is demonstrated by an immunofluorescent technique and is neither organ- nor species-specific. Some subtypes are occasionally found in autoimmune hepatitis and other autoimmune diseases.
- *Nucleic, smooth muscle (actin), liver/kidney microsomal antibodies* can be found in the serum in high titre in patients with autoimmune hepatitis. These antibodies can be found in the serum in other autoimmune conditions and other liver diseases.
- *Anti-nuclear cytoplasmic antibodies* (ANCA) are present in primary sclerosing cholangitis.

Markers of liver fibrosis

Fibrosis plays a key role in the outcome of certain chronic liver diseases, particularly non-alcoholic fatty liver disease (NAFLD) and chronic hepatitis C disease. Blood tests are being used to decrease the reliance on liver biopsy as the definitive way of detecting fibrosis.

Fibrotest/fibrosure tests measure α_2-macroglobulin, α_2-haptoglobulin, γ–globulin, apoprotein A_1, γ-GT and total bilirubin. These results are formulated to determine a fibrosis index. The index is sensitive and specific (>90%) for the absence of fibrosis, and has 80% sensitivity and specificity for severe fibrosis.

Markers of matrix deposition include procollagen I and III peptide and Type IV collagen. *Markers of matrix degradation*, e.g. matrix metalloproteinase (MMP) 2, 3, 9, and tissue inhibitors of metalloproteinases (TIMPS), e.g. TIMP1 and 2 all are being used as markers of fibrosis.

Genetic analysis

These tests are performed routinely for haemochromatosis (*HFE* gene) and for α_1-antitrypsin deficiency. Markers are also available for Wilson's disease (see p. 357).

Urine tests

Dipstick tests are available for bilirubin and urobilinogen. Bilirubinuria is due to the presence of conjugated (soluble) bilirubin. It is found in the jaundiced patient with hepatobiliary disease; its absence implies that the jaundice is due to increased unconjugated bilirubin. Urobilinogen in the urine is, in practice, of little value but suggests haemolysis or hepatic dysfunction of any cause.

Imaging techniques

Ultrasound examination

This is a non-invasive, safe and relatively cheap technique. It involves the analysis of the reflected ultrasound beam detected by a probe moved across the abdomen. The normal liver appears as a relatively homogeneous structure. The gall bladder, common bile duct, pancreas, portal vein and other structures in the abdomen can be visualized. Abdominal ultrasound is useful in:

- a jaundiced patient (p. 329)
- hepatomegaly/splenomegaly
- the detection of gallstones (Fig. 7.6)
- focal liver disease – lesions >1 cm
- general parenchymal liver disease
- assessing portal and hepatic vein patency
- lymph node enlargement.

Other abdominal masses can be delineated and biopsies obtained under ultrasonic guidance. Colour Doppler ultrasound will demonstrate the vascularity of a lesion and the direction of blood flow in the portal and hepatic veins. Ultrasound contrast agents, most of which are based on the pro-

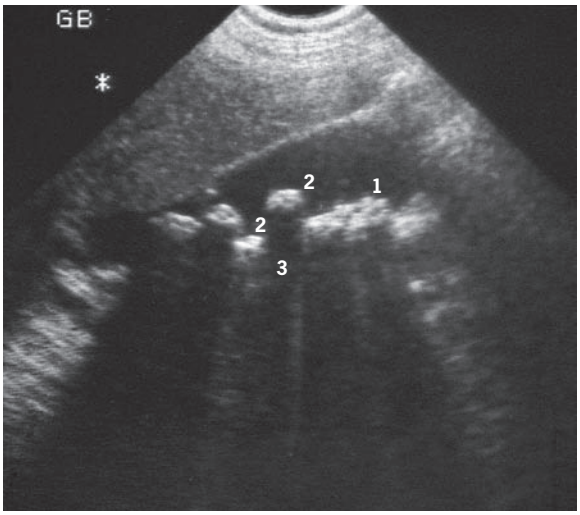


Fig. 7.6 **Gall bladder ultrasound with multiple echogenic gallstones causing well-defined acoustic shadowing.** 1, gall bladder; 2, gallstones; 3, echogenic shadow.

duction of microbubbles within the flowing blood, enhance the vascularity within a lesion, allowing abnormal circulation to be detected within liver nodules, giving a more specific diagnosis of hepatocellular carcinoma.

Hepatic stiffness (elastography). Using an ultrasound transducer, a vibration of low frequency and amplitude is passed through the liver, the velocity of which correlates directly with hepatic stiffness. Stiffness (measured in kPa) increases with increasing liver fibrosis (sensitivity and specificity 80–95% compared to liver biopsy). It cannot be used in the presence of ascites and morbid obesity.

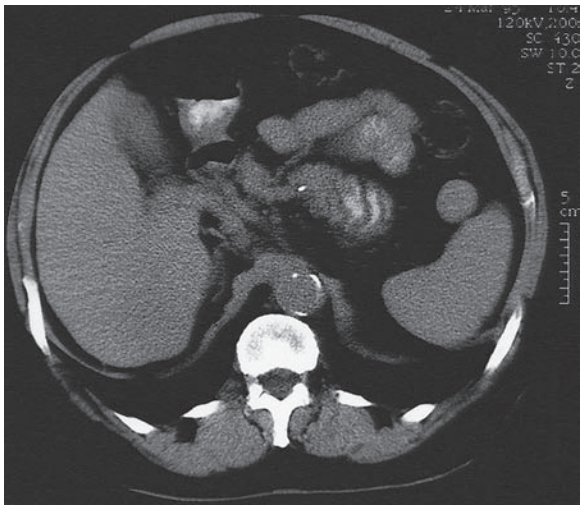
Endoscopic ultrasound (EUS)

In this technique, a small high-frequency ultrasound probe is mounted on the tip of an endoscope and placed by direct vision into the duodenum. The close proximity of the probe to the pancreas and biliary tree permits high-resolution ultrasound imaging. It allows accurate staging of small, potentially operable, pancreatic tumours and offers a less invasive method for bile duct imaging. It has a high accuracy in detection of small neuroendocrine tumours of the pancreas. EUS-guided fine-needle aspiration of tumours provides cytological/histological tissue for confirmation of malignancy. EUS is also used to place transmural tubes to drain pancreatic and peripancreatic fluid collections.

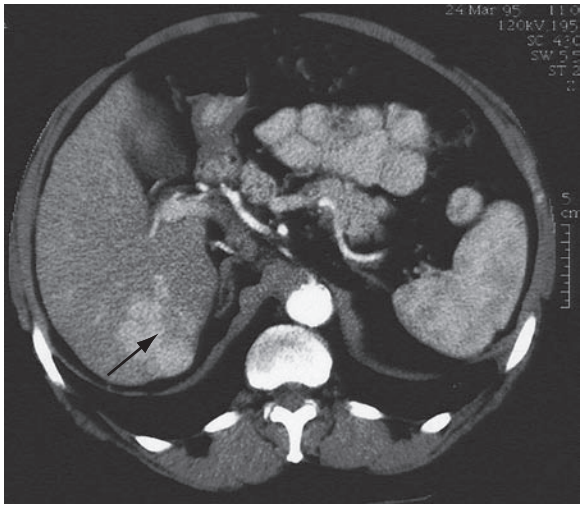
Computed tomography (CT) examination

This technique is complementary to ultrasound, which is usually performed first. It provides excellent visualization of the liver, pancreas, spleen, lymph nodes and lesions in the porta hepatis but involves a high radiation dose. CT allows assessment of the size, shape and density of the liver and can characterize focal lesions in terms of their vascularity. CT is more sensitive in detecting calcification than plain X-rays. Ultrasound is usually more valuable for lesions in the bile duct and gall bladder. CT has advantages in obese subjects.

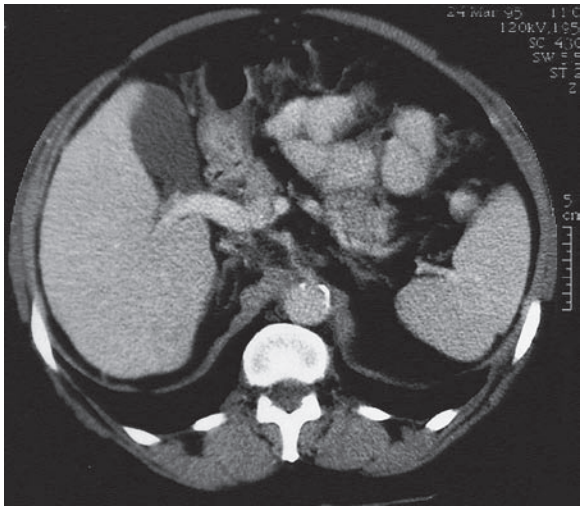
Spiral CT involves rapid acquisition of a volume of data during or immediately after intravenous contrast injection using multiple detectors. Data can thus be acquired in both arterial and portal venous phases of enhancement, enabling more precise characterization of a lesion and its vascular supply (Fig. 7.7). Retrospective analysis of data allows mul-



(a) Unenhanced.



(b) Arterial phase (note high density contrast in the aorta).



(c) Portal venous phase scan through the right lobe of the liver.

Fig. 7.7 **Use of contrast-enhanced spiral CT.** There is an irregular mass (arrow) in the posterior aspect of the right lobe of the liver which is only well seen on the early arterial phase enhanced scan (b).

tiple overlapping slices to be obtained with no increase in the radiation dose. Multi-planar and three-dimensional reconstruction in the arterial phase can create a CT angiogram, often making formal invasive angiography unnecessary. CT also provides guidance for biopsy. In general, lesions over 2 cm can usually be biopsied under ultrasound guidance, which is quicker and more cost-effective.

Magnetic resonance imaging (MRI)

(see also p. 1120)

MRI produces cross-sectional images in any plane within the body and does not involve radiation. Diffuse liver disease alters the T1 and T2 characteristics, and MRI is probably the most sensitive investigation of focal liver disease. Other fat-suppression modes such as STIR allow good differentiation between haemangiomas and other lesions. Contrast agents such as intravenous gadolinium allow further characterization of lesions, are suitable for those with iodine allergy, and provide angiography and venography of the splanchnic circulation. This has superseded direct arteriography.

Magnetic resonance cholangiopancreatography (MRCP)

This technique involves the manipulation of a volume of data acquired by MRI. A heavily T2-weighted sequence enhances visualization of the 'water-filled' bile ducts and pancreatic ducts to produce high-quality images of ductal anatomy. This non-invasive technique is replacing diagnostic (but not therapeutic) ERCP (see below).

Plain X-rays of the abdomen

These are rarely requested but may show:

- gallstones – 10% contain enough calcium to be seen
- air in the biliary tree owing to its recent instrumentation, surgery or to a fistula between the intestine and the gall bladder
- pancreatic calcification
- rarely, calcification of the gall bladder (porcelain gall bladder).

Radionuclide imaging – scintiscanning

In a technetium-99m (^{99m}Tc) colloid scan, the colloid is injected intravenously to be taken up by the reticuloendothelial cells of the liver and spleen. In chronic liver disease there is poor uptake in the liver and most of the colloid is taken up in the spleen and bone marrow. Ultrasound has largely replaced this technique.

In a ^{99m}Tc-IODIDA scan, technetium-labelled iododiethyl IDA is taken up by the hepatocytes and excreted rapidly into the biliary system. Its main uses are in the diagnosis of:

- acute cholecystitis
- jaundice due to either biliary atresia or hepatitis in the neonatal period.

Endoscopy

Upper GI endoscopy is used for the diagnosis and treatment of varices, for the detection of portal hypertensive gastropathy, and for associated lesions such as peptic ulcers. Colonoscopy may show portal hypertensive colopathy.

Endoscopic retrograde cholangiopancreatography (ERCP)

This technique is used to outline the biliary and pancreatic ducts. It involves the passage of an endoscope into the second part of the duodenum and cannulation of the ampulla. Contrast is injected into both systems and the patient is screened radiologically. Contrast medium with a low iodine content of 1.5 mg/mL is used for the common bile duct so that gallstones are not obscured; a higher iodine content of 2.8 mg/mL is used for the pancreatic duct. In addition, other diagnostic and therapeutic procedures can be carried out:

- Common bile duct stones can be removed after a diathermy cut to the sphincter has been performed to facilitate their withdrawal. Sphincterotomy has a serious morbidity rate of 3–5%: acute pancreatitis is the commonest, severe haemorrhage is rare. There is an overall mortality of 0.4%.
- The biliary system can be drained by passing a tube (stent) through an obstruction, or placement of a nasobiliary drain.
- Brachytherapy can be administered after placement at ERCP for therapy of cholangiocarcinoma.

The complication rate in diagnostic ERCP is 2–3%.

A raised serum amylase is often seen and pancreatitis is the most common complication. Cholangitis with or without septicaemia is also seen, and broad-spectrum antibiotics (e.g. 500 mg ciprofloxacin ×2) should be given prophylactically to all patients with suspected biliary obstruction, or a history of cholangitis.

Percutaneous transhepatic cholangiography (PTC)

Under a local anaesthetic, a fine flexible needle is passed into the liver. Contrast is injected slowly until a biliary radicle is identified and then further contrast is injected to outline the whole of the biliary tree. In patients with dilated ducts, the success rate is near 100%. ERCP is the preferred first investigation because therapy (e.g. stone removal) can be undertaken at the same time.

In difficult cases the two techniques are sometimes combined, PTC showing the biliary anatomy above the obstruction, with ERCP showing the more distal anatomy. If an obstruction in the bile ducts is seen, a bypass stent can sometimes be inserted, draining either externally or, for long-term use, internally. Contraindications are as for liver biopsy (see below). The main complications are bleeding and cholangitis with septicaemia, and prophylactic antibiotics should be given as for ERCP.

Angiography

This is performed by selective catheterization of the coeliac axis and hepatic artery. It detects the abnormal vasculature of hepatic tumours, but spiral CT and magnetic resonance angiography (MRA) have replaced this in many cases. The portal vein can be demonstrated with increased definition using subtraction techniques, and splenoportography (by direct splenic puncture) is rarely performed. In digital vascular imaging (DVI), contrast given intravenously or intra-arterially can be detected in the portal system using computerized subtraction analysis. Hepatic venous cannulation allows abnormal hepatic veins to be diagnosed in patients with Budd–Chiari syndrome and also serves as an indirect measurement of portal pressure. There is a 1:1 relationship of occluded (by balloon) hepatic venous pressure with portal pressure in patients with alcoholic or viral-related cirrhosis. The height of portal pressure has been shown to have prognostic value for survival and a reduced portal

pressure 20% from baseline values has been associated with protection from rebleeding. Retrograde CO_2 portography is used when there is doubt about portal vein patency and can be combined with transjugular biopsy and hepatic venous pressure measurement.

Liver biopsy (see Practical Box 7.1)

Histological examination of the liver is valuable in the differential diagnosis of diffuse or localized parenchymal disease. Liver biopsy can be performed on a day-case or overnight-stay basis. The indications and contraindications are shown in Table 7.2. The mortality rate is less than 0.02% when performed by experienced operators.

Practical Box 7.1 Needle biopsy of the liver

This should be performed only by experienced doctors and with sterile precautions. Patient consent must be obtained following explanation of the procedure.

- The patient's coagulation status (prothrombin time, platelets) is checked.
- The patient's blood group is checked and serum saved for crossmatching.
- The patient lies on his back at the edge of the couch.
- The liver margins are delineated using percussion. Alternatively ultrasound examination can be used to confirm liver margins and position of the gall bladder.
- Local anaesthetic is injected at the point of maximum dullness in the mid-axillary line through the intercostal space during expiration. Anaesthetic (1% lidocaine, approximately 5 mL) should be injected down to the liver capsule.
- A tiny cut is made in the skin with a scalpel blade.
- A special needle (Menghini, Trucut or Surecut) is used to obtain the liver biopsy whilst the patient holds his breath in expiration.
- The biopsy is laid on filter paper and placed in 10% formalin. If a culture of the biopsy is required it should be placed in a sterile pot.
- The patient should be observed, with pulse and blood pressure measurements taken regularly for at least 6 h.

Table 7.2 Indications and contraindications for liver biopsy

Indications
Liver disease
Unexplained hepatomegaly
Some cases of jaundice
Persistently abnormal liver biochemistry
Occasionally in acute hepatitis
Chronic hepatitis
Cirrhosis
Drug-related liver disease
Infiltrations
Tumours: primary or secondary
Infections (e.g. tuberculosis)
Storage disease (e.g. glycogen storage)
Pyrexia of unknown origin
Usual contraindications to percutaneous needle biopsy
Uncooperative patient
Prolonged prothrombin time (by more than 3 s)
Platelets $<80\times10^9$/L
Ascites
Extrahepatic cholestasis

Liver biopsy guided by ultrasound or CT is also performed particularly when specific lesions need to be biopsied. Laparoscopy with guided liver biopsy is performed through a small incision in the abdominal wall under local anaesthesia (general anaesthesia is preferred in some centres). A transjugular approach is used when liver histology is essential for management but coagulation abnormalities or ascites prevent the percutaneous approach.

Most complications of liver biopsy occur within 24 hours (usually in the first 2 hours). They are often minor and include abdominal or shoulder pain which settles with analgesics. Minor intraperitoneal bleeding can occur, but this settles spontaneously. Rare complications include major intraperitoneal bleeding, haemothorax and pleurisy, biliary peritonitis, haemobilia and transient septicaemia. Haemobilia produces biliary colic, jaundice and melaena within 3 days of the biopsy.

FURTHER READING

British Society of Gastroenterology. *Guidelines on the use of liver biopsy in clinical practice*, 2004. http://www.bsg.org.uk

Chen J, Raymond K. Nuclear receptors, bile acid detoxification, and cholestasis. *Lancet* 2006; **367**: 454–456.

SYMPTOMS OF LIVER DISEASE

Acute liver disease

This may be asymptomatic and anicteric. Symptomatic disease, which is often viral, produces generalized symptoms of malaise, anorexia and fever. Jaundice may appear as the illness progresses.

Chronic liver disease

Patients may be asymptomatic or complain of non-specific symptoms, particularly fatigue. Specific symptoms include:

- right hypochondrial pain due to liver distension
- abdominal distension due to ascites
- ankle swelling due to fluid retention
- haematemesis and melaena from gastrointestinal haemorrhage
- pruritus due to cholestasis – this is often an early symptom of primary biliary cirrhosis
- breast swelling (gynaecomastia), loss of libido and amenorrhoea due to endocrine dysfunction
- confusion and drowsiness due to neuropsychiatric complications (portosystemic encephalopathy).

SIGNS OF LIVER DISEASE

Acute liver disease

There may be few signs apart from jaundice and an enlarged liver. Jaundice is a yellow coloration of the skin and mucous membranes and is best seen in the conjunctivae and sclerae. In the cholestatic phase of the illness, pale stools and dark urine are present. Spider naevi and liver palms usually indicate chronic disease but they can occur in severe acute disease.

Chronic liver disease

The physical signs are shown in Figure 7.8. However, it is possible for the physical examination to be normal in patients with advanced chronic liver disease.

The skin

The chest and upper body may show *spider naevi*. These are telangiectases that consist of a central arteriole with radiating

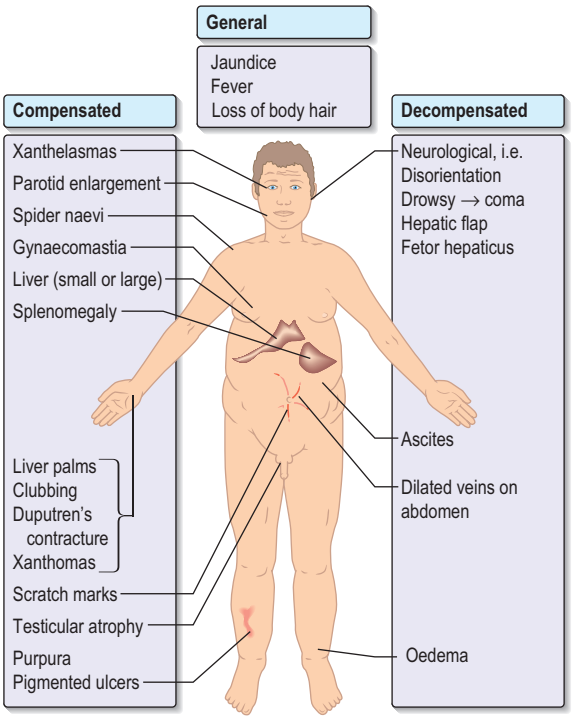


Fig. 7.8 **Physical signs in chronic liver disease.**

small vessels. They are found in the distribution of the superior vena cava (i.e. above the nipple line). They are also found in pregnancy. In haemochromatosis the skin may have a slate-grey appearance.

The hands may show *palmar erythema*, which is a non-specific change indicative of a hyperdynamic circulation; it is also seen in pregnancy, thyrotoxicosis or rheumatoid arthritis. Clubbing occasionally occurs, and a Dupuytren's contracture is often seen in alcoholic cirrhosis.

Xanthomas (cholesterol deposits) are seen in the palmar creases or above the eyes in primary biliary cirrhosis.

The abdomen

Initial hepatomegaly will be followed by a small liver in well-established cirrhosis. Splenomegaly is seen with portal hypertension.

The endocrine system

Gynaecomastia (occasionally unilateral) and testicular atrophy may be found in males. The cause of gynaecomastia is complex, but it is probably related to altered oestrogen metabolism or to treatment with spironolactone.

In decompensated cirrhosis, additional signs that can be seen are shown in Figure 7.8.

JAUNDICE

Jaundice (icterus) is detectable clinically when the serum bilirubin is greater than 50 μmol/L (3 mg/dL). The usual division of jaundice into prehepatic, hepatocellular and obstructive (cholestatic) is an oversimplification as in hepatocellular jaundice there is invariably cholestasis, and the clinical problem is whether the cholestasis is intrahepatic or extrahepatic. Jaundice will therefore be considered under the following headings:

- haemolytic jaundice – increased bilirubin load for the liver cells
- congenital hyperbilirubinaemias – defects in conjugation
- cholestatic jaundice, including hepatocellular (parenchymal) liver disease and large duct obstruction.

Haemolytic jaundice

The increased breakdown of red cells (see p. 403) leads to an increase in production of bilirubin. The resulting jaundice is usually mild (serum bilirubin of 68–102 μmol/L or 4–6 mg/dL) as normal liver function can easily handle the increased bilirubin derived from excess haemolysis. Unconjugated bilirubin is not water-soluble and therefore will not pass into the urine; hence the term 'acholuric jaundice'. Urinary urobilinogen is increased.

The causes of haemolytic jaundice are those of haemolytic anaemia (p. 405). The clinical features depend on the cause; anaemia, jaundice, splenomegaly, gallstones and leg ulcers may be seen.

Investigations show features of haemolysis (p. 403). The level of unconjugated bilirubin is raised but the serum ALP, transferases and albumin are normal. Serum haptoglobulins are low. The differential diagnosis is from other forms of jaundice.

Congenital hyperbilirubinaemias (non-haemolytic)

Unconjugated

Gilbert's syndrome

This is the most common familial hyperbilirubinaemia and affects 2–7% of the population. It is asymptomatic and is usually detected as an incidental finding of a slightly raised bilirubin (17–102 μmol/L or 1–6 mg/dL) on a routine check. All the other liver biochemistry is normal and no signs of liver disease are seen. There is a family history of jaundice in 5–15% of patients. Hepatic glucuronidation is approximately 30% of normal, resulting in an increased proportion of bilirubin monoglucuronide in bile. Most patients have reduced levels of UDP-glucuronosyl transferase (UGT-1) activity, the enzyme that conjugates bilirubin with glucuronic acid. Mutations occur in the gene (*HUG-Br1*) encoding this enzyme, with an expanded nucleotide repeat consisting of two extra bases in the upstream 5′ promoter element. This abnormality appears to be necessary for the syndrome, but is not in itself sufficient for the phenotypic expression of the syndrome.

The major importance of establishing this diagnosis is to inform the patient that this is not a serious disease and to prevent unnecessary investigations. The raised unconjugated bilirubin is diagnostic and rises on fasting and during a mild illness. The reticulocyte count is normal, excluding haemolysis, and no treatment is necessary.

Crigler–Najjar syndrome

This is very rare. Only patients with type II (autosomal dominant) with a decrease rather than absence (type I – autosomal recessive) of UDP-glucuronosyl transferase can survive into adult life. Mutation of the *HUG-Br1* gene for UDP-glucuronosyl transferase has been demonstrated in the coding region. Liver histology is normal. Transplantation is the only effective treatment.

Conjugated

Dubin–Johnson (autosomal recessive) and *Rotor's* (possibly autosomal dominant) *syndromes* are due to defects in bilirubin handling in the liver. The prognosis is good in both. In the Dubin–Johnson syndrome there are mutations in both *MRP2* (p. 322) transporter genes.

The liver is black owing to melanin deposition.

Benign recurrent intrahepatic cholestasis

This is rare and presents in early adulthood. Recurrent attacks of acute cholestasis occur without progression to chronic liver disease. Jaundice, severe pruritus, steatorrhoea and weight loss develop. Serum γ-GT is normal. The gene has been mapped to the FIC1 locus, but the precise relation to cholestasis is unclear. It may be associated with intrahepatic cholestasis of pregnancy (p. 362).

Progressive familial intrahepatic cholestasis syndromes

This is a heterogeneous group of conditions defined by defective secretion of bile acids. They are autosomal recessive. In *type 1*, with cholestasis in infancy (previously known as Byler's disease), γ-GT is normal. The gene is on the familial intrahepatic cholestasis-1 gene (FIC1) locus, but has been mapped to a region encoding P type ATPases (ATP8BI) on chromosome 18q21. *Type 2* has been mapped to the bile salt export pump gene (*BSEP*). The protein is located in the canalicular domain of the plasma membrane of the hepatocyte. The phenotypic expression frequently presents as a non-specific giant cell hepatitis progressing to cholestasis. *Type 3* is due to the multidrug resistance protein 3-P-glycoprotein *PGY3* (MDR-3) gene mutation leading to deficient canalicular phosphatidylcholine transport and thus toxic bile acids causing liver damage. Liver transplantation is the only cure for these syndromes.

FURTHER READING

Jansen PLM, Muller M. The molecular genetics of familial intrahepatic cholestasis. *Gut* 2000; **47**: 1–5.

Cholestatic jaundice (acquired)

This can be divided into extrahepatic and intrahepatic cholestasis. The causes are shown in Figure 7.9.

- Extrahepatic cholestasis is due to large duct obstruction of bile flow at any point in the biliary tract distal to the bile canaliculi.
- Intrahepatic cholestasis occurs owing to failure of bile secretion. A number of cellular mechanisms in cholestasis have been described in animal models, including inhibition of the Na^+/K^+-ATPase in the basolateral membranes, decreased fluidity of the sinusoidal plasma membrane, disruption of the microfilaments responsible for canalicular tone, and damage to the tight junctions. In addition, inflammatory change in ductular cells interferes with bile flow.

Clinically in both types there is jaundice with pale stools and dark urine, and the serum bilirubin is conjugated. However, intrahepatic and extrahepatic cholestatic jaundice must be differentiated as their clinical management is entirely different.

Differential diagnosis of jaundice

The history often gives a clue to the diagnosis. Certain causes of jaundice are more likely in particular categories of people. For example, a young person is more likely to have hepatitis, so questions should be asked about drug and alcohol use, and sexual behaviour. An elderly person with gross weight loss is more likely to have a carcinoma. All patients may complain of malaise. Abdominal pain occurs in patients with biliary obstruction by gallstones, and sometimes with an enlarged liver there is pain resulting from distension of the capsule.

Questions should be appropriate to the particular situation, and the following aspects of the history should be covered.

- *Country of origin*. The incidence of hepatitis B virus (HBV) infection is increased in many parts of the world (p. 335).
- *Duration of illness*. A history of jaundice with prolonged weight loss in an older patient suggests malignancy. A short history, particularly with a prodromal illness of malaise, suggests a hepatitis.
- *Recent outbreak of jaundice*. An outbreak in the community suggests hepatitis A virus (HAV).
- *Recent consumption of shellfish*. This suggests HAV infection.
- *Intravenous drug use, or recent injections or tattoos*. These all increase the chance of HBV and hepatitis C virus (HCV) infection.

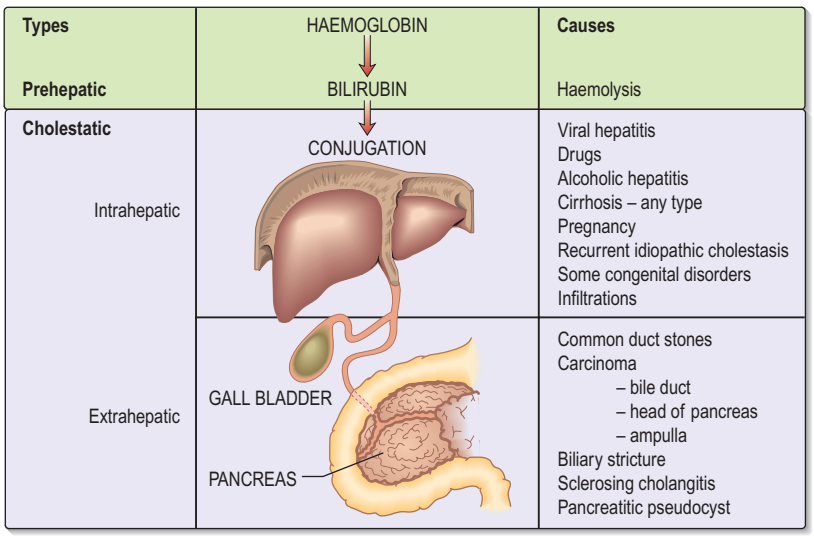


Fig. 7.9 **Causes of jaundice.**

Ask the authors
www.studentconsult.com

- *Men having sex with men*. This increases the chance of HBV infection.
- *Female prostitution*. This increases the chance of HBV infection.
- *Blood transfusion or infusion of pooled blood products*. Increased risk of HBV and HCV. In developed countries all donors are screened for HBV and HCV.
- *Alcohol consumption*. A history of drinking habits should be taken, although many patients often understate their consumption.
- *Drugs taken (particularly in the previous 2–3 months)*. Many drugs cause jaundice (see p. 365).
- *Travel*. Certain areas have an increased risk of HAV infection as well as hepatitis E (HEV) infection (this has a high mortality in pregnancy).
- *A recent anaesthetic*. Halothane (named patient only in UK) and occasionally isoflurane, and the recently introduced sevoflurane may cause jaundice, particularly in those already sensitive to halogenated anaesthetics. The risk with desflurane appears remote.
- *Family history*. Patients with, for example, Gilbert's disease may have family members who get recurrent jaundice.
- *Recent surgery* on the biliary tract or for carcinoma.
- *Environment*. People engaged in recreational activities in rural areas, as well as farm and sewage workers, are at risk for leptospirosis and hepatitis E.
- *Fevers or rigors*. These are suggestive of cholangitis or possibly a liver abscess.

Clinical features

The signs of acute and chronic liver disease should be looked for (p. 329). Certain additional signs may be helpful:

- **Hepatomegaly**. A smooth tender liver is seen in hepatitis and with extrahepatic obstruction, but a knobbly irregular liver suggests metastases. Causes of hepatomegaly are shown in Table 7.3.
- **Splenomegaly**. This indicates portal hypertension in patients when signs of chronic liver disease are present. The spleen can also be 'tipped' occasionally in viral hepatitis.
- **Ascites**. This is found in cirrhosis but can also be due to carcinoma (particularly ovarian) and many other causes (see Table 7.14).

A palpable gall bladder can suggest a carcinoma of the pancreas obstructing the bile duct. Generalized lymphadenopathy suggests a lymphoma.

Cold sores may suggest a herpes simplex virus hepatitis.

Table 7.3 Causes of hepatomegaly

Apparent
Low-lying diaphragm
Reidel's lobe

Cirrhosis (early)

Inflammation
Hepatitis
Schistosomiasis
Abscesses (pyogenic or amoebic)

Cysts
Hydatid
Polycystic

Metabolic
Fatty liver
Amyloid
Glycogen storage disease

Haematological
Leukaemias
Lymphoma
Myeloproliferative disorders
Thalassaemia

Tumours: primary and secondary carcinoma

Venous congestion
Heart failure
Hepatic vein occlusion

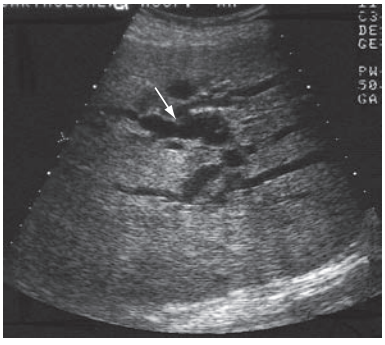
Biliary obstruction (particularly extrahepatic)

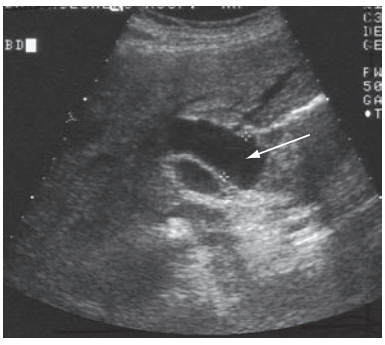
Investigations

Jaundice is not itself a diagnosis and the cause should always be sought. The two most useful tests are the viral markers for HAV, HBV and HCV (in high-risk groups), plus an ultrasound examination. Liver biochemistry confirms the jaundice and may help in the diagnosis.

An ultrasound examination should always be performed to exclude an extrahepatic obstruction, and to diagnose any features compatible with chronic liver disease except when hepatitis A is strongly suspected in a young patient. Ultrasound will demonstrate:

- the size of the bile ducts, which are dilated in extrahepatic obstruction (Fig. 7.10)
- the level of the obstruction


(a)



(b)

Fig. 7.10 **Liver ultrasound showing (a) dilated intrahepatic bile ducts (arrow), (b) common bile duct (arrow).** The normal bile duct measures 6 mm at the porta hepatis.

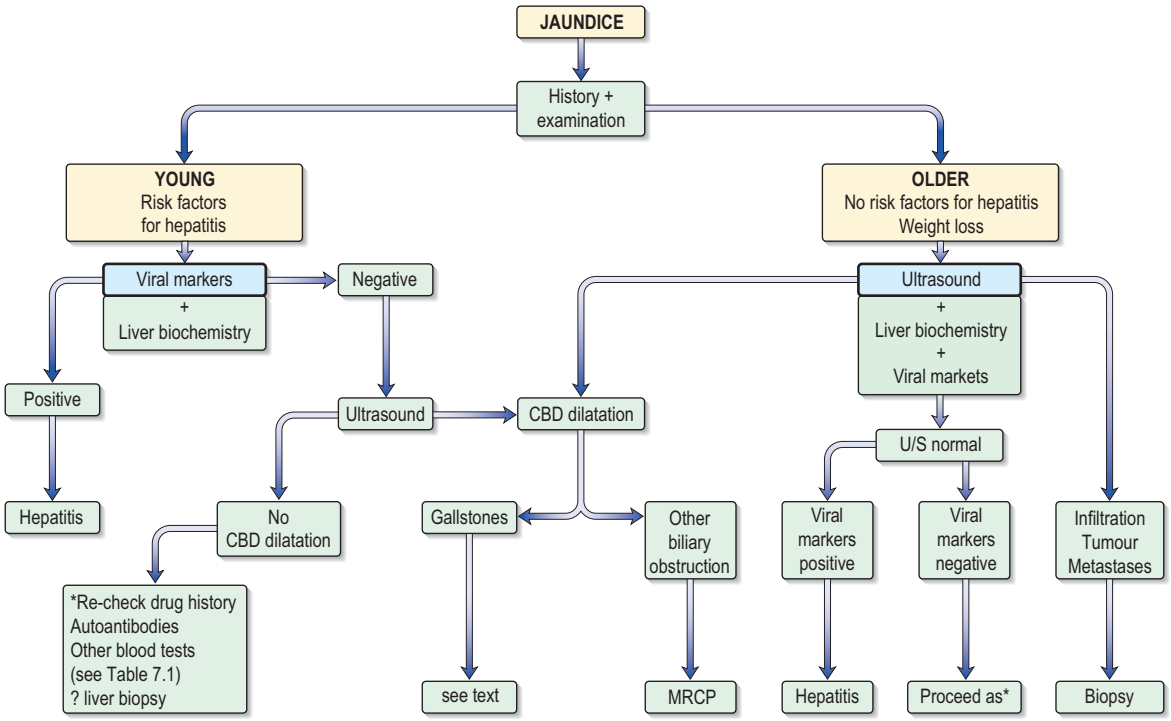


Fig. 7.11 **Approach to patient with jaundice.** CBD, common bile duct; US, ultrasound; MRCP, magnetic resonance cholangiopancreatography.

- the cause of the obstruction in virtually all patients with tumours and in 75% of patients with gallstones.

The pathological diagnosis of any mass lesion can be made by fine-needle aspiration cytology (sensitivity approximately 60%) or by needle biopsy using a spring-loaded device (sensitivity approximately 90%).

A flow diagram for the general investigation of the jaundiced patient is shown in Figure 7.11.

Liver biochemistry

In hepatitis, the serum AST or ALT tends to be high early in the disease with only a small rise in the serum ALP. Conversely, in extrahepatic obstruction the ALP is high with a smaller rise in aminotransferases. These findings cannot, however, be relied on alone to make a diagnosis in an individual case. The prothrombin time (PT) is often prolonged in long-standing liver disease, and the serum albumin is also low.

Haematological tests

In haemolytic jaundice the bilirubin is raised and the other liver biochemistry is normal. A raised white cell count may indicate infection (e.g. cholangitis). A leucopenia often occurs in viral hepatitis, while abnormal mononuclear cells suggest infectious mononucleosis and a Monospot test should be performed.

Other blood tests

These include tests to exclude unusual causes of liver disease (e.g. cytomegalovirus antibodies), autoimmune antibodies, e.g. anti-mitochondrial antibodies (AMA) for the diagnosis of primary biliary cirrhosis, and α-fetoprotein for a hepatocellular carcinoma.

HEPATITIS

Acute parenchymal liver damage can be caused by many agents (Fig. 7.12).

Chronic hepatitis is defined as any hepatitis lasting for 6 months or longer and is often classified according to the aetiology (Table 7.4). Chronic viral hepatitis is the principal cause of chronic liver disease, cirrhosis and hepatocellular carcinoma world-wide.

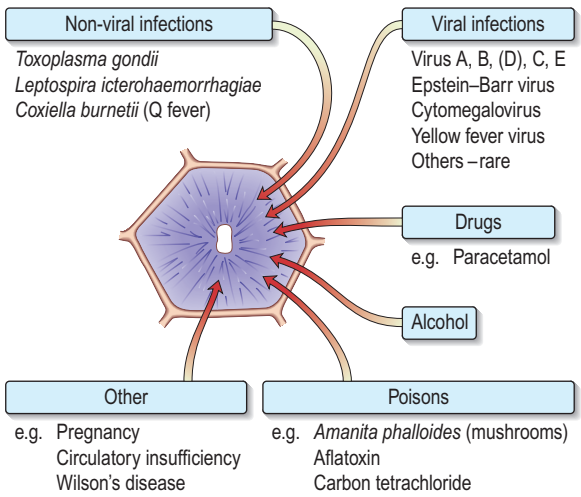


Fig. 7.12 **Some causes of acute parenchymal damage.**

Table 7.4 Causes of chronic hepatitis

Viral
Hepatitis B ± hepatitis D
Hepatitis C
Autoimmune

Drugs
(e.g. methyldopa, isoniazid, ketoconazole, nitrofurantoin)

Hereditary
Wilson's disease

Others
Inflammatory bowel disease – ulcerative colitis
Alcohol (rarely)

Acute hepatitis

Pathology

Although some histological features are suggestive of the aetiological factor, most of the changes are essentially similar whatever the cause. Hepatocytes show degenerative changes (swelling, cytoplasmic granularity, vacuolation), undergo necrosis (becoming shrunken, eosinophilic Councilman bodies) and are rapidly removed. The distribution of these changes varies somewhat with the aetiological agent, but necrosis is usually maximal in zone 3. The extent of the damage is very variable between individuals affected by the same agent: at one end of the spectrum, single and small groups of hepatocytes die (spotty or focal necrosis), while at the other end there is multiacinar necrosis involving a substantial part of the liver (massive hepatic necrosis) resulting in fulminant hepatic failure. Between these extremes there is limited confluent necrosis with collapse of the reticulin framework resulting in linking (bridging) between the central veins, the central veins and portal tracts, and between the portal tracts. The extent of the inflammatory infiltrate is also variable, but portal tracts and lobules are infiltrated mainly by lymphocytes. Other variable features include cholestasis in zone 3 and fatty change, the latter being prominent in hepatitis that is due to alcohol or certain drugs.

Chronic hepatitis

Pathology

Chronic inflammatory cell infiltrates comprising lymphocytes, plasma cells and sometimes lymphoid follicles are usually present in the portal tracts. The amount of inflammation varies from mild to severe. In addition, there may be:

- loss of definition of the portal/periportal limiting plate – interface hepatitis (damage is due to apoptosis rather than necrosis)
- lobular change, focal lytic necrosis, apoptosis and focal inflammation
- confluent necrosis
- fibrosis which may be mild, bridging (across portal tracts) or severe cirrhosis.

The overall severity of the hepatitis is judged by the degree of the hepatitis and inflammation (grading) and the severity of the fibrosis or cirrhosis (staging) using various scoring systems. For example, the *Knodell Scoring System* (histological activity index) uses the sum of four factors (periportal or bridging necrosis, intralobular degeneration and focal necrosis, portal inflammation and fibrosis). The *Ishak score* stages fibrosis from 0 (none) to 6 (cirrhosis). Scoring systems are useful for drug trials and for assessing progression of disease but have limitations.

VIRAL HEPATITIS

The differing features of the common forms of viral hepatitis are summarized in Table 7.5.

HEPATITIS A

Epidemiology

Hepatitis A is the most common type of viral hepatitis occurring world-wide, often in epidemics. The disease is commonly seen in the autumn and affects children and young adults. Spread of infection is mainly by the faeco-oral route and arises from the ingestion of contaminated food or water (e.g. shellfish). Overcrowding and poor sanitation facilitate spread. There is no carrier state. In the UK it is a notifiable disease.

Hepatitis A virus (HAV)

HAV is a picornavirus, having the structure shown in Figure 7.13. It has a single serotype as only one epitope is immunodominant. It replicates in the liver, is excreted in bile and is then excreted in the faeces of infected persons for about 2 weeks before the onset of clinical illness and for up to 7 days after. The disease is maximally infectious just before the onset of jaundice. HAV particles can be demonstrated in the faeces by electron microscopy.

Clinical features

The viraemia causes the patient to feel unwell with nonspecific symptoms that include nausea, anorexia and a distaste for cigarettes. Many recover at this stage and remain anicteric.

After 1 or 2 weeks some patients become jaundiced and symptoms often improve. Persistence of nausea, vomiting or any mental confusion warrants assessment in hospital. As the jaundice deepens, the urine becomes dark and the stools pale owing to intrahepatic cholestasis. The liver is moderately enlarged and the spleen is palpable in about 10% of patients. Occasionally, tender lymphadenopathy is seen, with a transient rash in some cases. Thereafter the jaundice lessens and in the majority of cases the illness is over within 3–6 weeks. Extrahepatic complications are rare but include arthritis, vasculitis, myocarditis and renal failure. A biphasic illness occasionally occurs, with the return of jaundice. Rarely the disease may be very severe with fulminant hepatitis, liver coma and death. The typical sequence of events after HAV exposure is shown in Figure 7.14.

Investigations

Liver biochemistry

Prodromal stage: the serum bilirubin is usually normal. However, there is bilirubinuria and increased urinary urobilinogen. A raised serum AST or ALT, which can sometimes be very high, precedes the jaundice.

Table 7.5 Some features of viral hepatitis

	A	B	D	C	E
Virus	RNA 27 nm	DNA 42 nm	RNA 36 nm (with HBsAg coat)	RNA approx. 50 nm	RNA 27 nm
	Picorna	Hepadna	Deltaviridae	Flavi	Hepevirus
Spread					
Faeco-oral	Yes	No	No	No	Yes
Blood/blood products	Rare	Yes	Yes	Yes	No
Vertical	No	Yes	Rare	Occasional	No
Saliva	Yes	Yes	Yes	? No	?
Sexual	Rare	Yes	Yes (rare)	Uncommon	No
Incubation	Short (2–3 weeks)	Long (1–5 months)	Long	Intermediate	Short
Age	Young	Any	Any	Any	Any
Carrier state	No	Yes	Yes	?	No
Chronic liver disease	No	Yes	Yes	Yes	No
Liver cancer	No	Yes	Rare	Yes	No
Mortality (acute)	<0.5%	<1%		<1%	1–2% (pregnant women10–20%)
Immunization					
Passive	Normal immunoglobulin serum i.m. (0.04–0.06 mL/kg)	Hepatitis B immunoglobulin (HBIg)	No	No	No
Active	Vaccine	Vaccine	HBV vaccine	No	No

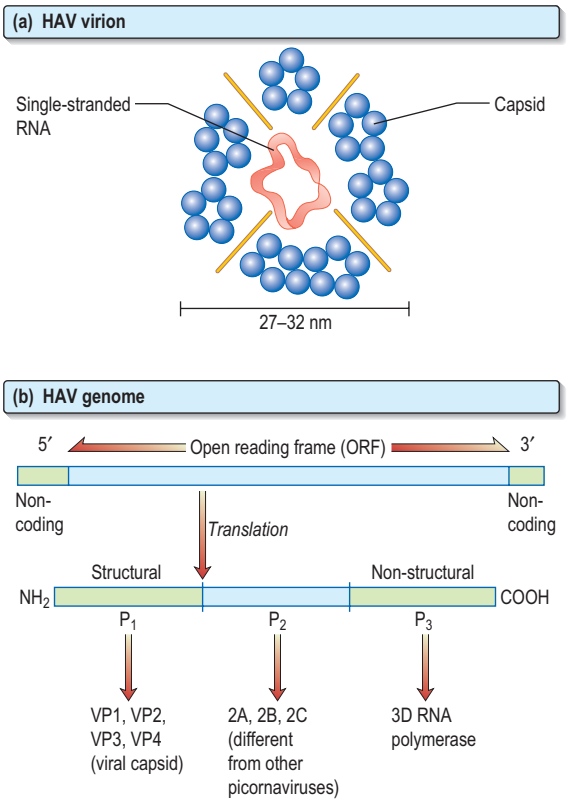


Fig. 7.13 **(a)** The hepatitis A (HAV) virion consists of four polypeptides (VP1–VP4) which form a tight protein shell, or capsid, containing the RNA. The major antigenic component is associated with VP1. **(b)** Arrangement of HAV genome.

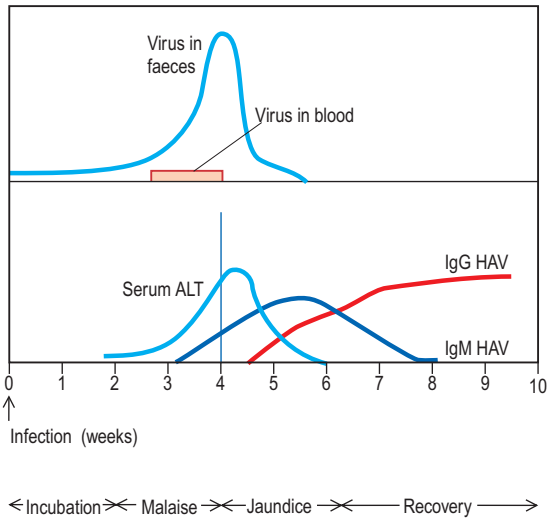


Fig. 7.14 **HAV – sequence of events after exposure.**

Icteric stage; the serum bilirubin reflects the level of jaundice. Serum AST reaches a maximum 1–2 days after the appearance of jaundice, and may rise above 500 IU/L. Serum ALP is usually less than 300 IU/L.

After the jaundice has subsided, the aminotransferases may remain elevated for some weeks and occasionally for up to 6 months.

Haematological tests

There is leucopenia with a relative lymphocytosis. Very rarely there is a Coombs'-positive haemolytic anaemia or an associated aplastic anaemia. The prothrombin time (PT) is pro-

longed in severe cases. The erythrocyte sedimentation rate (ESR) is raised.

Viral markers: antibodies to HAV

IgG antibodies are common in the general population over the age of 50 years, but an anti-HAV IgM means an acute infection. In areas of high prevalence most children have antibodies by the age of 3 years following asymptomatic infection.

Other tests

Further tests are not necessary in the presence of an IgM antibody, but liver biochemistry must be followed to establish a return to normal levels.

Differential diagnosis is from all other causes of jaundice – but in particular from other types of viral and drug-induced hepatitis.

Course and prognosis

The prognosis is excellent, with most patients making a complete recovery. The mortality in young adults is 0.1% but it increases with age. Death is due to fulminant hepatic necrosis. During convalescence, 5–15% of patients may have relapse of the hepatitis but this settles spontaneously. Occasionally a more severe jaundice with cholestasis will run a prolonged course of 7–20 weeks and is called 'cholestatic viral hepatitis'.

There is no reason to stop alcohol consumption other than for the few weeks when the patient is ill. Patients may complain of debility for several months following resolution of the symptoms and biochemical parameters. This is known as the post-hepatitis syndrome; it is a functional illness. Treatment is by reassurance. HAV hepatitis never progresses to chronic liver disease.

Treatment

There is no specific treatment, and rest and dietary measures are unhelpful. Corticosteroids have no benefit. Admission to hospital is not usually necessary.

Prevention and prophylaxis

Control of hepatitis depends on good hygiene. The virus is resistant to chlorination but is killed by boiling water for 10 minutes.

Active immunization

A formaldehyde-inactivated HAV vaccine is given to people travelling frequently to endemic areas, patients with chronic liver disease, people with haemophilia, and workers in frequent contact with hepatitis cases (e.g. in residential institutions for patients with learning difficulties). Community outbreaks can be interrupted by vaccination. A single dose produces antibodies that persist for at least 1 year, with immunity lasting beyond 10 years. This obviates the need for a booster injection in healthy individuals. Universal vaccination has been suggested.

Passive immunization

Normal human immunoglobulin (0.02 mL/kg i.m.) is used if exposure to HAV is <2 weeks. HAV vaccine should also be given.

HEPATITIS B

Epidemiology

The hepatitis B virus (HBV) is present world-wide with an estimated 360 million carriers. The UK and the USA have a low carrier rate (0.5–2%), but it rises to 10–20% in parts of Africa, the Middle and the Far East.

Vertical transmission from mother to child in utero, during parturition or soon after birth, is the usual means of transmission world-wide. This is related to the HBV replicative state of the mother (90% HbeAg+, 30% HbeAg–ve) and is uncommon in Africa where horizontal transmission (sib to sib) is common. HBV is not transmitted by breast feeding.

Horizontal transmission occurs particularly in children through minor abrasions or close contact with other children, and HBV can survive on household articles, e.g. toys, toothbrushes, for prolonged periods so transmission may be possible.

HBV spread also occurs by the intravenous route (e.g. by transfusion of infected blood or blood products, or by contaminated needles used by drug users, tattooists or acupuncturists), or by close personal contact, such as during sexual intercourse, particularly in men having sex with men (25% of cases in the USA). The virus can be found in semen and saliva.

Hepatitis B virus (HBV)

The complete infective virion or Dane particle is a 42 nm particle comprising an inner core or nucleocapsid (27 nm) surrounded by an outer envelope of surface protein (HBsAg). This surface coat is produced in excess by the infected hepatocytes and can exist separately from the whole virion in serum and body fluid as 22 nm particles or 22 nm tubules.

HBsAg contains a major 'a' antigenic determinant as well as several subtypes: 'd', 'y', 'w' and 'r'. Combinations of these subdeterminants (e.g. adr, adw, ayw and ayr) are used to classify HBV genotypes A–H, of which the main types are type A (35%), B (22%), C (31%) and D (10%). There is a strong correlation between genotypes and geographical areas. Genotype A is mainly seen in North West Europe, North America and Central Africa; B in South East Asia (including China, Taiwan and Japan); genotype C in South East Asia; D in southern Europe, India and the Middle East; E in West Africa; F in South and Central America, in American Indians and in Polynesia; G in France and the USA; and H in Central and South America. These genotypes have a bearing on, for example, the time to HBeAg seroconversion (B < C), response to interferon treatment (A > B; C > D) and chronic liver disease (A < D).

The core or nucleocapsid is formed of core protein (HBcAg) containing incompletely double-stranded circular DNA and DNA polymerase/reverse transcriptase. One strand is almost a complete circle and contains overlapping genes that encode both structural proteins (pre-S, surface (S), core (C)) and replicative proteins (polymerase and X). The other strand is variable in length. DR1 and DR2 are direct repeats necessary for HBV synthesis during viral replication (Fig. 7.15).

HBeAg is a protein formed via specific self-cleavage of the pre-core/core gene product which is secreted separately by the cell.

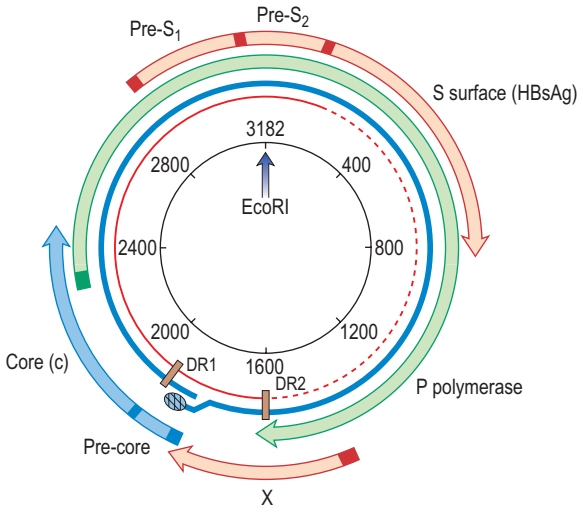


Fig. 7.15 **Hepatitis B virus (HBV) genome.** The viral DNA is partially double-stranded (red incomplete circle and blue circle). The long strand (blue) encodes seven proteins from four overlapping reading frames (S, surface (Pre-S_1, Pre-S_2, S); C, core (Pre-C, C) P, polymerase (P) and X gene (X)). EcoRI restriction-enzyme-binding site is included as a reference point. DR, direct repeats.

Table 7.6 HBV proteins

HBV protein	Significance
Core	Protein of core particle; kinase activity (role in replication?)
Pre-core (HBeAg)	Pre-core/core cleaves to HBeAg; good marker of active replication and role in inducing immunotolerance
Surface (HBsAg)	Envelope protein of HBV; basis of current vaccine
Pre-S_2 / Pre-S_1	HBV binding and entry into hepatocytes
Polymerase	Viral replication
X protein	Transcriptional and transactivator activity

Hepatitis B mutants

Mutations occur in the various reading frames of the HBV genome (see Fig. 7.15). These mutants can emerge in patients with chronic HBV infection (escape mutants) or can be acquired by infection.

HBsAg mutants are produced by alterations in the 'a' determinants of the HbsAg proteins with usually a substitution of glycine for arginine at position 145. This results in changes in the antibody binding domain and the usual tests for HBsAg may be affected.

A mutation in the *pre-core region* when a guanosine (G) to adenosine (A) change creates a stop codon (TAG) prevents the production of HBeAg, but the synthesis of HBcAg is unaffected. To detect infectivity, HBV-DNA must always be measured as no eAg will be present.

DNA polymerase mutants occur, particularly after lamivudine therapy.

Pathogenesis

Pre-S_1 and pre-S_2 regions are involved in attachment to an unknown receptor on the hepatocyte. After penetration into the cell, the virus loses its coat and the virus core is transported to the nucleus without processing. The transcription of HBV into mRNA takes place by the HBV DNA being converted into a closed circular form (Yc DNA), which acts as a template for RNA transcription.

Translation into HBV proteins (Table 7.6) as well as replication of the genome takes place in the endoplasmic reticulum; they are then packaged together and exported from the cell. There is an excess production of non-infective HbsAg particles which are extruded into the circulation.

The HBV is not directly cytopathic and the liver damage produced is by the cellular immune response of the host.

Cytotoxic T cells recognize the viral antigen via HLA class I molecules on the infected hepatocytes; Th1 responses (interleukin-2, gamma-interferon) are thought to be associated with the clearance of the virus and Th2 (interleukins 4, 5, 6, 10, 13) responses with the development of chronic infection and severity of the disease. Viral persistence in patients with a very poor cell-mediated response leads to a healthy inactive chronic HBV infective state. A better response, however, results in continuing hepatocellular damage with the development of chronic hepatitis.

Chronic HBV infection goes through a replicative and an integrated phase. In the *replicative phase* there is active viral replication with hepatic inflammation and the patient is highly infectious with HBeAg and HBV DNA positivity. At some stage the viral genome becomes *integrated* into the host DNA and the viral genes are then transcribed along with those of the host. At this stage, the level of HBV DNA in the serum is low and the patient is HBeAg negative and HBe antibody positive. The aminotransferases are now normal or only slightly elevated and liver histology shows little inflammation, often with cirrhosis. Hepatocellular carcinoma (HCC) develops in patients with this late-stage disease, but the mechanism is still unclear. The REVEAL Study (Risk Evaluation of Viral load Elevation and Associated Liver disease) showed that the risk of HCC was related to levels of HBV DNA rather than a raised aminotransferase (ALT). Integration of the viral DNA with the host-cell chromosomal DNA does appear to have a major role in carcinogenesis. There is evidence to implicate inactivation of *p53*-induced apoptosis by protein X (Table 7.6), allowing accumulation of abnormal cells and, eventually, carcinogenesis.

Clinical features of acute hepatitis

The sequence of events following acute HBV infection is shown in Figure 7.16. In many cases, however, the infection is subclinical. When HBV infection is acquired perinatally, an acute hepatitis usually does not occur as there is a high level of immunological tolerance and the virus persists in over 90%. If there is an acute clinical episode the virus is cleared in approximately 99% of patients as there is a good immune reaction. The clinical picture is the same as that found in HAV infection, although the illness may be more severe. In addition, a serum sickness-like immunological syndrome may be seen. This consists of rashes (e.g. urticaria or a maculopapular rash) and polyarthritis affecting small joints occurring in up to 25% of cases in the prodromal period. Fever is usual. Extrahepatic immune complex-mediated conditions such as an arteritis or glomerulonephritis are occasionally seen.

Investigations

These are generally the same as for hepatitis A.

Specific tests. The markers for HBV are shown in Table 7.7. HBsAg is looked for initially; if it is found, a full viral profile

is then performed. In acute infection, as HBsAg may be cleared rapidly, anti-HBc IgM is diagnostic. HBV DNA is the most sensitive index of viral replication. HBV DNA has been shown to persist (using polymerase chain reaction (PCR) techniques) even when the e antibody has developed.

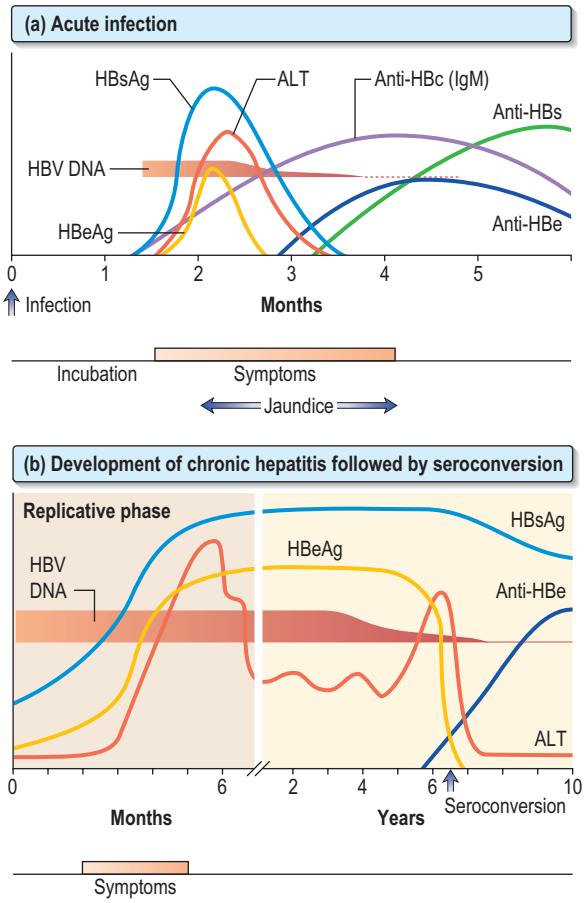


Fig. 7.16 **Time course of the events and serological changes seen following infection with hepatitis B virus.**

(a) Acute infection

Antigens

HBsAg appears in the blood from about 6 weeks to 3 months after an acute infection and then disappears.

HBeAg rises early and usually declines rapidly.

Antibodies

Anti-HBs appears late and indicates immunity.

Anti-HBc is the first antibody to appear and high titres of IgM anti-HBc suggest an acute and continuing viral replication. It persists for many months. IgM anti-HBc may be the only serological indicator of recent HBV infection in a period when HBsAg has disappeared and anti-HBs is not detectable in the serum.

Anti-HBe appears after the anti-HBc and its appearance relates to a decreased infectivity, i.e. a low risk.

(b) Development of chronic hepatitis followed by seroconversion

HBsAg persists and indicates a chronic infection (or carrier state).

HBeAg persists and correlates with increased severity and infectivity and the development of chronic liver disease. When anti-HBe develops (seroconversion) the Ag disappears and there is a rise in ALT.

HBV DNA suggests continual viral replication. For **mutants** – see text.

Table 7.7 Significance of viral markers in hepatitis B

Antigens	
HBsAg	Acute or chronic infection
HBeAg	Acute hepatitis B Persistence implies: continued infectious state development of chronicity increased severity of disease
HBV DNA	Implies viral replication Found in serum and liver
Antibodies	
Anti-HBs	Immunity to HBV; previous exposure; vaccination
Anti-HBe	Seroconversion
Anti-HBc	
IgM	Acute hepatitis B (high titre) Chronic hepatitis B (low titre)
IgG	Past exposure to hepatitis B (HBsAg-negative)

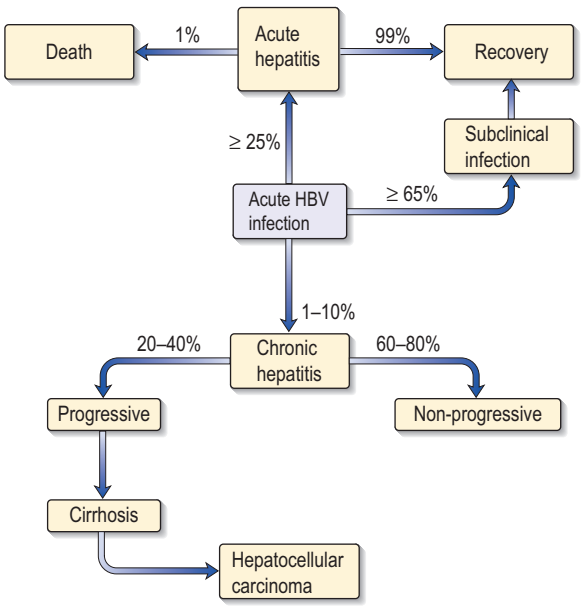


Fig. 7.17 **Clinical course of hepatitis B infection acquired in adults.**

Course

The majority of patients recover completely, fulminant hepatitis occurring in up to 1%. Some patients go on to develop chronic hepatitis (p. 338), cirrhosis (p. 345) and hepatocellular carcinoma (p. 363) or have inactive chronic HBV infection (Fig. 7.17). The outcome depends upon several factors, including the virulence of the virus and the immunocompetence and age of the patient. Some genetic factors, e.g. the presence of MHC class II genotype, may alter host defence to HBV. Complete eradication of HBV infection is probably rare.

Treatment for acute hepatitis

This is mainly symptomatic. Antiviral agents, e.g. lamivudine, have been used in very ill patients but there are no controlled data on their efficacy. However, patients should have their HBV markers monitored and many experts suggest that nucleoside analogues should be given for the persistent presence of HbeAg beyond 12 weeks.

Prevention and prophylaxis

Prevention depends on avoiding risk factors (see above). These include not sharing needles and having safe sex. Vertical transmission is discussed below. Infectivity is highest in those with the e antigen and/or HBV DNA in their blood. These patients should be counselled about their infection. In developing countries, blood and blood products are still a hazard. Standard safety precautions in laboratories and hospitals must be enforced strictly to avoid accidental needle punctures and contact with infected body fluids.

Passive and active immunization

Vaccination is universally available in most developed countries as well as countries with high endemicity. Groups at high risk are: all healthcare personnel; members of emergency and rescue teams; morticians and embalmers; children in high-risk areas; people with haemophilia; patients in some psychiatric units; patients with chronic kidney disease/on dialysis units; long-term travellers; homosexual and bisexual men and prostitutes; intravenous drug users. Routine vaccination is available in the USA.

Combined prophylaxis (i.e. vaccination and immunoglobulin) should be given to: staff with accidental needle-stick injury; all newborn babies of HBsAg-positive mothers; regular sexual partners of HBsAg-positive patients, who have been found to be HBV-negative.

For adults a dose of 500 IU of specific hepatitis B immunoglobulin (HBIG) (200 IU to newborns) is given and the vaccine (i.m.) is given at another site.

Active immunization

This is with a recombinant yeast vaccine produced by insertion of a plasmid containing the gene of HBsAg into a yeast.

Dosage regimen. Three injections (at 0, 1 and 6 months) are given into the deltoid muscle; this gives short-term protection in over 90% of patients. People who are over 50 years of age or clinically ill and/or immunocompromised (including those with HIV infection or AIDS) have a poor antibody response; more frequent and larger doses are required. Antibody levels should be measured at 7–9 months after the initial dose in all at-risk groups. Antibody levels fall steadily after vaccination and booster doses may be required after approximately 3–5 years. It is not cost-effective to check antibody levels prior to active immunization. There are few side-effects from the vaccine; soreness at the site of injection may occur, with very occasionally a fever, rash or a 'flu-like' illness.

Chronic HBV infection

Inactive chronic HBV infection

Following an acute HBV infection, which may be subclinical, approximately 1–10% of patients will not clear the virus and will develop inactive chronic HBV infection. This occurs more readily with neonatal (90%) or childhood (20–50% below the age of 5 years) infection than when HBV is acquired in adult life (<10%) (see Fig. 7.17). There is a vast geographical variation in the incidence of inactive chronic hepatitis B. In the UK, patients are usually discovered incidentally on blood tests, such as when they are screened for donating blood for transfusion or when attending genital medicine or antenatal clinics. Patients with inactive chronic hepatitis have HBsAg in their serum and are HBeAg negative, HBe antibody positive with lower levels of HBV DNA in their serum. They have no evidence of active liver disease (normal or only slightly raised transferases) and are not highly infective. Most remain HBsAg positive, but do not develop progressive liver disease, although some patients have episodes of reactive hepatitis. This develops when the immune balance changes and lymphocytes recognize infected hepatocytes causing hepatitis. There is an annual spontaneous clearance rate of HBsAg of 1–2%.

Active chronic HBV infection

These patients have raised serum aminotransferases with evidence of HBV replication; they have HBe antigen or HBV DNA in their serum. A liver biopsy shows chronic hepatitis.

Clinical features and investigations of active chronic hepatitis

Chronic hepatitis is more frequent in men and it is often not preceded by an acute attack. The condition may be asymptomatic or may present as a mild, slowly progressive hepatitis; 50% present with established chronic liver disease. Clinical relapses or 'flares' occur, sometimes associated with seroconversion (see below) of HBeAg to anti-HBe or vice versa.

Investigations show a moderate rise in aminotransferases and a slightly raised ALP. The serum bilirubin is often normal. HBsAg and HBV DNA are found in the serum, usually with HBe antigen, unless a mutant virus is involved (see p. 336).

Histologically, there is a full spectrum of changes from near normal with only a few lymphocytes and interface hepatitis to a full-blown cirrhosis. HBsAg may be seen as a 'ground-glass' appearance in the cytoplasm on haematoxylin and eosin staining, and this can be confirmed on orcein staining or more specifically with immunohistochemical staining. HBcAg can also be demonstrated in hepatocytes by appropriate immunohistochemical staining.

Treatment for chronic hepatitis

- Patients **positive for HBeAG**, HBsAg and HBV DNA (>20 000 iu/mL – for copies multiply by 5.6) with abnormal serum ALT (>2× normal) should be treated. Liver biopsy is not essential in this group.
- Similar patients with HBV DNA levels >20 000 iu/L who have ALT values <2× normal, should not be treated because the efficacy of therapy is low. A liver biopsy may help as treatment is sometimes given if inflammation is marked.
- Patients who are **negative for HBeAg** with a HBV DNA level of >20 000 should also be treated. Those with DNA levels >2000 should have a liver biopsy to help with the decision to treat or not.

The aim of treatment is the seroconversion of HBeAg (when present) and the reduction of HBV DNA to undetectable levels by PCR. In addition normalization of the serum ALT level and histological improvement in inflammation and fibrosis reflects a good response.

- Patients with **compensated cirrhosis** with HBV DNA >2000 iu/L should be treated and those with a DNA <2000 should only be treated if the ALT is high.
- Patients with **decompensated cirrhosis** can also be treated but liver transplantation may be required.

Table 7.8 Factors predictive of a sustained response to treatment in patients with chronic hepatitis B

	Chronic hepatitis B
Duration of disease	Short
Liver biochemistry	High serum aminotransferase concentrations
Histology	Active liver disease (mild to moderate)
Viral levels	Low HBV DNA levels
Other	Absence of immunosuppression Female gender Adult acquired Delta virus negative Rapidity of response to oral therapy

In patients in whom HBeAg disappears, remission is usually sustained for many years. The patient remains a carrier with HBsAg present, although some will eventually become HBsAg negative.

Antiviral agents

Interferon, lamivudine, adefovir, entecavir and tenofovir are the most commonly used drugs (p. 100). Response to therapy is judged by the HBV DNA level.

Pegylated α-2a interferon (180 μg once a week subcutaneously) gives response rates of 25–45% (depending on genotype – A and B respond best) after 12 months of treatment. Patients with concomitant HIV respond poorly and those with cirrhosis should not receive interferon.

Side-effects of treatment are many, with an acute flu-like illness occurring 6–8 hours after the first injection. This usually disappears after subsequent injections, but malaise, headaches and myalgia are common and depression, reversible hair loss and bone marrow depression and infection may occur. The platelet count should be monitored. These drug reactions occur in up to 30% of patients, and the dose may have to be lowered.

Overall, the *response rate* (i.e. disappearance of HbeAg) is 25–40%. The success rate depends on factors shown in Table 7.8.

In older patients HbeAg usually disappears (sometimes due to changes in the viral genome). Many such patients have inactive disease but some reactivate without regaining HbeAg (HbeAg negative disease). They respond poorly to interferon but can be treated with nucleoside and nucleotide analogues, i.e.:

- *Lamivudine* 100 mg/day given orally, is well tolerated. However, by 4 years, 80% develop viral resistance due to YMDD mutant (tyrosine (Y), methionine (M), aspartate (D)), which itself causes hepatitis. Lamivudine monotherapy is no longer recommended.
- *Entecavir* is more effective than lamivudine and reduces HBV DNA more quickly, and there is much less viral resistance. Serum HBV DNA becomes negative in 67% (HBeAg-positive patients) and 90% (HBeAg-negative patients) by 48 weeks.
- *Tenofovir* 300 mg per day is effective against lamivudine-resistant virus and, as monotherapy, has a similar potency to entecavir.

The duration of all treatment, and which combination of antivirals is optimal, is still being assessed.

In general, current treatment for HbeAg positive disease is with pegylated interferon or tenofovir or entecavir. Interferon should be given for 1 year. The oral antiviral agents should be used long term. HbeAg negative disease requires long-term (many years) oral therapy.

Prognosis

The clinical course of hepatitis B is very variable and treatments have improved survival. Progression from the acute to the chronic phase depends on the age at which infection is acquired. Established cirrhosis is associated with a poor prognosis. Hepatocellular carcinoma is a frequent association and is one of the most common carcinomas in HBV-endemic areas such as the Far East. This incidence is being reduced by routine HBV vaccination of all children.

HEPATITIS D

This is caused by the hepatitis D virus (HDV or delta virus) which is an incomplete RNA particle enclosed in a shell of HbsAg and belongs to the delta viridae family. It is unable to replicate on its own but is activated by the presence of HBV. It is particularly seen in intravenous drug users but can affect all risk groups for HBV infection. Hepatitis D viral infection can occur either as a co-infection with HBV or as a superinfection in an HBsAg-positive patient.

Co-infection of HDV and HBV is clinically indistinguishable from an acute icteric HBV infection, but a biphasic rise of serum aminotransferases may be seen. *Diagnosis* is confirmed by finding serum IgM anti-HDV in the presence of IgM anti-HBc. IgM anti-delta appears at 1 week and disappears by 5–6 weeks (occasionally 12 weeks) when serum IgG anti-delta is seen. The HDV RNA is an early marker of infection. The infection may be transient but the clinical course is variable.

Superinfection results in an acute flare-up of previously quiescent chronic HBV infection. A rise in serum AST or ALT may be the only indication of infection. *Diagnosis* is by finding HDV RNA or serum IgM anti-HDV at the same time as IgG anti-HBc. Active HBV DNA synthesis is reduced by delta superinfection and patients are usually negative for HBeAg with low HBV DNA.

Fulminant hepatitis can follow both types of infection but is more common after co-infection. HDV RNA in the serum and liver can be measured and is found in acute and chronic HDV infection.

Chronic D hepatitis

This is a relatively infrequent chronic hepatitis, but spontaneous resolution is rare. Between 60% and 70% of patients will develop cirrhosis, and more rapidly than with HBV infection alone. In 15% the disease is rapidly progressive with development of cirrhosis in only a few years. The diagnosis is made by finding anti-delta antibody in a patient with chronic liver disease who is HBsAg positive. It can be confirmed by finding HDV in the liver or HDV RNA in the serum by reverse transcription PCR. *Treatment* for patients with active liver disease (raised ALT levels and/or inflammation on biopsy) is with pegylated α–2a interferon for 12 months, but response is poor. Lamivudine and adefovir appear to be unhelpful; there are no large trials of HDV treatment so optimal therapy is not known.

HEPATITIS C

Epidemiology

The prevalence rate of infection in healthy blood donors is about 0.02% in Northern Europe, 1–3% in Southern Europe, possibly linked to intramuscular injections of vaccines or other medicines, and 6% in Africa, with rates as high as 19% in Egypt owing to parenteral antimony treatment for schistosomiasis. The virus is transmitted by blood and blood products and was common in people with haemophilia treated before screening of blood products was introduced. The incidence in intravenous drug users is high (50–60%). The low rate of hepatitis C (HCV) infection in high-risk groups – such as men who have sex with men, prostitutes and attendees at STI clinics – suggests a limited role for sexual transmission. Vertical transmission from healthy mother to child can occur, but is very rare. Other routes of community-acquired infection (e.g. close contact) are extremely rare. In 20% of cases the exact mode of transmission is unknown. An estimated 240 million people are infected with this virus world-wide.

Hepatitis C virus (HCV)

HCV is a single-stranded RNA virus of the Flaviviridae family. The RNA genome is approximately 10 Kb in length, encoding a polyprotein product consisting of structural (capsid and envelope) and non-structural viral proteins (Fig. 7.18). Comparisons of subgenomic regions, such as E1, NS4 or NS5, have allowed variants to be classified into six genotypes. Variability is distributed throughout the genome with the non-structural gene of different genotypes showing 30–50% nucleotide sequence disparity. Genotypes 1a and 1b account for 70% of cases in the USA and 50% in Europe. There is a rapid change in envelope proteins, making it difficult to develop a vaccine. Antigens from the nucleocapsid regions have been used to develop enzyme-linked immunosorbent assays (ELISA). The current assay, ELISA-3, incorporates antigens NS3, NS4 and NS5 regions.

Clinical features

Most acute infections are asymptomatic, with about 10% of patients having a mild flu-like illness with jaundice and a rise in serum aminotransferases. Most patients will not be diagnosed until they present, years later, with evidence of abnormal transferase values at health checks or with chronic liver disease. Extrahepatic manifestations are seen, including arthritis, glomerulonephritis associated with cryoglobulinaemia, and porphyria cutanea tarda. There is a higher incidence of diabetes, and associations with lichen planus, sicca syndrome and non-Hodgkin's lymphoma.

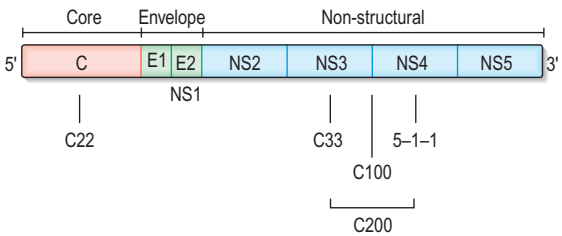


Fig. 7.18 **Hepatitis C virus.** Diagram showing a single-stranded RNA virus with viral proteins. C, core; NS, non-structural; E, envelope.

Diagnosis

This is frequently by exclusion in a high-risk individual with negative markers for HAV, HBV and other viruses. A drug cause for hepatitis should be excluded if possible. HCV RNA can be detected from 1 to 8 weeks after infection. Anti-HCV tests are usually positive 8 weeks from infection.

Treatment

Interferon has been used in acute cases to prevent chronic disease. Needle-stick injuries must be followed and treated early if there is evidence of HCV viraemia, usually re-tested for at 4 weeks.

Course

Eighty five to ninety per cent of asymptomatic patients develop chronic liver disease. A higher percentage of symptomatic patients 'clear' the virus with only 48–75% going on to chronic liver disease (p. 341). Cirrhosis develops in about 20–30% within 10–30 years and of these patients between 7% and 15% will develop hepatocellular carcinoma. The course is adversely affected by co-infection with HBV and/or HIV, and by alcohol consumption, which should be discouraged.

Chronic hepatitis C infection

Pathogenesis

As with hepatitis B infection, cytokines in the Th2 phenotypes are profibrotic and lead to the development of chronic infection. A dominant CD4 Th2 response with a weak CD8 gamma-interferon response may lead to rapid fibrosis. Th1 cytokines are antifibrotic and thus a dominant CD4 Th1 and CD8 cytolytic response may cause less fibrosis. Viral load and viral genotype do not affect rate of fibrosis; persistence of HCV infection has been shown to be associated with HLA-DRB1*0701 and DRB4*0101. Other factors also have an effect on the development of fibrosis, particularly male gender, high alcohol intake, a fatty liver and diabetes.

Clinical features

Patients with chronic hepatitis C infection are usually asymptomatic, the disease being discovered only following a routine biochemical test when mild elevations in the aminotransferases (usually ALT) are noticed (50%). The elevation in ALT may be minimal and fluctuating (Fig. 7.19) and some patients have a persistently normal ALT (25%), the disease being detected by checking HCV antibodies (e.g. in blood donors).

Severe chronic hepatitis (25%) and even cirrhosis can be present with only minimal elevation in aminotransferases, but progression is very uncommon in those with a persistently normal ALT. Fatigue is the commonest symptom with sometimes nausea, anorexia and weight loss, which do not correlate with disease activity.

Diagnosis

This is made by finding HCV antibody in the serum using third-generation ELISA-3 tests. HCV RNA should be assayed using quantitative HCV-RNA PCR. The viraemia is usually variable; less than 600 000 iu/mL signifies a greater likelihood of response to antiviral therapy.

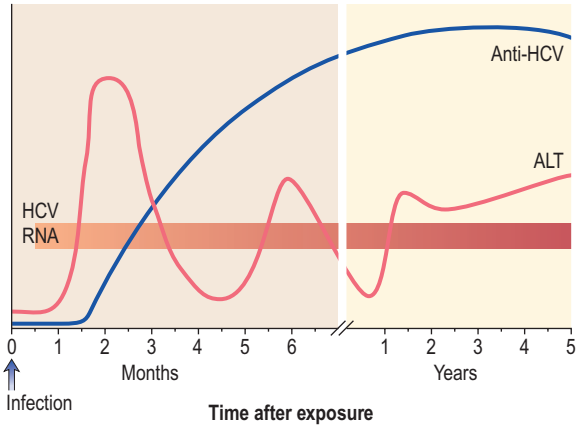


Fig. 7.19 **Time course of the events and serological changes seen following infection with hepatitis C virus.**

The HCV genotype should be characterized in patients who are to be given treatment (see below).

Liver biopsy is indicated if treatment is being considered, especially for genotype 1 and 4. Non-invasive methods for the diagnosis of fibrosis such as serum markers and elastography (p. 326) can replace the need for biopsy in many cases and are useful in follow-up. The changes on liver biopsy are highly variable. Sometimes only minimal inflammation is detected, but in most cases the features of chronic hepatitis are present, as described (p. 333). The METAVIR scoring system grades the inflammation from 0 (none) to 4 (severe), and fibrosis 0 (none) to 4 (cirrhosis) and is used to guide therapy. Lymphoid follicles are often present in the portal tracts, and fatty change is frequently seen.

Treatment (Fig. 7.20)

Treatment is appropriate for patients with chronic hepatitis on liver histology who have HCV RNA in their serum and who have raised serum aminotransferases for more than 6 months. Patients with persistently normal aminotransferases are also treated if they have abnormal histology. The presence of cirrhosis is not a contraindication, but therapeutic responses are less likely. Patients with decompensated cirrhosis should be considered for transplantation. The aim of treatment is to eliminate the HCV RNA from the serum in order to:

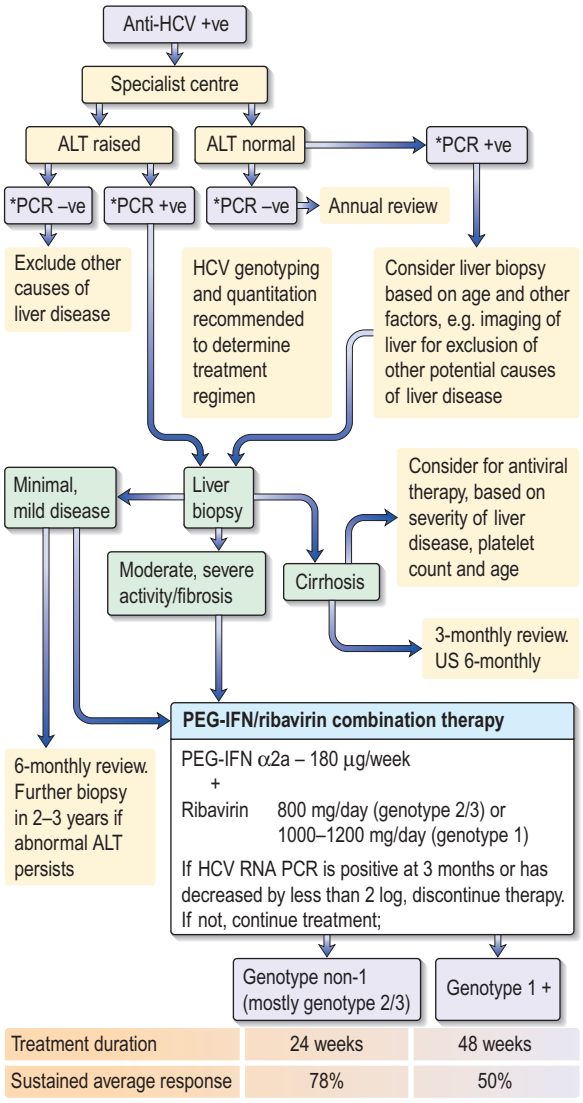
- stop the progression of active liver disease
- prevent the development of hepatocellular carcinoma.

Antiviral agents

Current treatment is combination therapy with pegylated interferon, which is interferon with a polyethyleneglycol tail (α–2a 180 μg/week or α–2b 1.5 μg/kg/week), and ribavirin (1000–1200 mg/day for genotype 1, 800 mg/day for genotype 2 or 3) in daily divided doses for 12 months for genotype 1, and 6 months for other genotypes. Efficacy is also determined by viral load, with HCV RNA >600 000 iu/L less likely to respond.

Side-effects of interferon are described on page 339. Ribavirin is usually well tolerated but side-effects include a dose-related haemolysis, pruritus and nasal congestion. Pregnancy should be avoided as ribavirin is teratogenic.

Monitoring results. An *early virological response* is defined as becoming HCV RNA negative or having at least a 2 log reduction in RNA following treatment in the first 12 weeks.



*PCR HCV RNA

Fig. 7.20 **Approach to a patient with hepatitis C.** ALT, alanine aminotransferase; PCR, polymerase chain reaction. Adapted from Dhumeaux D, Marcellin P, Lerebours E. Treatment of hepatitis C. The 2002 French consensus. *Gut* 2003; **52** with permission from the BMJ Publishing Group.

This occurs in about 80% of patients tolerating full dosage. If this is not achieved, the patient is unlikely to respond and treatment should be stopped. If HCV RNA is undetectable at 4 weeks, treatment for patients with genotype 2 can be stopped at 12–16 weeks and for genotype 1 (if <600 000 iu/mL at baseline) at 6 months.

A *sustained response* is clearance of HCV RNA at 6 months after the end of therapy. It is a good surrogate marker for the resolution of the hepatitis. This is achieved in 40–50% of patients with genotypes 1 and 4 and 80% in genotype 2 or 3. In sustained responders relapse is unlikely and histological progression is halted. Best results are obtained in young patients with low HCV RNA levels and genotype 2 or 3. For non-responders and relapsers, longer duration of therapy is being assessed.

HEPATITIS E

Hepatitis E virus (HEV) is an RNA virus (Hepevirus) (Fig. 7.21) which causes a hepatitis clinically very similar to hepatitis A.

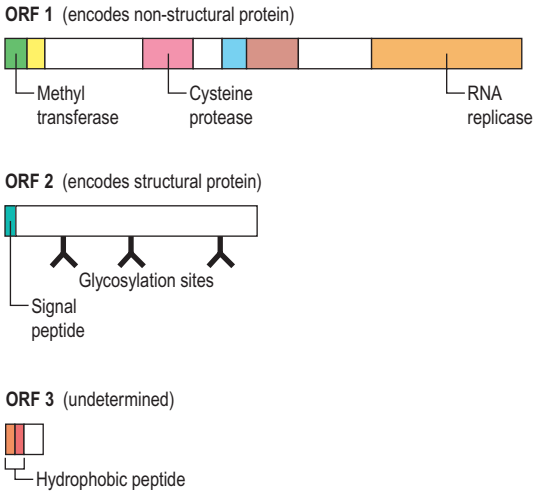


Fig. 7.21 **Hepatitis E genome – showing a single-stranded RNA genome with three open reading frames (ORF).**

It is enterally transmitted, usually by contaminated water, with 30% of dogs, pigs and rodents carrying the virus. Epidemics have been seen in many developing countries. It has a mortality from fulminant hepatic failure of 1–2%, which rises to 20% in pregnant women. There is no carrier state and it does not progress to chronic liver disease. An ELISA for IgG and IgM anti-HEV is available for diagnosis. HEV RNA can be detected in the serum or stools by PCR. Prevention and control depend on good sanitation and hygiene; a vaccine has been developed.

HEPATITIS NON-A–E

Approximately 10–15% of acute viral hepatitides cannot be typed and are described as hepatitis non-A–E. GB agent (HGV hepatitis G virus) and TTV (transfusion-transmitted virus) agents have not been documented as causing disease in humans.

ACUTE HEPATITIS DUE TO OTHER INFECTIOUS AGENTS

Abnormal liver biochemistry is frequently found in a number of acute infections. The abnormalities are usually mild and have no clinical significance.

Infectious mononucleosis (see also p. 107). This is due to the Epstein–Barr (EB) virus. Mild jaundice associated with minor abnormalities of liver biochemistry is extremely common, but 'clinical' hepatitis is rare. Hepatic histological changes occur within 5 days of onset; the sinusoids and portal tracts are infiltrated with large mononuclear cells but the liver architecture is preserved. A Paul–Bunnell or Monospot test is usually positive, and atypical lymphocytes are present in the peripheral blood. Treatment is of the symptoms.

Cytomegalovirus (CMV) (see also p. 107). This can cause acute hepatitis, particularly in a patient with an impaired immune response. The virus may be isolated from the urine. The liver biopsy shows intranuclear inclusions and giant cells.

Yellow fever (see also p. 114). This viral infection is carried by the mosquito *Aedes aegypti* and can cause acute hepatic necrosis. There is no specific treatment.

Herpes simplex (see also p. 104). Very occasionally the herpes simplex virus causes a generalized acute infection, particularly in the immunosuppressed patient, and occasionally in pregnancy. Aminotransferases are usually massively elevated. Liver biopsy shows extensive necrosis. Aciclovir is used for treatment.

Toxoplasmosis (see also p. 160). This produces a clinical picture similar to that of infectious mononucleosis, with abnormal liver biochemistry, but the Paul–Bunnell test is negative.

FURTHER READING

Chen C-J, Yang H-I, Su J et al. Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. *Journal of the American Medical Association* 2006; **295**: 65–73.

Deutsch M, Hadziyannis SJ. Old and emerging therapies in chronic hepatitis C: an update. *Journal of Viral Hepatitis* 2008; **15**(1): 2–11.

Dienstag JL. Hepatitis B viral infection. *New England Journal of Medicine* 2008; **359**: 1486–1500.

Keeffe EB. Future treatment of chronic hepatitis C. *Antiviral Therapy* 2007; **12**(7): 1015–1025.

Keeffe EB, Dieterich DT, Pawlotsky JM et al. Chronic hepatitis B: preventing, detecting, and managing viral resistance. *Clinical Gastroenterology and Hepatology* 2008; **6**(3): 268–274.

Maheshwari A, Ray S, Thuluvath PJ. Acute hepatitis C. *Lancet* 2008; **372**: 321–332.

Papatheodoridis GV, Manolakopoulos S, Dusheiko G et al. Therapeutic strategies in the management of patients with chronic hepatitis B virus infection. *Lancet Infectious Diseases* 2008; **8**(3): 167–178.

Shrestha MP, Scott RM, Joshi DM et al. Safety and efficacy of a recombinant hepatitis E vaccine. *Lancet* 2007; **356**: 895–903.

FULMINANT HEPATIC FAILURE (FHF)

This is defined as severe hepatic failure in which encephalopathy develops in under 2 weeks in a patient with a previously normal liver (occasionally in some patients with previous liver damage; e.g. D virus superinfection in a previous carrier of HBsAg and in Wilson's disease). Cases that evolve at a slower pace (2–12 weeks) are called subacute or subfulminant hepatic failure. FHF is a rare but often life-threatening syndrome that is due to acute hepatitis from many causes (Table 7.9). The causes vary throughout the world; most cases are due to viral hepatitis, but paracetamol overdose is commonly implicated in the UK (50% of cases). HCV does not cause FHF although exceptional cases have been reported from Japan.

Histologically there is multiacinar necrosis involving a substantial part of the liver. Severe fatty change is seen in pregnancy (p. 364), Reye's syndrome (p. 363) or following tetracycline administration intravenously.

Clinical features

Examination shows a jaundiced patient with a small liver and signs of hepatic encephalopathy. The mental state varies

Table 7.9 Causes of fulminant hepatic failure

Viruses
A, B, (D), E

Other viruses

Drugs (examples)
Analgesics (e.g. paracetamol)
Monoamine oxidase inhibitors
Halogenated anaesthetics
Antituberculosis (e.g. isoniazid)
Antiepileptic (e.g. valproate)
'Social' drugs (e.g. 'Ecstasy')

Toxins
Amanita poisoning
Halohydrocarbons

Miscellaneous
Wilson's disease
Acute fatty liver of pregnancy
Reye's syndrome
Budd–Chiari syndrome
Autoimmune hepatitis

Box 7.1 Transfer criteria to specialized units for patients with acute liver injury

- INR >3.0
- Presence of hepatic encephalopathy
- Hypotension after resuscitation with fluid
- Metabolic acidosis
- Prothrombin time (seconds) > interval (hours) from overdose (paracetamol cases)

from slight drowsiness, confusion and disorientation (grades I and II) to unresponsive coma (grade IV) with convulsions. Fetor hepaticus is common, but ascites and splenomegaly are rare. Fever, vomiting, hypotension and hypoglycaemia occur. Neurological examination shows spasticity and hyperreflexia; plantar responses remain flexor until late. Cerebral oedema develops in 80% of patients with FHF but is far less common with subacute failure and its consequences of intracranial hypertension and brain herniation are the most common causes of death. Other complications include bacterial and fungal infections, gastrointestinal bleeding, respiratory arrest, renal failure (hepatorenal syndrome and acute tubular necrosis) and pancreatitis.

Investigations

There is hyperbilirubinaemia, high serum aminotransferases and low levels of coagulation factors, including prothrombin and factor V. Aminotransferases are not useful indicators of the course of the disease as they tend to fall along with the albumin with progressive liver damage. An EEG is sometimes helpful in grading the encephalopathy. Ultrasound will define liver size and may indicate underlying liver pathology.

Treatment

There is no specific treatment, but patients should be managed in a specialized unit. Transfer criteria to such units are shown in Box 7.1. Supportive therapy as for hepatic encephalopathy is necessary (see p. 353). When signs of raised intracranial pressure (which is sometimes measured directly) are present, 20% mannitol (1 g/kg bodyweight) should be infused intravenously; this dose may need to be repeated. Dexamethasone is of no value. Hypoglycaemia, hypokalaemia, hypomagnesaemia, hypophosphataemia and hypocalcaemia should be anticipated and corrected with 10% dextrose infusion (checked by 2-hourly Dextrostix testing), potassium, calcium, phosphate and magnesium. Coagulopathy is managed with intravenous vitamin K, platelets, blood or fresh frozen plasma. Haemorrhage may be a problem and patients are given a proton pump inhibitor (PPI) to prevent gastrointestinal bleeding. Prophylaxis against bacterial and fungal infection is routine, and suspected infection should be treated immediately with suitable antibiotics. Renal and respiratory failure should be treated as necessary. Liver transplantation has been a major advance for patients with FHF. It is difficult to judge the timing or the necessity for transplantation, but there are guidelines based on validated prognostic indices of survival (see below).

Box 7.2 Poor prognostic variables in fulminant hepatic failure indicating a need for liver transplantation

Non-paracetamol (three of following five)

- Drug or non-A–E hepatitis
- Age <10 and >40 years
- Interval from onset of jaundice to encephalopathy >7 days
- Serum bilirubin >300 μmol/L
- Prothrombin time >50 s (or >100 s in isolation)

Paracetamol (acetaminophen)

- Arterial pH <7.3 (after resuscitation, 7.25 on *N*-acetylcysteine)
  Or
- Serum creatinine >300 μmol/L and
- PT >100 s and
- Grade III–IV encephalopathy

Course and prognosis

In mild cases (grades I and II encephalopathy with drowsiness and confusion), two-thirds of the patients will survive. The outcome of severe cases (grades III and IV encephalopathy with stupor or deep coma) is related to the aetiology. In special units, 70% of patients with paracetamol overdosage and grade IV coma survive, as do 30–40% patients with HAV or HBV hepatitis. Poor prognostic variables indicating a need to transplant the liver are shown in Box 7.2.

AUTOIMMUNE HEPATITIS

This condition occurs most frequently in women. In type I (see below) there is an association with other autoimmune diseases (e.g. pernicious anaemia, thyroiditis, coeliac disease and Coombs'-positive haemolytic anaemia), and 60% of cases are associated with HLA, DR3, DR52a loci, HLA-DRB1*0301 and HLA-DRB2*0401. In Asians, the condition is associated with HLA-DR4.

Pathogenesis

The cause is unknown. It is proposed, in a genetically predisposed person, that an environmental agent (perhaps a

virus) causes a sequence of T cell mediated events against liver antigens, producing a progressive necroinflammatory process which results in fibrosis and cirrhosis. In vitro observations have shown that there is a defect of suppressor (regulatory) T cells which may be primary or secondary. However, no clear mechanism causing the inflammation has been found.

Clinical features

There are two peaks in presentation. In the peri- and postmenopausal group, patients may be asymptomatic or present with fatigue, the disease being discovered by abnormalities in liver biochemistry or because of signs of chronic liver disease on routine examination. In the teens and early twenties the disease (often type II) presents as an acute hepatitis with jaundice and very high aminotransferases, which do not improve with time. This age group often has clinical features of cirrhosis with hepatosplenomegaly, cutaneous striae, acne, hirsuties, bruises and, sometimes, ascites. An ill patient can also have features of an autoimmune disease with a fever, migratory polyarthritis, glomerulonephritis, pleurisy, pulmonary infiltration or lung fibrosis.

There are overlap syndromes with primary biliary cirrhosis and primary sclerosing cholangitis, existing concomitantly or developing consecutively.

Investigations

Liver biochemistry

The serum aminotransferases are high, with lesser elevations in the ALP and bilirubin. The serum γ-globulins are high, frequently twice normal, particularly the IgG. The biochemical pattern is similar in both types.

Haematology

A mild normochromic normocytic anaemia with thrombocytopenia and leucopenia is present, even before portal hypertension and splenomegaly. The prothrombin time is often high.

Autoantibodies

Two types of autoimmune hepatitis have been recognized:

- Type I with antibodies (titres >1:80)
  (a) anti-nuclear (ANA)
  (b) anti-smooth muscle (anti-actin)
  (c) soluble liver antigen (anti SLA/LP) – 10–30% cases. This occurs most frequently in girls and young women.
- Type II with antibodies: anti-liver/kidney microsomal (anti-LKM1). The main target is cytochrome P4502D6 (CYP2D6) on liver cell plasma membranes. Approximately 13% of patients lack the above autoantibodies.

Liver biopsy

This shows the changes of chronic hepatitis described previously. The amount of interface hepatitis is variable, but tends to be high in untreated patients. Lymphoid follicles are less often seen than in hepatitis C, and plasma cell infiltration is frequent. Approximately one-third of patients have cirrhosis at presentation.

Treatment

Prednisolone 30 mg is given daily for 2 weeks, followed by a slow reduction and then a maintenance dose of 10–15 mg daily; azathioprine should be added, 1–2 mg/kg daily, as a steroid-sparing agent and in some patients as sole long-term maintenance therapy. Mycophenolate, ciclosporin and tacrolimus have been used in resistant cases.

Course and prognosis

Steroid and azathioprine therapy induce remission in over 80% of cases. Treatment is lifelong in most cases. Those with initial cirrhosis are more likely to relapse following treatment withdrawal and require indefinite therapy. Liver transplantation is performed if treatment fails, although the disease may recur. Hepatocellular carcinoma occurs less frequently than with viral-induced cirrhosis.

DRUG-INDUCED CHRONIC HEPATITIS

Several drugs can cause a CH which clinically bears many similarities to autoimmune hepatitis. Patients are often female, present with jaundice and hepatomegaly, have raised serum aminotransferases and globulin levels, and LE cells and anti-LKM1 antibodies may be detected. Improvement follows drug withdrawal but exacerbations can occur with drug reintroduction. Isoniazid, amiodarone and methotrexate can lead to chronic histological changes. With rare exceptions patients with pre-existing chronic liver disease are not more susceptible to drug injury.

Chronic alcoholic liver disease can occasionally have histological appearances more like a chronic hepatitis.

CHRONIC HEPATITIS OF UNKNOWN CAUSE

As more and more people are having routine blood tests, mild elevations in the serum aminotransferases and γ-GT are found. Many of these patients have no symptoms and no evidence of liver disease clinically. All known aetiological agents should be excluded (see above), and tests carried out to exclude primary biliary cirrhosis, primary sclerosing cholangitis, Wilson's disease, haemochromatosis and α_1-antitrypsin deficiency. Risk factors for NAFLD should be evaluated.

Liver biopsy should be performed if the elevation in the aminotransferases continues for over a year, to confirm the presence of chronic hepatitis, but often non-specific changes are found.

NON-ALCOHOLIC FATTY LIVER DISEASE (NAFLD)

This is an increasingly recognized condition that can lead to cirrhosis (in 1%) and hepatocellular carcinoma. NAFLD is estimated to affect 3–6% of the population in the USA and of these, 1–3% have non-alcoholic steatohepatitis (NASH).

Histological changes are similar to those of alcohol-induced hepatic injury and range from simple fatty change to fat and inflammation (steatohepatitis, NASH) and fibrosis.

FURTHER READING

Czaja AJ. Genetic factors affecting the occurrence, clinical phenotype, and outcome of autoimmune hepatitis. *Clinical Gastroenterology and Hepatology* 2008; **6**(4): 379–388.

Soloway RD, Hewlett AT. The medical treatment for autoimmune hepatitis through corticosteroids to new immunosuppressive agents: a concise review. *Annals of Hepatology* 2007; **6**(4): 204–207.

Drug-induced chronic hepatitis
Chronic hepatitis of unknown cause
Non-alcoholic fatty liver disease
Cirrhosis

Oxidative stress injury and other factors lead to lipid peroxidation in the presence of fatty infiltration and inflammation results. Fibrosis then may occur which is enhanced by insulin resistance, which induces connective tissue growth factor. *Risk factors* for NAFLD are obesity, hypertension, type 2 diabetes and hyperlipidaemia, such that NAFLD is considered the liver component of the metabolic syndrome. Insulin resistance is universal.

Most patients are asymptomatic; hepatomegaly may be present. Mild increases in serum transferases and/or γ-GT (with ALT > AST) are frequently the sole abnormality in the liver biochemistry.

Diagnosis is by demonstration of a fatty liver, usually on ultrasound, with the exclusion of other causes of liver injury, e.g. alcohol. *Liver biopsy* allows staging of the disease but when this should be performed is unclear as there are no definitive guidelines. Most would biopsy if the ALT is persistently over twice normal.

Elastography (p. 326) is being used to evaluate the degree of fibrosis.

Management. Currently weight loss, strict control of hypertension, diabetes and lipid levels are the only treatments. Fatty liver on its own does not progress. Factors indicating progression are unknown, but diabetic patients are most at risk. Thiazolidinediones (p. 1038) which ameliorate insulin resistance, are being evaluated – histological improvement has been documented in short term studies. Liver transplantation is reserved for end-stage cirrhosis, but the condition may recur. Regular follow-up is indicated, particularly for patients with NASH.

CIRRHOSIS

Cirrhosis results from the necrosis of liver cells followed by fibrosis and nodule formation. The liver architecture is diffusely abnormal and this interferes with liver blood flow and function. This derangement produces the clinical features of portal hypertension and impaired liver cell function.

Aetiology

The causes of cirrhosis are shown in Table 7.10. Alcohol is now the most common cause in the West, but viral infection is the most common cause world-wide. With the identification of HCV, and recognition of non-alcoholic fatty liver disease (NAFLD), idiopathic (cryptogenic) cirrhosis is diagnosed infrequently. Young patients with cirrhosis must be investigated carefully as the cause may be treatable (e.g. Wilson's disease).

Table 7.10 Causes of cirrhosis

Common	Others
Alcohol	Biliary cirrhosis:
Hepatitis B ± D	primary
Hepatitis C	secondary
	Autoimmune hepatitis
	Hereditary haemochromatosis
	Hepatic venous congestion
	Budd–Chiari syndrome
	Wilson's disease
	Drugs (e.g. methotrexate)
	α_1-Antitrypsin deficiency
	Cystic fibrosis
	Non-alcoholic fatty liver disease
	Galactosaemia
	Glycogen storage disease
	Veno-occlusive disease
	Idiopathic (cryptogenic)
	? Other viruses

Pathogenesis

Chronic injury to the liver results in inflammation, necrosis and, eventually, fibrosis (Fig. 7.22). Fibrosis is initiated by activation of the stellate cells (see p. 320). Kupffer cells, damaged hepatocytes and activated platelets are probably involved. Stellate cells are activated by many cytokines and their receptors, reactive oxygen intermediates and other paracrine and autocrine signals.

In the early stage of activation the stellate cells become swollen and lose retinoids with upregulation of receptors for proliferative and fibrogenic cytokines, such as platelet-derived growth factor (PDGF), and transforming growth factor β_1 (TGF-β_1). TGF-β_1 is the most potent fibrogenic mediator identified so far. Inflammatory cells contribute to fibrosis via cytokine secretion.

In the space of Disse, the normal matrix is replaced by collagens, predominantly types 1 and 3, and fibronectin. Subendothelial fibrosis leads to loss of the endothelial fenestrations (ports), and this impairs liver function. Collagenases (matrix metalloproteinases, MMPs) are able to degrade this collagen but are inhibited by tissue inhibitors of metalloproteinases (TIMPs), which are increased in human liver fibrosis. There is accumulating evidence that in the early stages liver fibrosis is reversible, particularly when inflammation is reduced, e.g. by suppressing or eliminating viruses.

Pathology

The characteristic features of cirrhosis are regenerating nodules separated by fibrous septa and loss of the normal lobular architecture within the nodules (Fig. 7.23a). Two types of cirrhosis have been described which give clues to the underlying cause:

- *Micronodular cirrhosis*. Regenerating nodules are usually less than 3 mm in size and the liver is involved uniformly. This type is often caused by ongoing alcohol damage or biliary tract disease.

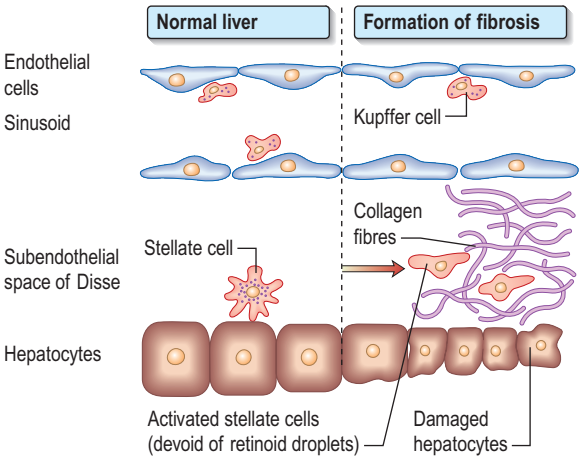
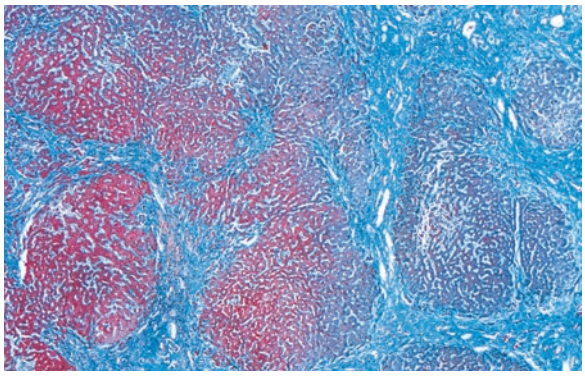


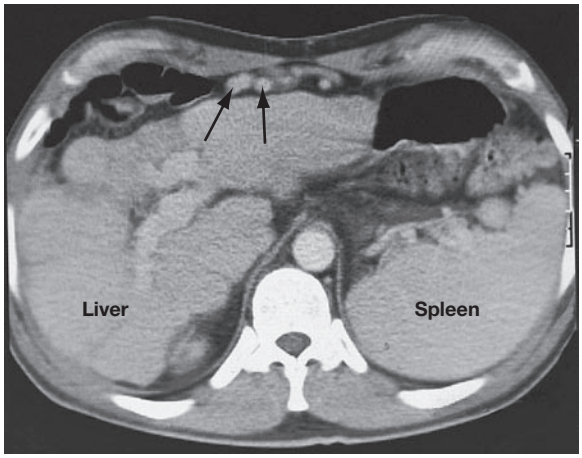
Fig. 7.22 **Pathogenesis of fibrosis.** Activation of the stellate cell is followed by proliferation of fibroblasts and the deposition of collagen.

FURTHER READING

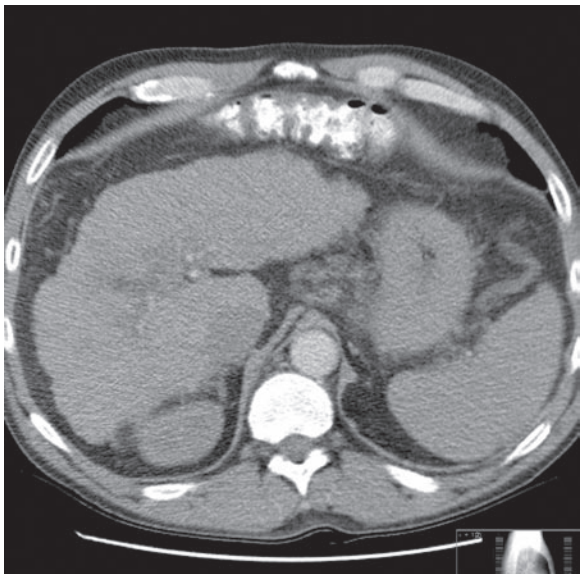
Torres DM, Harrison SA. Diagnosis and therapy of nonalcoholic steatohepatitis. *Gastroenterology* 2008; **134**(6): 1682–1698.



(a)



(b)



(c)

Fig. 7.23 **(a) Pathology of cirrhosis.** Histological appearance showing nodules of liver tissue of varying size surrounded by fibrosis.
(b) CT scan showing an irregular lobulated liver. There is splenomegaly and enlargement of collateral vessels beneath the anterior abdominal wall (arrows) as a result of portal hypertension.
(c) CT image showing cirrhosis, with a patent portal vein and no space-occupying lesion.

- *Macronodular cirrhosis*. The nodules are of variable size and normal acini may be seen within the larger nodules. This type is often seen following chronic viral hepatitis.

A mixed picture with small and large nodules is sometimes seen.

Symptoms and signs are described on page 328.

Investigations

These are performed to assess the severity and type of liver disease.

Severity

- *Liver function*. Serum albumin and prothrombin time are the best indicators of liver function: the outlook is poor with an albumin level below 28 g/L. The prothrombin time is prolonged commensurate with the severity of the liver disease (Box 7.3).
- *Liver biochemistry*. This can be normal, depending on the severity of cirrhosis. In most cases there is at least a slight elevation in the serum ALP and serum aminotransferases. In decompensated cirrhosis all biochemistry is deranged.
- *Serum electrolytes*. A low sodium indicates severe liver disease due to a defect in free water clearance or to excess diuretic therapy.
- *Serum creatinine*. An elevated concentration >130 μmol/L is a marker of worse prognosis.

In addition, serum α-fetoprotein if >200 ng/mL is strongly suggestive of the presence of a hepatocellular carcinoma.

Type

This can be determined by:

- viral markers
- serum autoantibodies
- serum immunoglobulins
- iron indices and ferritin
- copper, caeruloplasmin (p. 357)
- α_1-antitrypsin (p. 358).

Serum copper and serum α_1-antitrypsin should always be measured in young cirrhotics. Total iron-binding capacity (TIBC) and ferritin should be measured to exclude hereditary haemochromatosis; genetic markers are also available (p. 356).

Imaging

- **Ultrasound examination**. This can demonstrate changes in size and shape of the liver. Fatty change and fibrosis produce a diffuse increased echogenicity. In established cirrhosis there may be marginal nodularity of the liver surface and distortion of the arterial vascular architecture. The patency of the portal and hepatic veins can be evaluated. It is useful in detecting hepatocellular carcinoma. *Elastography* is being used in diagnosis and follow-up to avoid liver biopsy (p. 326).
- **CT scan** (see p. 326). Figure 7.23b and c shows hepatosplenomegaly, and dilated collaterals are seen in chronic liver disease. Arterial phase-contrast-enhanced scans are useful in the detection of hepatocellular carcinoma.
- **Endoscopy** is performed for the detection and treatment of varices, and portal hypertensive

www.studentconsult.com
Ask the authors

Box 7.3 Scoring systems in cirrhosis

(a) Modified Child's–Pugh classification

Score	1	2	3
Ascites	None	Mild	Moderate/severe
Encephalopathy	None	Mild	Marked
Bilirubin (μmol/L)	<34	34–50	>50
Albumin (g/L)	>35	28–35	<28
Prothrombin time (seconds over normal)	<4	4–6	>6

Add above scores for your patient for survival figures below

Grade (scores)	% survival		
	1 year	5 years	10 years
Child's A (< 7)	82	45	25
Child's B (7–9)	62	20	7
Child's C (10+)	42	20	0

(b) Model of end-stage liver disease (MELD)

$3.8 \times \text{LN}(\text{bilirubin in mg/dL}) + 9.6 \times \text{LN}(\text{creatinine in mg/dL}) + 11.2 \times \text{LN}(\text{INR}) + 6.4$

To convert:

- bilirubin from μmol/L to mg/dL divide by 17
- creatinine from μmol/L to mg/dL divide by 88.4

LN, natural logarithm; INR, international normalized ratio. MELD scores (with no complications): 1-year survival 97% (score <10); 70% (score 30–40).

gastropathy. Colonoscopy is occasionally performed for colopathy.

- **MRI scan.** This is useful in the diagnosis of benign tumours such as haemangiomas. MR angiography can demonstrate the vascular anatomy and MR cholangiography the biliary tree.

Liver biopsy

This is usually necessary to confirm the severity and type of liver disease. The core of liver often fragments and sampling errors may occur in macronodular cirrhosis. Special stains are required for iron and copper, and various immunocytochemical stains can identify viruses, bile ducts and angiogenic structures. Chemical measurement of iron and copper is necessary to confirm diagnosis of iron overload or Wilson's disease. Adequate samples in terms of length and number of complete portal tracts are necessary for diagnosis and for staging/grading of chronic viral hepatitis.

Management

Management is that of the complications seen in decompensated cirrhosis. Patients should have 6-monthly ultrasound to detect the early development of a hepatocellular carcinoma (see p. 363), as all therapeutic strategies work best with small single tumours.

Treatment of the underlying cause may arrest or occasionally reverse the cirrhotic changes (see below). Patients with compensated cirrhosis should lead a normal life. The only dietary restriction is to reduce salt intake. Aspirin and NSAIDs should be avoided. Alcohol should be avoided, although if the cirrhosis is not due to alcohol and not due to viral hepatitis, small amounts not taken on a regular basis are probably not harmful.

Table 7.11 Poor prognostic indicators in cirrhosis

Blood tests

Low albumin (<28 g/L)

Low serum sodium (<125 mmol/L)

Prolonged prothrombin time >6 seconds above normal value

Raised creatinine >130 μmol/L

Clinical

Persistent jaundice

Failure of response to therapy

Ascites

Haemorrhage from varices, particularly with poor liver function

Neuropsychiatric complications developing with progressive liver failure

Small liver

Persistent hypotension

Aetiology (e.g. alcoholic cirrhosis, if the patient continues drinking)

Course and prognosis

This is extremely variable, depending on many factors, including the aetiology and the presence of complications. Poor prognostic indicators are given in Table 7.11. Development of any complication usually worsens the prognosis. In general, the 5-year survival rate is approximately 50%, but this also varies depending on the aetiology and the stage at which the diagnosis is made.

There are a number of prognostic classifications based on modifications of Child's grading (A, B and C; see Box 7.3) and the model for end-stage disease (MELD), based on serum bilirubin, creatinine and INR, which is widely used as a predictor of mortality in patients awaiting liver transplantation.

LIVER TRANSPLANTATION

This is an established treatment for a number of liver diseases. Shortage of donors is a major problem in all developed countries and in some, such as Japan, living related donors form the majority of transplant operations.

Indications include the following:

Acute liver disease. Patients with fulminant hepatic failure of any cause, including acute viral hepatitis (p. 333), may be considered.

Chronic liver disease. The indications for transplantation are usually for complications of cirrhosis, no longer responsive to therapy. Timing of the transplant depends on donor availability. All patients with end-stage (Child's grade C) cirrhosis should be referred to a transplant centre and also those with debilitating symptoms. In addition specific extrahepatic complications of cirrhosis, even with preserved liver function, such as hepatopulmonary syndrome (shunting in the lung leading to hypoxia) and porto-pulmonary hypertension, can be reversed by liver transplantation.

- *Primary biliary cirrhosis.* Patients with this disease should be transplanted when their serum bilirubin is persistently >100 μmol/L or symptoms such as itching are intolerable.
- *Chronic hepatitis B if HBV DNA negative* or levels falling under therapy. Following transplantation, recurrence of hepatitis is prevented by hepatitis B immunoglobulin

and nucleoside analogues in combination to prevent escape mutants (see p. 336).

- *Chronic hepatitis C* is the most common indication. Universal HCV reinfection occurs with chronic hepatitis of varying severity and cirrhosis occurs in 10–20% at 5 years. Antiviral agents may delay this progression if sustained viral response occurs.
- *Autoimmune hepatitis*. In patients who have failed to respond to medical treatment or have major side-effects of corticosteroid therapy. It can reoccur.
- *Alcoholic liver disease*. Well-motivated patients who have stopped drinking without improvement of liver disease are offered a transplant, with concomitant and frequent counselling before and after transplant.
- *Primary metabolic disorders*. Examples are Wilson's disease, hereditary haemochromatosis and α_1-antitrypsin deficiency.
- *Other conditions*, such as sclerosing cholangitis.

Contraindications

Absolute contraindications include active sepsis outside the hepatobiliary tree, malignancy outside the liver, liver metastases (except neuroendocrine), and if the patient is not psychologically committed.

Relative contraindications are mainly anatomical considerations that would make surgery more difficult, such as extensive splanchnic venous thrombosis. With exceptions, patients aged 65 years or over are not usually transplanted. In hepatocellular carcinoma the recurrence rate is high unless there are fewer than three small (<3 cm) lesions, or a solitary nodule of <5 cm.

Surgical procedure

Pretransplant work-up includes confirmation of the diagnosis, ultrasound and cross-sectional imaging, radiological demonstration of the hepatic arterial and biliary tree as well as assessment of cardiorespiratory and renal status. Because of the ethical and financial implications of this operation, regular psychosocial support is vital, and psychiatric counselling may be necessary in some cases.

The donor should be ABO compatible (but no HLA matching is necessary) and have no evidence of active sepsis, malignancy, HIV, HBV or HCV infection. Younger donors (<50 years) result in better graft function. The liver is cooled and stored on ice; its preservation time can be up to 20 hours. The recipient operation takes approximately 8 hours and may require a large blood transfusion, but sometimes none at all. Cadaveric donor livers may consist of whole graft, split grafts (for two recipients) or reduced grafts from non-heart-beating donors. Live donors may be healthy individuals or patients with, for example, familial amyloid polyneuropathy, whose livers can then be transplanted into others (domino transplant). Right lobe donors have a mortality between 1 in 200 and 1 in 400.

The operative mortality is low. Most postoperative deaths occur in the first 3 months. Sepsis and haemorrhage can be serious complications. Opportunistic infections are still a problem owing to immunosuppression. Various immunosuppressive agents have been used, but microemulsified ciclosporin, tacrolimus in combination with either azathioprine or mycophenolate mofetil, steroids and sirolimus are the most common. A pretransplant serum creatinine above 160 μmol/L (2 mg/dL) is the best predictor of post-transplant death.

Rejection

Acute or cellular rejection is usually seen 5–10 days post-transplant; it can be asymptomatic or there may be a fever. Histologically, there is a pleomorphic portal infiltrate with prominent eosinophils, bile duct damage and endothelialitis of the blood vessels. This type of rejection responds to immunosuppressive therapy.

Chronic ductopenic rejection is seen from 6 weeks to 9 months post-transplant, with disappearing bile ducts (vanishing bile duct syndrome, VBDS) and an arteriopathy with narrowing and occlusion of the arteries. Early ductopenic rejection may rarely be reversed by immunosuppression, but often requires retransplantation.

Graft-versus-host disease is extremely rare.

Prognosis

Elective liver transplantation in low-risk patients has a 90% 1-year survival. Five-year survivals are as high as 70–85% largely owing to the introduction of ciclosporin and tacrolimus. Patients require lifelong immunosuppression, although the doses can be reduced over time without significant problems.

COMPLICATIONS AND EFFECTS OF CIRRHOSIS

These are shown in Table 7.12.

Portal hypertension

The portal vein is formed by the union of the superior mesenteric and splenic veins. The pressure within it is normally 5–8 mmHg with only a small gradient across the liver to the hepatic vein in which blood is returned to the heart via the inferior vena cava. Portal hypertension can be classified according to the site of obstruction:

- *prehepatic* – due to blockage of the portal vein before the liver
- *intrahepatic* – due to distortion of the liver architecture, which can be presinusoidal (e.g. in schistosomiasis) or postsinusoidal (e.g. in cirrhosis)
- *posthepatic* – due to venous blockage outside the liver (rare).

As portal pressure rises above 10–12 mmHg, the compliant venous system dilates and collaterals occur within the systemic venous system. The main sites of the collaterals are at the gastro-oesophageal junction, the rectum, the left renal vein, the diaphragm, the retroperitoneum and the anterior abdominal wall via the umbilical vein.

The collaterals at the gastro-oesophageal junction (varices) are superficial in position and tend to rupture. Por-

Table 7.12 Complications and effects of cirrhosis

Portal hypertension and gastrointestinal haemorrhage
Ascites
Portosystemic encephalopathy
Renal failure
Hepatocellular carcinoma
Bacteraemias, infections
Malnutrition

tosystemic anastomoses at other sites seldom give rise to symptoms. Rectal varices are found frequently (30%) if carefully looked for and can be differentiated from haemorrhoids, which are lower in the anal canal. The microvasculature of the gut becomes congested giving rise to portal hypertensive gastropathy and colopathy, in which there is punctate erythema and sometimes erosions, which can bleed.

Pathophysiology

Portal vascular resistance is increased in chronic liver disease. During liver injury, stellate cells are activated and transform into myofibroblasts. In these cells there is de novo expression of the specific smooth muscle protein α-actin. Under the influence of mediators, such as endothelin, nitric oxide or prostaglandins, the contraction of these activated cells contributes to abnormal blood flow patterns and increased resistance to blood flow. In addition the balance of fibrogenic and fibrolytic factors is shifted towards fibrogenesis. This increased resistance leads to portal hypertension and opening of portosystemic anastomoses in both precirrhotic and cirrhotic livers. Neoangiogenesis also occurs. Patients with cirrhosis have a hyperdynamic circulation. This is thought to be due to the release of mediators, such as nitric oxide and glucagon, which leads to peripheral and splanchnic vasodilatation. This effect is followed by plasma volume expansion due to sodium retention (see the discussion on ascites, p. 351), and this has a significant effect in maintaining portal hypertension.

Causes (see Table 7.13)

The most common cause is cirrhosis. Other causes include the following.

Prehepatic causes

Extrahepatic blockage is due to portal vein thrombosis. The cause is often unidentified, but some cases are due to portal vein occlusion secondary to congenital portal venous abnormalities or neonatal sepsis of the umbilical vein. Many are due to inherited defects causing prothrombotic conditions, e.g. factor V Leiden.

Table 7.13 Causes of portal hypertension

Prehepatic
Portal vein thrombosis
Intrahepatic
Pre-sinusoidal
Schistosomiasis
Sarcoidosis
Primiary biliary cirrhosis
Sinusoidal
Cirrhosis (e.g. alcoholic)
Partial nodular transformation
Post-sinusoidal
Veno-occlusive disease
Budd–Chiari syndrome
Posthepatic
Right heart failure (rare)
Constrictive pericarditis
IVC obstruction

Patients usually present with bleeding, often at a young age. They have normal liver function and, because of this, their prognosis following bleeding is excellent.

The portal vein blockage can be identified by ultrasound with Doppler imaging; CT and MR angiography are also used.

Treatment is usually repeated endoscopic therapy or non-selective beta-blockade. Splenectomy is only performed if there is isolated splenic vein thrombosis. Anticoagulation prevents further thrombosis, and does not increase the risk of bleeding; it is used when there is a high risk of recurrent thrombosis.

Intrahepatic causes

Although cirrhosis is the most common intrahepatic cause of portal hypertension, there are other causes:

- *Non-cirrhotic portal hypertension*. Patients present with portal hypertension and variceal bleeding but without cirrhosis. Histologically, the liver shows mild portal tract fibrosis. The aetiology is unknown, but arsenic, vinyl chloride, antiretroviral therapy and other toxic agents have been implicated. A similar disease is found frequently in India. The liver lesion does not progress and the prognosis is therefore good.
- *Schistosomiasis* with extensive pipe-stem fibrosis is the commonest cause, but is confined to endemic areas such as Egypt and Brazil. However, often there may be concomitant liver disease such as HCV infection.
- Other causes include congenital hepatic fibrosis, nodular regenerative hyperplasia and partial nodular transformation. The last two conditions are rare. They share the common features of hyperplastic liver cell growth in the form of nodules, but in contrast to cirrhosis, fibrosis is typically absent. A wedge liver biopsy is usually required to establish the diagnosis. In none of these conditions are hormones implicated in aetiology or progression.

Posthepatic causes

Prolonged severe heart failure with tricuspid incompetence and constrictive pericarditis can both lead to portal hypertension. The Budd–Chiari syndrome is described on page 359.

Clinical features

Patients with portal hypertension are often asymptomatic and the only clinical evidence of portal hypertension is splenomegaly. Clinical features of chronic liver disease are usually present (see p. 328). Presenting features may include:

- haematemesis or melaena from rupture of gastro-oesophageal varices or portal hypertensive gastropathy
- ascites
- encephalopathy
- breathlessness due to porto-pulmonary hypertension or hepatopulmonary syndrome (rare).

Variceal haemorrhage

Approximately 90% of patients with cirrhosis will develop gastro-oesophageal varices, over 10 years, but only one-third of these will bleed from them. Bleeding is likely to occur with large varices, red signs on varices (diagnosed at endoscopy) and in severe liver disease.

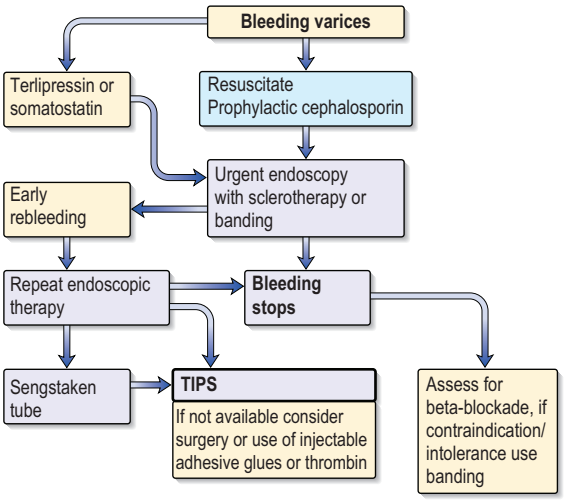
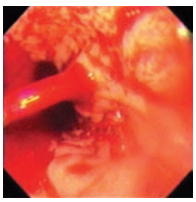


Fig. 7.24 **Management of gastrointestinal haemorrhage due to oesophageal varices.** TIPS, transjugular intrahepatic portosystemic shunt.

(a)



(b)

(c)

(d)

Fig. 7.25 **Endoscopic pictures of oesophageal varices.** **(a)** Gross varices. **(b)** Blood spurting from a varix. **(c)** Following a recent bleed. **(d)** Varices with a band in place (arrow). (a, c and d Courtesy of Dr Peter Fairclough.)

Management

Management can be divided into the active bleeding episode, the prevention of rebleeding, and prophylactic measures to prevent the first haemorrhage. Despite all the therapeutic techniques available, the prognosis depends on the severity of the underlying liver disease, with an overall mortality from variceal haemorrhage of 25%, reaching 50% in Child's grade C.

Initial management of acute variceal bleeding (Fig. 7.24)

See also the discussion of the general management of gastrointestinal haemorrhage on page 267.

Resuscitation

- Assess the general condition of the patient – pulse and blood pressure.
- Insert an intravenous line and obtain blood for grouping and crossmatching, haemoglobin, PT/INR, urea, electrolytes, creatinine, liver biochemistry and blood cultures.
- Restore blood volume with plasma expanders or, if possible, blood transfusion. These measures are discussed in more detail in the treatment of shock (p. 904). Prompt correction of hypovolaemia is necessary in patients with cirrhosis as their baroreceptor reflexes are diminished.
- Ascitic tap.
- Monitor for alcohol withdrawal. Give thiamine i.v.
- Start prophylactic antibiotics – third generation cephalosporins, e.g. cefotaxime. These treat and prevent infection and early rebleeding and reduce mortality.

Urgent endoscopy

Endoscopy should be performed to confirm the diagnosis of varices (Fig. 7.25). It also excludes bleeding from other sites (e.g. gastric ulceration) or portal hypertensive (or congestive) gastropathy. The latter term is used for chronic gastric congestion, punctate erythema and gastric erosions. It is a source of bleeding but varices may or may not be present. Propranolol (see below) is the best treatment for this gastropathy.

Injection sclerotherapy or variceal banding

The varices should be injected with a sclerosing agent that may arrest bleeding by producing vessel thrombosis. A needle is passed down the biopsy channel of the endoscope and a sclerosing agent is injected into the varices. Alternatively, the varices can be banded by mounting a band on the tip of the endoscope, sucking the varix just into the end of the scope and dislodging the band over the varix using a trip-wire mechanism (Fig. 7.25d).

Acute variceal sclerotherapy and banding are the treatment of choice; they arrest bleeding in 80% of cases and reduce early rebleeding. Between 15% and 20% of bleeding comes from gastric varices and here results of sclerotherapy and banding are poor. Injection of tissue glue is preferable.

Other measures available

Vasoconstrictor therapy

The main use of this is for emergency control of bleeding whilst waiting for endoscopy and in combination with endoscopic techniques. The aim of vasoconstrictor agents is to restrict portal inflow by splanchnic arterial constriction.

- *Terlipressin*. This is the only vasoconstrictor shown to reduce mortality. The dose is 2 mg 6-hourly, reducing to 1 mg 4-hourly after 48 hours if a prolonged dosage regimen is used. It should not be given to patients with ischaemic heart disease. The patient will complain of abdominal colic, will defecate and have facial pallor owing to the generalized vasoconstriction.
- *Somatostatin*. This drug has few side-effects. An infusion of 250–500 μg/h appears to reduce bleeding, but has no effect on mortality. It should be used if there are contraindications to terlipressin.

Balloon tamponade

Balloon tamponade is used mainly to control bleeding if endoscopic therapy or vasoconstrictor therapy has failed or is contraindicated or if there is exsanguinating haemorrhage. The usual balloon tube is a four lumen Sengstaken–Blakemore, which should be left in place for no more than 12 hours and removed in the endoscopy room prior to the endoscopic procedure. The tube is passed into the stomach and the gastric balloon is inflated with air and pulled back. It should be positioned in close apposition to the gastro-oesophageal junction to prevent the cephalad variceal blood flow to the bleeding point. The oesophageal balloon should be inflated only if bleeding is not controlled by the gastric balloon alone.

This technique is successful in up to 90% of patients and is very useful in the first few hours of bleeding. However, it can have serious complications such as aspiration pneumonia, oesophageal rupture and mucosal ulceration, which lead to a 5% mortality. The procedure is very unpleasant for the patient. A self-expanding covered metal stent introduced orally and placed over the varices is currently being evaluated, and has the advantage that swallowing is not impaired.

Additional management of acute episode

- *Measures to prevent encephalopathy*. Portosystemic encephalopathy (PSE) can be precipitated by a large

bleed (since blood contains protein). The management is described on page 353.

- *Nursing.* Patients require high-dependency/intensive care nursing. They should be nil by mouth until bleeding has stopped.
- *Sucralfate.* 1 g four times daily is given to reduce oesophageal ulceration following endoscopic therapy.

Management of an acute rebleed

Rebleeding occurs in 20–30% within 5 days after a single session of therapeutic endoscopy. The source of rebleeding should be established by endoscopy. It is sometimes due to a sclerotherapy induced ulcer or slippage of a ligation band. Management starts with repeat endoscopic therapy – once only to control rebleeding (further sessions of sclerotherapy or banding are not advisable).

Transjugular intrahepatic portocaval shunt (TIPS)

TIPS is used when bleeding cannot be stopped after two sessions of endoscopic therapy within 5 days. In this technique, a guidewire is passed from the jugular vein into the liver and an expandable covered metal shunt is placed over it to form a channel between the systemic and portal venous systems. It reduces the hepatic sinusoidal and portal vein pressure by creating a total shunt and does not have the risks of general anaesthesia and major surgery. However there is an increased risk of portal systemic encephalopathy. Recurrent portal hypertension due to stent stenosis or thrombosis is far less frequent with 'covered' compared to 'bare' stents. Collaterals arising from the splenic or portal veins can be selectively embolized.

Emergency surgery

This is used when other measures fail or if TIPS is not available and, particularly, if the rebleeding is from gastric fundal varices. Oesophageal transection and ligation of the feeding vessels to the bleeding varices is the most common surgical technique. Acute portosystemic shunt surgery (see below) is infrequently performed.

Prevention of recurrent variceal bleeding

Following an episode of variceal bleeding, the risk of recurrence is 60–80% over a 2-year period with an approximate mortality of 20% per episode. These facts justify the use of measures to prevent rebleeding.

Long-term measures

Non-selective beta-blockade. Oral propranolol in a dose sufficient to reduce resting pulse rate by 25% has been shown to decrease portal pressure. Portal inflow is reduced by two mechanisms: by a decrease in cardiac output (β_1), and by the blockade of β_2 vasodilator receptors on the splanchnic arteries, leaving an unopposed vasoconstrictor effect. This decreases the frequency of rebleeding, and is as effective as sclerotherapy and ligation as it also prevents bleeding from portal hypertensive gastropathy. It is the treatment of first choice, but a substantial number of patients either have contraindications or are intolerant of treatment. Significant reduction of hepatic venous pressure gradient (HVPG, measured by hepatic vein catheterization) is associated with very low rates or absence of rebleeding, particularly if <12 mmHg. Assessment of HVPG target reduction has prognostic specificity but poor sensitivity and thus poor clinical applicability.

Endoscopic treatment. The use of repeated courses of banding at 2-weekly intervals leads to obliteration of the varices. This markedly reduces rebleeding, most instances occurring before the varices have been fully obliterated. Between 30% and 40% of varices return per year, so follow-up endoscopy with ablation should be performed. Banding is superior to sclerotherapy.

Although a reduction in bleeding episodes occurs, the effect on survival is controversial and probably small. Complications include oesophageal ulceration, mediastinitis and rarely strictures. Combined medical and endoscopic therapy is often used in practice.

Transjugular portosystemic stent shunts. These reduce rebleeding rates compared to endoscopic techniques, but do not improve survival and increase encephalopathy. They are used if endoscopic or medical therapy fails.

Surgical procedures

Surgical portosystemic shunting is associated with an extremely low risk of rebleeding, and is used if TIPS is not available. Hepatic encephalopathy is a significant complication. Operative mortality is low in patients with Child's grade A (0–5%) but rises with worsening liver disease. The 'shunts' performed are usually an end-to-side portocaval anastomosis or a selective distal splenorenal shunt (Warren shunt), which transiently maintains hepatic blood flow via the superior mesenteric vein.

Devascularization procedures including oesophageal transection do not produce encephalopathy, and can be used when there is splanchnic venous thrombosis.

Liver transplantation (p. 347) is the best option when there is poor liver function.

Prophylactic measures

Patients with cirrhosis and varices that have not bled should be prescribed non-selective beta-blockers (e.g. propranolol). This reduces the chances of upper GI bleeding, may increase survival and is cost-effective. If there are contraindications or intolerance, variceal banding is an option. Beta-blockers do not prevent development of varices.

Ascites

Ascites is fluid within the peritoneal cavity and is a common complication of cirrhosis. The pathogenesis of ascites in liver disease is secondary to renal sodium and water retention. Several factors are involved.

- *Sodium and water retention* results from peripheral arterial vasodilatation and consequent reduction in the effective blood volume. Nitric oxide and other substances (e.g. atrial natriuretic peptide and prostaglandins) act as vasodilators. The reduction in effective blood volume activates various neurohumoral pressor systems such as the sympathetic nervous system and the renin–angiotensin system, thus promoting salt and water retention (Fig. 12.3).
- *Portal hypertension* exerts a local hydrostatic pressure and leads to increased hepatic and splanchnic production of lymph and transudation of fluid into the peritoneal cavity.

Box 7.4 The serum–ascites albumin gradient

High serum–ascites albumin gradient (>11 g/L)

Portal hypertension, e.g. hepatic cirrhosis
Hepatic outflow obstruction
Budd–Chiari syndrome
Hepatic veno-occlusive disease
Cardiac ascites
Tricuspid regurgitation
Constrictive pericarditis
Right-sided heart failure

Low serum–ascites albumin gradient (<11 g/L)

Peritoneal carcinomatosis
Peritoneal tuberculosis
Pancreatitis
Nephrotic syndrome

Modified from Chung et al. *New England Journal of Medicine* 2006; **354**: 2166.

- *Low serum albumin* (a consequence of poor synthetic liver function) may further contribute by a reduction in plasma oncotic pressure.

In patients with ascites, urine sodium excretion rarely exceeds 5 mmol in 24 hours. Loss of sodium from extrarenal sites accounts for approximately 30 mmol in 24 hours. The normal daily dietary sodium intake may vary between 120 and 200 mmol, resulting in a positive sodium balance of approximately 90–170 mmol in 24 hours (equivalent to 600–1300 mL of fluid retained).

Clinical features

The abdominal swelling associated with ascites develops over many weeks or as rapidly as a few days. Precipitating factors include a high sodium diet or the development of a hepatocellular carcinoma or splanchnic vein thrombosis. Mild generalized abdominal pain and discomfort are common but, if more severe, should raise the suspicion of spontaneous bacterial peritonitis (see below). Respiratory distress accompanies tense ascites, and also causes difficulty in eating.

The presence of fluid is confirmed by the demonstration of shifting dullness. Many patients also have peripheral oedema. A pleural effusion (usually on the right side) may infrequently be found and arises from the passage of ascitic fluid through congenital diaphragmatic defects.

Investigations

A diagnostic aspiration of 10–20 mL of fluid should be obtained and the following performed:

- **Cell count**. A neutrophil count above 250 cells/mm^3 is indicative of an underlying (usually spontaneous) bacterial peritonitis.
- **Gram stain and culture** – for bacteria and acid-fast bacilli.
- **Protein**. A high serum–ascites albumin gradient of >11 g/L suggests portal hypertension, and a low gradient <11 g/L is associated with abnormalities of the peritoneum, e.g. inflammation, infections, neoplasia (Box 7.4).
- **Cytology** – for malignant cells.
- **Amylase** – to exclude pancreatic ascites.

The differential diagnosis of ascites is listed in Table 7.14.

Table 7.14 Causes of ascites divided according to the type of ascitic fluid

Straw-coloured
Malignancy (most common cause)
Cirrhosis
Infective
  Tuberculosis
  Following intra-abdominal perforation – any bacterium may be found (e.g. *E. coli*)
  Spontaneous in cirrhotics
Hepatic vein obstruction (Budd–Chiari syndrome) protein level high in fluid
Chronic pancreatitis
Congestive cardiac failure
Constrictive pericarditis
Meigs' syndrome (ovarian tumour)
Hypoproteinaemia (e.g. nephrotic syndrome)

Chylous
Obstruction of main lymphatic duct (e.g. by carcinoma) – chylomicrons are present
Cirrhosis

Haemorrhagic
Malignancy
Ruptured ectopic pregnancy
Abdominal trauma
Acute pancreatitis

Management

The aim is to both reduce sodium intake and increase renal excretion of sodium, producing a net reabsorption of fluid from the ascites into the circulating volume. The maximum rate at which ascites can be mobilized is 500–700 mL in 24 hours (see below). The management is as follows:

- Check serum electrolytes and creatinine at the start and every other day; weigh patient and measure urinary output daily.
- Bed rest alone will lead to a diuresis in a small proportion of people by improving renal perfusion, but in practice is not helpful.
- By dietary sodium restriction it is possible to reduce sodium intake to 40 mmol in 24 hours and still maintain an adequate protein and calorie intake with a palatable diet.
- Drugs: many contain significant amounts of sodium (up to 50 mmol daily). Examples include antacids, antibiotics (particularly the penicillins and cephalosporins) and effervescent tablets. Sodium-retaining drugs (non-steroidals, corticosteroids) should be avoided.
- Fluid restriction is probably not necessary unless the serum sodium is under 128 mmol/L (see below).
- The diuretic of first choice is the aldosterone antagonist spironolactone, starting at 100 mg daily. Chronic administration produces gynaecomastia. Eplerenone 25 mg once daily does not cause gynaecomastia.

The aim of diuretic therapy should be to produce a net loss of fluid approaching 700 mL in 24 hours (0.7 kg weight loss or 1.0 kg if peripheral oedema is present). Although 60% of patients respond with this regimen, diuresis is often poor and

the spironolactone can be increased gradually to 500 mg daily providing there is no hyperkalaemia. A loop diuretic, such as furosemide 20–40 mg or bumetanide 1 mg daily, may be added if response is poor. These loop diuretics have several potential disadvantages, including hyponatraemia, hypokalaemia and volume depletion.

Ascitic fluid is mobilized more slowly than interstitial fluid, and diuretics should be given with great care in those without peripheral oedema.

Diuretics should be temporarily discontinued if a rise in serum creatinine level occurs, representing overdiuresis and hypovolaemia, or if there is hyperkalaemia or the development of precoma. Hyponatraemia occurring during therapy almost always represents haemodilution secondary to a failure to clear free water (usually a marker of reduced renal perfusion) and should be treated by stopping the diuretics if the sodium level falls below approximately 128 mmol/L as well as introducing water restriction. Vaptans (p. 663), a class of drugs that increase free water clearance by inhibition of vasopressin receptors, are being evaluated in cirrhosis.

Paracentesis

This is used to relieve symptomatic tense ascites. It is also used as a means of rapid therapy in patients with ascites and peripheral oedema, thus avoiding prolonged hospital stay. The main complication is hypovolaemia and renal dysfunction (post-paracentesis circulatory dysfunction) as the ascites reaccumulates at the expense of the circulating volume; this is more likely with more than 5 litres removal and worse liver function. In patients with normal renal function and without hyponatraemia, this is overcome by infusing albumin (8 g per litre of ascitic fluid removed). In practice, up to 20 L can be removed over 4–6 hours with albumin infusion.

Shunts

A transjugular intrahepatic portosystemic shunt (TIPS) is used for resistant ascites providing there is no spontaneous portosystemic encephalopathy and minimal disturbance of renal function. Frequency of paracentesis and diuretic use is usually reduced. Survival may improve. The use of a peritoneo-venous shunt has been abandoned in most centres due to a high rate of blockage.

Spontaneous bacterial peritonitis (SBP)

This represents a serious complication of ascites with cirrhosis and occurs in approximately 8%. The infecting organisms gain access to the peritoneum by haematogenous spread; most are *Escherichia coli, Klebsiella* or enterococci. The condition should be suspected in any patient with ascites who clinically deteriorates. Features such as pain and pyrexia are frequently absent. Diagnostic aspiration should always be performed (see above). A raised neutrophil count in ascites is alone sufficient evidence to start treatment immediately. A third-generation cephalosporin, such as cefotaxime or ceftazidime, is used and is modified on the basis of culture results. Mortality is 10–15%. Recurrence is common (70% within a year) and an oral quinolone, e.g. norfloxacin 400 mg daily, is given for prevention, prolonging the survival. Primary prophylaxis of SBP in patients with ascites protein <10 g/L or severe liver disease also improves survival.

SBP is an indication to refer to a liver transplant centre.

Portosystemic encephalopathy (PSE)

This is a chronic neuropsychiatric syndrome secondary to cirrhosis. Acute encephalopathy can occur in acute hepatic failure (see p. 342). PSE can occur in portal hypertensive patients due to spontaneous 'shunting', or in those with surgical or TIPS shunts. Encephalopathy is potentially reversible.

Pathogenesis

The mechanism is unknown but several factors are involved. In cirrhosis, the portal blood bypasses the liver via the collaterals and the 'toxic' metabolites pass directly to the brain to produce the encephalopathy.

Many 'toxic' substances may be causative factors, including ammonia, free fatty acids, mercaptans and accumulation of false neurotransmitters (octopamine) or activation of the γ-aminobutyric acid (GABA) inhibitory neurotransmitter system. Increased blood levels of aromatic amino acids (tyrosine and phenylalanine) and reduced branched-chain amino acids (valine, leucine and isoleucine) also occur. Ammonia has a major role; ammonia-induced alteration of brain neurotransmitter balance – especially at the astrocyte–neurone interface – is the leading pathophysiological mechanism. Ammonia is produced by intestinal bacteria breaking down protein. The factors precipitating PSE are shown in Table 7.15.

Clinical features

An acute onset often has a precipitating factor (Table 7.15). The patient becomes increasingly drowsy and comatose.

Chronically, there is a disorder of personality, mood and intellect, with a reversal of normal sleep rhythm. These changes may fluctuate, and a history from a relative must be obtained. The patient is irritable, confused, disorientated and has slow slurred speech. General features include nausea, vomiting and weakness. Coma occurs as the encephalopathy becomes more marked, but there is always hyperreflexia and increased tone. Convulsions are so very rare that other causes must be looked for.

Signs include:

- fetor hepaticus (a sweet smell to the breath)
- a coarse flapping tremor seen when the hands are outstretched and the wrists hyperextended (asterixis)

Table 7.15 Factors precipitating portosystemic encephalopathy

High dietary protein
Gastrointestinal haemorrhage
Constipation
Infection, including spontaneous bacterial peritonitis
Fluid and electrolyte disturbance due to:
diuretic therapy
paracentesis
Drugs (e.g. any CNS depressant)
Portosystemic shunt operations, TIPS
Any surgical procedure
Progressive liver damage
Development of hepatocellular carcinoma

TIPS, transjugular intrahepatic portocaval shunt

- constructional apraxia, with the patient being unable to write or draw, for example, a five-pointed star
- decreased mental function, which can be assessed by using the serial-sevens test (see p. 1189). A trail-making (or connection) test (the ability to join numbers and letters (in chronological order) with a pen within a certain time – a standard psychological test for brain dysfunction) is prolonged.

Diagnosis is clinical. Routine liver biochemistry merely confirms the presence of liver disease, not the presence of encephalopathy.

Additional investigations

- Electroencephalography (EEG) shows a decrease in the frequency of the normal α-waves (8–13 Hz) to δ-waves of 1.5–3 Hz. These changes occur before coma supervenes.
- Visual evoked responses (see p. 1121) also detect subclinical encephalopathy.
- Arterial blood ammonia can be useful for the differential diagnosis of coma and to follow a patient with PSE, but is not readily available.

Management

- Identify and remove the possible precipitating cause, such as drugs with cerebral depressant properties, constipation or electrolyte imbalance due to overdiuresis.
- Give purgation and enemas to empty the bowels of nitrogenous substances. Lactulose (10–30 mL three times daily) is an osmotic purgative that reduces the colonic pH and limits ammonia absorption. Lactilol (β-galactoside sorbitol 30 g daily) is metabolized by colonic bacteria and is comparable in efficacy to lactulose.
- Maintain nutrition with adequate calories, given if necessary via a fine-bore nasogastric tube, and do not restrict protein for more than 48 hours.
- Give antibiotics. Rifaximin is mainly unabsorbed and well tolerated long term. Metronidazole (200 mg four times daily) is also effective in the acute situation. Neomycin should be avoided. Stop or reduce diuretic therapy.
- Give intravenous fluids as necessary (beware of too much sodium).
- Treat any infection.
- Increase protein in the diet to the limit of tolerance as the encephalopathy improves.

Course and prognosis

Acute encephalopathy in acute liver failure has a very poor prognosis as the disease itself has a high mortality. In cirrhosis, chronic PSE is very variable and adversely affects prognosis. Very rarely with chronic portosystemic shunting an organic syndrome with cerebellar signs or choreoathetosis can develop, as well as a myelopathy leading to a spastic paraparesis due to demyelination. Patients should be referred to a liver transplant centre.

Renal failure (hepatorenal syndrome)

The hepatorenal syndrome occurs typically in a patient with advanced cirrhosis, portal hypertension with jaundice and ascites. The urine output is low with a low urinary sodium concentration, a maintained capacity to concentrate urine (i.e. tubular function is intact) and almost normal renal histology. The renal failure is described as 'functional'. It is sometimes precipitated by overvigorous diuretic therapy, NSAIDs, diarrhoea or paracentesis, and infection, particularly spontaneous bacterial peritonitis.

The mechanism is similar to that producing ascites. The initiating factor is thought to be extreme peripheral vasodilatation, possibly due to nitric oxide, leading to an extreme decrease in the effective blood volume and hypotension (p. 351). This activates the homeostatic mechanisms, causing a rise in plasma renin, aldosterone, norepinephrine (noradrenaline) and vasopressin, leading to vasoconstriction of the renal vasculature. There is an increased preglomerular vascular resistance causing the blood flow to be directed away from the renal cortex. This leads to a reduced glomerular filtration rate and plasma renin remains high. Salt and water retention occur with reabsorption of sodium from the renal tubules.

Other mediators have been incriminated in the pathogenesis of the hepatorenal syndrome, in particular the eicosanoids. This has been supported by the precipitation of the syndrome by inhibitors of prostaglandin synthase such as non-steroidal anti-inflammatory drugs (NSAIDs).

Diuretic therapy should be stopped and intravascular hypovolaemia corrected, preferably with albumin. Terlipressin or noradrenaline with intravenous albumin improves renal function in one-third of patients. Liver transplantation is the best option.

Hepatopulmonary syndrome

This is defined as a hypoxaemia occurring in patients with advanced liver disease. It is due to intrapulmonary vascular dilatation with no evidence of primary pulmonary disease. The patients have features of cirrhosis with spider naevi and clubbing as well as cyanosis. Most patients have no respiratory symptoms, but with more severe disease, patients are breathless on standing. Transthoracic ECHO shows intrapulmonary shunting, and arterial blood gases confirm the arterial oxygen desaturation. These changes are improved with liver transplantation.

Porto-pulmonary hypertension

This must be distinguished from the hepatopulmonary syndrome as in this group there is pulmonary hypertension. It occurs in 1–2% of patients with cirrhosis related to portal hypertension. It may respond to medical therapy. Severe pulmonary hypertension is a contraindication for liver transplantation.

Primary hepatocellular carcinoma

This is discussed on page 363.

TYPES OF CIRRHOSIS

Alcoholic cirrhosis

This is discussed in the section on alcoholic liver disease (p. 358).

FURTHER READING

Friedman SL, Bansal MB. Reversal of hepatic fibrosis – fact or fantasy. *Hepatology* 2006; **43**: 582–588.

Rodriguez-Roisin R, Krowka MJ. Hepato-pulmonary syndrome. *New England Journal of Medicine* 2008; **358**: 2378–2387.

Schuppan D, Afdhal NH. Liver cirrhosis. *Lancet* 2008; **371**: 838–851.

Shawcross DL, Jalan R. Dispelling myths in the treatment of hepatic encephalopathy. *Lancet* 2005; **365**: 431–433.

Primary biliary cirrhosis

Primary biliary cirrhosis (PBC) is a chronic disorder in which there is a progressive destruction of bile ducts, eventually leading to cirrhosis. Ninety per cent of those affected are women in the age range 40–50 years. PBC is frequently being diagnosed in its milder forms. The prevalence is approximately 7.5 per 100 000, with a 1–6% increase in first-degree relatives. PBC has been called 'chronic non-suppurative destructive cholangitis'; this term is more descriptive of the early lesion and emphasizes that true cirrhosis occurs only in the later stages of the disease.

Aetiology

The aetiology is unknown, but immunological mechanisms play a part. Serum anti-mitochondrial antibodies (AMA) are found in almost all patients with PBC, and of the mitochondrial proteins involved, the antigen M2 is specific to PBC.

Five M2-specific antigens have been further defined using immunoblot techniques, of which the E2 component of the pyruvate dehydrogenase complex (PDC) is the major M2 autoantigen (72 kDa E2 subunit (PDC,E2)). The presence of AMA in high titre is unrelated to the clinical or histological picture and its role in pathogenesis is unclear. Antibodies against nuclear antigens, e.g. anti gp210, are present in 50% of patients and correlate with progression towards liver failure.

It seems likely that an environmental factor acts on a genetically predisposed host via molecular mimicry initiating autoimmunity. *E. coli* and *N. aromaticivorans* antibodies are present in high titre. Halogenated hydrocarbons mimic the PDC autoepitopes.

Although damage to bile ducts is a feature, antibodies to bile ductules are not specific to PBC. Biliary epithelium from patients with PBC expresses aberrant class II HLAs, but it is not known whether this expression is the cause or result of the inflammatory response. Cell-mediated immunity is impaired (demonstrated both in vitro and by skin testing); cytotoxic $DC4^+$ and $CD8^+$ T lymphocytes directly produce biliary epithelium damage. They recognize the inner lipoyl domain and lipoic acid also recognized by AMA. There is an increased synthesis of IgM, thought to be due to a failure of the switch from IgM to IgG antibody synthesis. No specific associated Class 2 MHC loci have been found.

Asymptomatic patients are discovered on routine examination or screening and may have hepatomegaly, a raised serum alkaline phosphatase or autoantibodies.

Pruritus is often the earliest symptom, preceding jaundice by a few years. Fatigue, which is often disabling, may accompany pruritus, particularly in progressive cases. When jaundice appears, hepatomegaly is usually found. In the later stages, patients are pigmented and jaundiced with severe pruritus. Pigmented xanthelasma on eyelids or other deposits of cholesterol in the creases of the hands may be seen.

Associations

Autoimmune disorders (e.g. Sjögren's syndrome, scleroderma, thyroid disease) occur with increased frequency. Keratoconjunctivitis sicca (dry eyes and mouth) is seen in 70% of cases. Renal tubular acidosis, membranous glomerulonephritis, coeliac disease and interstitial pneumonitis are also associated with PBC.

Investigations

- **Mitochondrial antibodies** – measured routinely by ELISA (in titres >1:160) – are present in over 95% of patients; M2 antibody is specific. Other non-specific antibodies (e.g. anti-nuclear factor and smooth muscle) may also be present.
- **High serum alkaline phosphatase** is often the only abnormality in the liver biochemistry.
- **Serum cholesterol** is raised.
- **Serum IgM** may be very high.
- **Ultrasound** can show a diffuse alteration in liver architecture.
- **Liver biopsy** shows characteristic histological features of a portal tract infiltrate, mainly of lymphocytes and plasma cells; approximately 40% have granulomas. Most of the early changes are in zone 1. Later, there is damage to and loss of small bile ducts with ductular proliferation. Portal tract fibrosis and, eventually, cirrhosis is seen.

Hepatic granulomas are not specific and are also seen in sarcoidosis, tuberculosis, schistosomiasis, drug reactions, brucellosis, parasitic infestation (e.g. strongyloidiasis) and other conditions.

Differential diagnosis

The classical picture presents little difficulty with diagnosis (high serum alkaline phosphatase and the presence of AMA); this can be confirmed by the characteristic histological features although this is not necessary except in doubtful cases. There is a group of patients with the histological changes of PBC but the serology of autoimmune hepatitis. This has been given the name of autoimmune cholangitis and responds to steroids and azathioprine.

In the jaundiced patient, extrahepatic biliary obstruction should be excluded by ultrasound and, if there is doubt about the diagnosis, MRCP (or ERCP) should be performed to make sure that the bile ducts are normal.

Treatment

Ursodeoxycholic acid (10–15 mg/kg) improves bilirubin and aminotransferase levels. It should be given early in the asymptomatic phase. It is not clear if prognosis is altered. Symptoms are not improved. Steroids improve biochemical and histological disease but may lead to increased osteoporosis and other side-effects and should not be used.

Malabsorption of fat-soluble vitamins (A, D and K) occurs and supplementation is required when deficiency is detected and prophylactically in the jaundiced patient. Bisphosphonates are required for osteoporosis. Despite raised serum lipid concentrations there is no increased risk from cardiovascular disease, although this has been recently disputed by one group.

Pruritus is difficult to control, but colestyramine, one 4 g sachet three times daily, can be helpful, although it is unpalatable. Rifampicin, naloxone hydrochloride and naltrexone (opioid antagonists) have been shown to be of benefit. Intractable pruritus can be relieved by plasmapheresis or a molecular absorbent re-circulating system (MARS).

The lack of effective medical therapy has made PBC a major indication for liver transplantation (p. 347).

Complications

The complications are those of cirrhosis. In addition, osteoporosis, and rarely osteomalacia and a polyneuropathy can also occur.

Course and prognosis

This is very variable. Asymptomatic patients and those presenting with pruritus will survive for more than 20 years. Symptomatic patients with jaundice have a more rapidly progressive course and die of liver failure or bleeding varices in approximately 5 years. Liver transplantation should therefore be offered when the serum bilirubin reaches 100 μmol/L. Transplantation has a 5-year survival of at least 80%.

Secondary biliary cirrhosis

Cirrhosis can result from prolonged (for months) large duct biliary obstruction. Causes include bile duct strictures, gallstones and sclerosing cholangitis. An ultrasound examination, followed by ERCP or PTC, is performed to outline the ducts and any remedial cause is dealt with.

Hereditary haemochromatosis

(see also p. 394)

Hereditary haemochromatosis (HH) is an inherited disease characterized by excess iron deposition in various organs leading to eventual fibrosis and functional organ failure.

Prevalence and aetiology

HH is transmitted by an autosomal recessive gene with a prevalence in Caucasians of homozygotes (affected) of 1 in 400, but very variable phenotypic expression and a heterozygote (carrier) frequency of 1 in 10. It is the most common single gene disorder in Caucasians. The most common form of HH is due to a mutation in a gene, *HFE*, on the short arm of chromosome 6.

Between 85% and 90% of patients with overt HH are homozygous for the Cys 282 Tyr (C282Y mutation). A second mutation (His 63 Asp (H63D)) occurs in about 25% of the population and is in complete linkage disequilibrium with Cys 282 Tyr. Another form of haemochromatosis occurs in Southern Europe and is associated with TfR2, a transferrin receptor isoform. Rarer types are ferroportin related, and there is also a juvenile form.

Dietary intakes of iron and chelating agents (ascorbic acid) may be relevant. Iron overload may be present in alcoholics, but alcohol excess per se does not cause HH although there is a history of excess alcohol intake in 25% of patients.

Mechanism of damage. This is still unclear. The *HFE* gene protein interacts with the transferrin receptor 1, which is a mediator in intestinal iron absorption (see Fig. 8.8). Iron is taken up by the mucosal cells inappropriately, exceeding the binding capacity of transferrin. Hepatic expression of the hepcidin gene is decreased in HFE haemochromatosis, facilitating liver iron overload. Excess iron is then taken up by the liver and other tissues gradually over a long period. It seems likely that it is the iron itself that precipitates fibrosis.

Pathology

In symptomatic patients the total body iron content is 20–40 g, compared with 3–4 g in a normal person. The iron content is particularly increased in the liver and pancreas (50–100 times normal) but is also increased in other organs (e.g. the endocrine glands, heart and skin).

In established cases the liver shows extensive iron deposition and fibrosis. Early in the disease, iron is deposited in the periportal hepatocytes (in pericanalicular lysosomes). Later it is distributed widely throughout all acinar zones, biliary duct epithelium, Kupffer cells and connective tissue. Cirrhosis is a late feature.

Clinical features

The course of the disease depends on a number of factors, including sex, dietary iron intake, presence of associated hepatotoxins (especially alcohol) and genotypes. Overt clinical manifestations occur more frequently in men; the reduced incidence in women is probably explained by physiological blood loss and a smaller dietary intake of iron. Most affected individuals present in the fifth decade. The classic triad of bronze skin pigmentation (due to melanin deposition), hepatomegaly and diabetes mellitus is only present in cases of gross iron overload.

Hypogonadism secondary to pituitary dysfunction is the most common endocrine feature. Deficiency of other pituitary hormones is also found, but symptomatic endocrine deficiencies, such as loss of libido, are very rare. Cardiac manifestations, particularly heart failure and arrhythmias, are common, especially in younger patients. Calcium pyrophosphate is deposited asymmetrically in both large and small joints (chondrocalcinosis) leading to an arthropathy. The exact relationship of chondrocalcinosis to iron deposition is uncertain.

Complications

Thirty per cent of patients with cirrhosis will develop primary hepatocellular carcinoma (HCC). HCC has only very rarely been described in non-cirrhotic patients in whom the excess iron stores have been removed. Early diagnosis is vital.

Investigations

Homozygotes

- **Serum iron** is elevated (>30 μmol/L) in 90% with a reduction in the TIBC and a transferrin saturation of >45%.
- **Serum ferritin** is elevated (usually >500 μg/L or 240 nmol/L).
- **Liver biochemistry** is often normal, even with established cirrhosis.

Heterozygotes

Heterozygotes may have normal biochemical tests or modest increases in serum iron transferrin saturation (>45%) or serum ferritin (usually >400 μg/L).

Genetic testing

If iron studies are abnormal, genetic testing is performed.

Liver biopsy

This is not required for diagnosis, but is useful to assess the extent of tissue damage, assess tissue iron, and measure the hepatic iron concentration (>180 μmol/g dry weight of liver indicates haemochromatosis).

Mild degrees of parenchymal iron deposition in patients with other forms of cirrhosis, particularly due to alcohol, can often cause confusion with true homozygous HH.

Magnetic resonance imaging

MRI shows dramatic reduction in the signal intensity of the liver and pancreas owing to the paramagnetic affect of ferritin and haemosiderin. A highly T2 weighted gradient recalled echo (GRE) technique detects all clinically relevant liver iron overload (>60 μmol/g of liver). In secondary iron overload (haemosiderosis), which involves the reticuloendothelial cells, the pancreas is spared, enabling distinction between these two conditions.

Treatment and management

Venesection

This prolongs life and may reverse tissue damage; the risk of malignancy still remains if cirrhosis is present. All patients should have excess iron removed as rapidly as possible. This is achieved using venesection of 500 mL performed twice-weekly for up to 2 years, i.e. 160 units with 250 mg of iron per unit, equals 40 g removed. During venesection, serum iron and ferritin and the mean corpuscular volume (MCV) should be monitored. These fall only when available iron is depleted. Three or four venesections per year are required to prevent reaccumulation of iron. Serum ferritin should remain within the normal range.

Manifestations of the disease usually improve or disappear, except for diabetes, testicular atrophy and chondrocalcinosis. The requirements for insulin often diminish in diabetic patients. Testosterone replacement is often helpful.

In the rare patient who cannot tolerate venesection (because of severe cardiac disease or anaemia), chelation therapy with desferrioxamine, either intermittently or continuously by infusion, has been successful in removing iron.

Screening

In all cases of HH, all first-degree family members must be screened to detect early and asymptomatic disease. *HFE* mutation analysis is performed with measurement of transferrin saturation and serum ferritin.

In the general population, the serum iron and transferrin saturation are the best and cheapest tests available.

Wilson's disease (progressive hepatolenticular degeneration)

Dietary copper is normally absorbed from the stomach and upper small intestine. It is transported to the liver loosely bound to albumin. Here it is incorporated into apocaeruloplasmin forming caeruloplasmin, a glycoprotein synthesized in the liver, and secreted into the blood. The remaining copper is normally excreted in the bile and excreted in faeces.

Wilson's disease is a very rare inborn error of copper metabolism that results in copper deposition in various organs, including the liver, the basal ganglia of the brain and the cornea. It is potentially treatable and all young patients with liver disease must be screened for this condition.

Aetiology

It is an autosomal recessive disorder with a molecular defect within a copper-transporting ATPase encoded by a gene (designated *ATP7B*) located on chromosome 13, affecting between 1 in 30 000 and 1 in 100 000 individuals. Over 300 mutations have been identified, the most frequent being His1069 Gly (H1069Q) found in approximately 50% of Caucasian patients, with compound heterozygotes being frequent. This mutation is rare in India and Asia. Wilson's disease occurs world-wide, particularly in countries where consanguinity is common. There is a failure of both incorporation of copper into procaeruloplasmin, which leads to low serum caeruloplasmin, and biliary excretion of copper. There is a low serum caeruloplasmin in over 80% of patients but this is not the cause of the copper deposition. The precise mechanism for the failure of copper excretion is not known.

Pathology

The liver histology is not diagnostic and varies from that of chronic hepatitis to macronodular cirrhosis. Stains for copper show a periportal distribution but this can be unreliable (see below). The basal ganglia are damaged and show cavitation, the kidneys show tubular degeneration, and erosions are seen in bones.

Clinical features

Children usually present with hepatic problems, whereas young adults have more neurological problems, such as tremor, dysarthria, involuntary movements and eventually dementia. The liver disease varies from episodes of acute hepatitis, especially in children, which can go on to fulminant hepatic failure, to chronic hepatitis or cirrhosis.

Typical signs are of chronic liver disease with neurological signs of basal ganglia involvement (p. 1148). A specific sign is the Kayser–Fleischer ring, which is due to copper deposition in Descemet's membrane in the cornea. It appears as a greenish brown pigment at the corneoscleral junction and frequently requires slit-lamp examination for identification. It may be absent in young children.

Investigations

- **Serum copper and caeruloplasmin** are usually reduced but can be normal.
- **Urinary copper** is usually increased 100–1000 μg in 24 hours (1.6–16 μmol); normal levels <40 μg (0.6 μmol).
- **Liver biopsy.** The diagnosis depends on measurement of the amount of copper in the liver (>250 μg/g dry weight), although high levels of copper are also found in the liver in chronic cholestasis.
- **Haemolysis and anaemia** may be present.
- **Genetic analysis** is limited but selected exons are screened according to population group.

Treatment

Lifetime treatment with penicillamine, 1–1.5 g daily, is effective in chelating copper. If treatment is started early, clinical and biochemical improvement can occur. Urine copper levels should be monitored and the drug dose adjusted downwards after 2–3 years. Serious side-effects of the drug occur in 10% and include skin rashes, leucopenia, skin changes and renal damage. Trientine (1.2–1.8 g/day) and zinc acetate (150 mg/day) have been used as maintenance therapy and for asymptomatic cases. All siblings and children of patients should be screened (*ATP7B* mutation analysis is useful) and treatment given even in the asymptomatic if there is evidence of copper accumulation.

Prognosis

Early diagnosis and effective treatment have improved the outlook. Neurological damage is, however, permanent. Ful-

minant hepatic failure or decompensated cirrhosis should be treated by liver transplantation.

α_1-Antitrypsin deficiency (see also p. 837)

A deficiency of α_1-antitrypsin (α_1AT) is sometimes associated with liver disease and pulmonary emphysema (particularly in smokers). α_1AT is a glycoprotein, part of a family of serine protease inhibitors, or serpin, superfamily. α_1AT-deficiency is inherited as an autosomal dominant and 1 in 10 northern Europeans carries an abnormal gene.

The protein is a 394-amino acid 52 kDa acute-phase protein that is synthesized in the liver and constitutes 90% of the serum α_1-globulin seen on electrophoresis. Its main role is to inhibit the proteolytic enzyme, neutrophil elastase.

The gene is located on chromosome 14. The genetic variants of α_1AT are characterized by their electrophoretic mobilities as medium (M), slow (S) or very slow (Z). The normal genotype is protease inhibitor MM (PiMM), the homozygote for Z is PiZZ, and the heterozygotes are PiMZ and PiSZ. S and Z variants are due to a single amino acid replacement of glutamic acid at positions 264 and 342 of the polypeptide, respectively. This results in decreased synthesis and secretion of the protein by the liver as protein–protein interactions occur between the reactive centre loop of one molecule and the β-pleated sheet of a second (loop sheet polymerization).

How this causes liver disease is uncertain. It is postulated that the failure of secretion of the abnormal protein leads to an accumulation in the liver, causing liver damage.

Clinical features

The majority of patients with clinical disease are homozygotes with a PiZZ phenotype. Some may present in childhood and a few require transplantation. Approximately 10–15% of adult patients will develop cirrhosis, usually over the age of 50 years, and 75% will have respiratory problems. Approximately 5% of patients die of their liver disease. Heterozygotes (e.g. PiSZ or PiMZ) may develop liver disease, but the risk is small.

Investigations

- **Serum α_1-antritrypsin** is low, at 10% of the normal level in the PiZZ phenotypes, and 60% of normal in the S variant.

Histologically periodic acid–Schiff (PAS)-positive, diastase-resistant globules which contain α_1AT are seen in periportal hepatocytes. Fibrosis and cirrhosis can be present.

Treatment

There is no treatment apart from dealing with the complications of liver disease. Patients with hepatic decompensation should be considered for liver transplantation. Patients should stop smoking (see p. 830).

FURTHER READING

Adams PC, Barton JC. Haemochromatosis. *Lancet* 2007; **370**: 1855–1860.

Ala A, Walker AP, Ashkan K et al. Wilson's disease. *Lancet* 2007; **369**: 397–408.

ALCOHOLIC LIVER DISEASE

This section gives the pathology and clinical features of alcoholic liver disease. The amounts needed to produce liver damage, alcohol metabolism, and other clinical effects of alcohol are described on page 239.

Ethanol is metabolized in the liver by two pathways, resulting in an increase in the NADH/NAD ratio. The altered redox potential results in increased hepatic fatty acid synthesis with decreased fatty acid oxidation, both events leading to hepatic accumulation of fatty acid that is then esterified to glycerides.

The changes in oxidation–reduction also impair carbohydrate and protein metabolism and are the cause of the centrilobular necrosis of the hepatic acinus typical of alcohol damage.

Acetaldehyde is formed by the oxidation of ethanol, and its effect on hepatic proteins may well be a factor in producing liver cell damage. The exact mechanism of alcoholic hepatitis and cirrhosis is unknown, but since only 10–20% of people who drink heavily will develop cirrhosis, a genetic predisposition is recognized. Immunological mechanisms have also been proposed.

Alcohol can enhance the effects of toxic metabolites of drugs (e.g. paracetamol) on the liver, as it induces microsomal metabolism via the microsomal ethanol oxidizing system (MEOS) (p. 238).

Pathology

Alcohol can produce a wide spectrum of liver disease from fatty change to hepatitis and cirrhosis.

Fatty change

The metabolism of alcohol invariably produces fat in the liver, mainly in zone 3. This is minimal with small amounts of alcohol, but with larger amounts the cells become swollen with fat (steatosis). There is no liver cell damage. The fat disappears on stopping alcohol. Steatosis is also seen in non-alcoholic fatty liver disease.

In some cases collagen is laid down around the central hepatic veins (perivenular fibrosis) and this can sometimes progress to cirrhosis without a preceding hepatitis. Alcohol directly affects stellate cells, transforming them into collagen-producing myofibroblast cells (p. 320).

Alcoholic hepatitis

In addition to fatty change there is infiltration by polymorphonuclear leucocytes and hepatocyte necrosis mainly in zone 3. Dense cytoplasmic inclusions called Mallory bodies are sometimes seen in hepatocytes and giant mitochondria are also a feature. Mallory bodies are suggestive of, but not specific for, alcoholic damage as they can be found in other liver disease, such as Wilson's disease and PBC. If alcohol consumption continues, alcoholic hepatitis may progress to cirrhosis.

Alcoholic cirrhosis

This is classically of the micronodular type, but a mixed pattern is also seen accompanying fatty change, and evidence of pre-existing alcoholic hepatitis may be present.

Clinical features

Fatty liver

There are often no symptoms or signs. Vague abdominal symptoms of nausea, vomiting and diarrhoea are due to the more general effects of alcohol on the gastrointestinal tract. Hepatomegaly, sometimes huge, can occur together with other features of chronic liver disease.

Alcoholic hepatitis

The clinical features vary in degree:

- The patient may be well, with few symptoms, the hepatitis only being apparent on the liver biopsy in addition to fatty change.
- Mild to moderate symptoms of ill-health, occasionally with mild jaundice, may occur. Signs include all the features of chronic liver disease. Liver biochemistry is deranged and the diagnosis is made on liver histology.
- In the severe case, usually superimposed on alcoholic cirrhosis, the patient is ill, with jaundice and ascites. Abdominal pain is frequently present, with a high fever associated with the liver necrosis. On examination there is deep jaundice, hepatomegaly, sometimes splenomegaly, and ascites with ankle oedema. The signs of chronic liver disease are also present.

Alcoholic cirrhosis

This represents the final stage of liver disease from alcohol abuse. Nevertheless, patients can be very well with few symptoms. On examination, there are usually signs of chronic liver disease. The diagnosis is confirmed by liver biopsy.

Usually the patient presents with one of the complications of cirrhosis. In many cases there are features of alcohol dependency (see p. 1212) as well as evidence of involvement of other systems, such as polyneuropathy.

Investigations

Fatty liver

An elevated MCV often indicates heavy drinking. Liver biochemistry shows mild abnormalities with elevation of both serum aminotransferase enzymes. The γ-GT level is a sensitive test for determining whether the patient is taking alcohol. With severe fatty infiltration, marked changes in all liver biochemical parameters can occur. Ultrasound or CT will demonstrate fatty infiltration, as will liver histology. Elastography can be used to estimate the degree of fibrosis.

Alcoholic hepatitis

Investigations show a leucocytosis with markedly deranged liver biochemistry with elevated:

- serum bilirubin
- serum AST and ALT
- serum alkaline phosphatase
- prothrombin time (PT).

A low serum albumin may also be found. Rarely, hyperlipidaemia with haemolysis (Zieve's syndrome) may occur.

Liver biopsy, if required, is performed by the transjugular route because of prolonged PT.

Alcoholic cirrhosis

Investigations are as for cirrhosis in general.

Management and prognosis

General management

Patients should be advised to stop drinking. Delirium tremens (a withdrawal symptom) is treated with diazepam. Intravenous thiamin should be given empirically to prevent Wernicke–Korsakoff encephalopathy. Bed rest with a diet high in protein and vitamin supplements is given. Dietary protein rarely needs to be limited because of encephalopathy. Patients need to be advised to participate in alcohol cessation programmes. The likelihood of abstention is dependent on many factors, particularly social and family ones.

Fatty liver

The patient is advised to stop drinking alcohol; the fat will disappear and the liver biochemistry usually returns to normal. Small amounts of alcohol can be drunk subsequently as long as patients are aware of the problems and can control their consumption.

Alcoholic hepatitis

In severe cases the patient requires admission to hospital. Nutrition must be maintained with enteral feeding if necessary and vitamin supplementation given. Steroid therapy does improve outcome in more severe cases as judged by a discriminant function (DF).

Discriminant function (DF)

DF = [4.6 × prothrombin time above control in seconds] + bilirubin (mg/dL)

Bilirubin μmol/L ÷ 17 to convert to mg/dL

Severe = >32

Infection must be excluded or concomitantly treated. Treatment for encephalopathy and ascites is commenced. Antifungal prophylaxis should also be used.

Patients are advised to stop drinking for life, as this is undoubtedly a precirrhotic condition. The prognosis is variable and, despite abstinence, the liver disease is progressive in many patients.

In severe cases the mortality is at least 50%, and with a PT twice the normal, progressive encephalopathy and renal failure, the mortality approaches 90%.

Alcoholic cirrhosis

The management of cirrhosis is described on page 347. Again, all patients are advised to stop drinking for life. Abstinence from alcohol results in an improvement in prognosis, with a 5-year survival of 90%, but with continued drinking this falls to 60%. With advanced disease (i.e. jaundice, ascites and haematemesis) the 5-year survival rate falls to 35%, with most of the deaths occurring in the first year. Liver transplantation results in good survival – recurrence of cirrhosis due to recidivism is rare. Patients often sign a contract with their clinicians regarding their abstinence, both before and after transplantation.

A trial of abstention to establish if liver disease can improve is mandatory, but transplantation should not be denied if the patient continues to deteriorate. Specific follow-up regarding alcohol use is recommended.

Hepatocellular carcinoma is a complication, particularly in men.

BUDD–CHIARI SYNDROME

In this condition there is obstruction to the venous outflow of the liver owing to occlusion of the hepatic vein. In one-third of patients the cause is unknown, but specific causes include hypercoagulability states (e.g. paroxysmal nocturnal haemoglobinuria, polycythaemia vera) or thrombophilia (p. 442), taking the contraceptive pill, or leukaemia. Other causes include occlusion of the hepatic vein owing to posterior abdominal wall sarcomas, renal or adrenal tumours, hepatocellular carcinoma, hepatic infections (e.g. hydatid cyst), congenital venous webs, radiotherapy, or trauma to the liver.

The acute form presents with abdominal pain, nausea, vomiting, tender hepatomegaly and ascites (a fulminant form occurs particularly in pregnant women). In the chronic form there is enlargement of the liver (particularly the caudate lobe), mild jaundice, ascites, a negative hepatojugular reflux, and splenomegaly with portal hypertension.

Investigations

Investigations show a high protein content in the ascitic fluid and characteristic liver histology with centrizonal congestion, haemorrhage, fibrosis and cirrhosis. Ultrasound, CT or MRI will demonstrate hepatic vein occlusion with diffuse abnormal parenchyma on contrast enhancement. The caudate lobe is spared because of its independent blood supply and venous drainage. There may be compression of the inferior vena cava. Pulsed Doppler sonography or colour Doppler is useful as they show abnormalities of flow in the hepatic vein. Thrombophilia screening is mandatory. Multiple defects of coagulation are found. Thrombosis of the portal vein is present in 2% of patients.

Differential diagnosis

A similar clinical picture can be produced by inferior vena caval obstruction, right-sided cardiac failure or constrictive pericarditis, and appropriate investigations should be performed.

Treatment

In the acute situation, thrombolytic therapy can be given. Ascites should be treated, as should any underlying cause (e.g. polycythaemia). Congenital webs should be treated radiologically or resected surgically. A transjugular intrahepatic portosystemic shunt (TIPS) is the first treatment of choice as caval compression does not prejudice the efficacy of TIPS. Surgical porto-caval shunts are reserved for those who fail this treatment providing there is no caval obstruction or severe caval compression when a caval stent can be inserted. Liver transplantation is the treatment of choice for chronic Budd–Chiari syndrome and for the fulminant form. Lifelong anticoagulation is mandatory following TIPS and transplantation.

Prognosis

The prognosis depends on the aetiology, but some patients can survive for several years.

FURTHER READING

Senzolo M, Cholongitas E, Patch D et al. Update on the classification, assessment, prognosis and therapy of Budd–Chiari syndrome. *Nature Clinical Practice Gastroenterology and Hepatology* 2005; **2**: 182–190.

VENO-OCCLUSIVE DISEASE

This is due to injury of the hepatic veins and presents clinically like the Budd–Chiari syndrome. It was originally described in Jamaica, where the ingestion of toxic pyrrolizidine alkaloids in bush tea (made from plants of the genera *Senecio, Heliotropium* and *Crotalaria*) caused damage to the hepatic veins. It can be seen in other parts of the world. It is also seen as a complication of chemotherapy and total body irradiation before allogeneic bone marrow transplantation. The development of veno-occlusive disease after transplantation carries a high mortality. Treatment is supportive with control of ascites and hepatocellular failure. Defibrotide has been tried and TIPS has been used in a few cases.

FIBROPOLYCYSTIC DISEASES

These diseases are usually inherited and lead to the presence of cysts or fibrosis in the liver, kidney and occasionally the pancreas, and other organs.

Polycystic disease of the liver

Multiple cysts can occur in the liver as part of autosomal dominant polycystic disease of the kidney (p. 642). These cysts are usually asymptomatic but occasionally cause abdominal pain and distension. Liver function is normal and complications such as oesophageal varices are very rare. The prognosis is excellent and depends on the kidney disease.

Solitary cysts

These are usually found by chance during imaging and are mainly asymptomatic.

Congenital hepatic fibrosis

In this rare condition the liver architecture is normal but there are broad collagenous fibrous bands extending from the portal tracts. It is often inherited as an autosomal recessive condition but can also occur sporadically. It usually presents in childhood with hepatosplenomegaly, and portal hypertension is common. It may present later in life and can be misdiagnosed as cirrhosis.

A wedge biopsy of the liver may be required to confirm the diagnosis. The outlook is good and the condition should be distinguished from cirrhosis. Patients who bleed do well after endoscopic therapy of varices or a portocaval anastomosis because of their good liver function.

Congenital intrahepatic biliary dilatation (Caroli's disease)

In this rare, non-familial disease there are saccular dilatations of the intrahepatic or extrahepatic ducts. It can present at any age (although usually in childhood) with fever, abdominal pain and recurrent attacks of cholangitis with Gram-negative septicaemia. Jaundice and portal hypertension are absent. Diagnosis is by ultrasound, PTC, ERCP or MRCP. There is an increased risk of biliary malignancy.

LIVER ABSCESS

Pyogenic abscess

These abscesses are uncommon, but may be single or multiple. The most common was a portal pyaemia from intra-abdominal sepsis (e.g. appendicitis or perforations), but now in many cases the aetiology is not known. In the elderly, biliary sepsis is a common cause. Other causes include trauma, bacteraemia and direct extension from, for example, a perinephric abscess.

The organism found most commonly is *E. coli. Streptococcus milleri* and anaerobic organisms such as *Bacteroides* are often seen. Other organisms include *Enterococcus fae-*

calis, *Proteus vulgaris* and *Staphylococcus aureus*. Often the infection is mixed.

Clinical features

Some patients are not acutely ill and present with malaise lasting several days or even months. Others can present with fever, rigors, anorexia, vomiting, weight loss and abdominal pain. In these patients a Gram-negative septicaemia with shock can occur. On examination there may be little to find. Alternatively, the patient may be toxic, febrile and jaundiced. In such patients, the liver is tender and enlarged and there may be signs of a pleural effusion or a pleural rub in the right lower chest.

Investigations

Patients who are not acutely ill are often investigated as a 'pyrexia of unknown origin' (PUO) and most investigations will be normal. Often the only clue to the diagnosis is a raised serum alkaline phosphatase.

- **Serum bilirubin** is raised in 25% of cases.
- **Normochromic normocytic anaemia** may occur, usually accompanied by a polymorphonuclear leucocytosis.
- **Serum alkaline phosphatase and ESR** are often raised.
- **Serum B_{12}** is very high, as vitamin B_{12} is stored in and subsequently released from the liver.
- **Blood cultures** are positive in only 30% of cases.

Imaging

Ultrasound is useful for detecting abscesses. A CT scan may be of value in complex and multiple lesions (Fig. 7.26). A chest X-ray will show elevation of the right hemidiaphragm with a pleural effusion in the severe case. Depending on age, imaging of the colon may be necessary to find the source of the infection.

Management

Aspiration of the abscess should be attempted under ultrasound control. Antibiotics should initially cover Gram-positive, Gram-negative and anaerobic organisms until the causative organism is identified.

Further drainage via a large-bore needle under ultrasound control or surgically may be necessary if resolution is difficult or slow. Any underlying cause must also be treated.

Prognosis

The overall mortality depends on the nature of the underlying pathology and has been reduced to approximately 16% with needle aspiration and antibiotics. A unilocular abscess in the right lobe has the better prognosis. Scattered multiple abscesses have a very high mortality, with only one in five patients surviving.

Amoebic abscess (see also p. 161)

This condition occurs world-wide and must be considered in patients travelling from endemic areas. *Entamoeba histolytica* (p. 161) can be carried from the bowel to the liver in the portal venous system. Portal inflammation results, with the development of multiple microabscesses and eventually single or multiple large abscesses.

Clinically the onset is usually gradual but may be sudden. There is fever, anorexia, weight loss and malaise. There is often no history of dysentery. On examination the patient looks ill and has tender hepatomegaly and signs of an effusion or consolidation in the base of the right side of the chest. Jaundice is unusual.

Investigations

These are as for pyogenic abscess, plus:

- Serological tests for amoeba (e.g. haemagglutination, amoebic complement fixation test, ELISA). These are always positive, particularly if there are bowel symptoms, and remain positive after a clinical cure and therefore do not indicate current disease. A repeat negative test, however, is good evidence against an amoebic abscess.
- Diagnostic aspiration of fluid looking like anchovy sauce.

Treatment

Metronidazole 800 mg three times daily is given for 10 days. Aspiration is used in patients failing to respond, in multiple and sometimes large abscesses, and in those with abscesses in the left lobe of the liver or impending rupture.

Complications

Complications include rupture, secondary infection and septicaemia.

OTHER INFECTIONS OF THE LIVER

Schistosomiasis (see also p. 168)

Schistosoma mansoni and *S. japonicum* affect the liver, but *S. haematobium* rarely does so. During their life cycle the ova reach the liver via the venous system and obstruct the portal branches, producing granulomas, fibrosis and inflammation but not cirrhosis.

Clinically there is hepatosplenomegaly and portal hypertension, which is particularly severe with *S. mansoni*. In Egypt there is frequently concomitant chronic hepatitis C infection.

Investigations show a raised serum alkaline phosphatase and ova can be found in the stools (centrifuged deposits) and in rectal and liver biopsies. Skin tests and other immunological tests often have false results and may also be positive because of past infection.

Treatment is with praziquantel, but fibrosis still remains with a potential risk of portal hypertension, characteristically pre-sinusoidal due to intense portal fibrosis.

Hydatid disease (see also p. 171)

Cysts caused by *Echinococcus granulosus* are single or multiple. They usually occur in the lower part of the right lobe. The cyst has three layers: an outside layer derived from the host, an intermediate laminated layer, and an inner germinal layer that buds off brood capsules to form daughter cysts.

Clinically there may be no symptoms or a dull ache and swelling in the right hypochondrium. Investigations show a

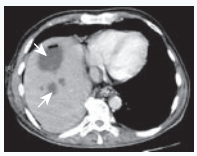


Fig. 7.26 **CT of liver abscesses** in the right lobe of the liver (arrowed), secondary to partial biliary obstruction.

Ask the authors
www.studentconsult.com

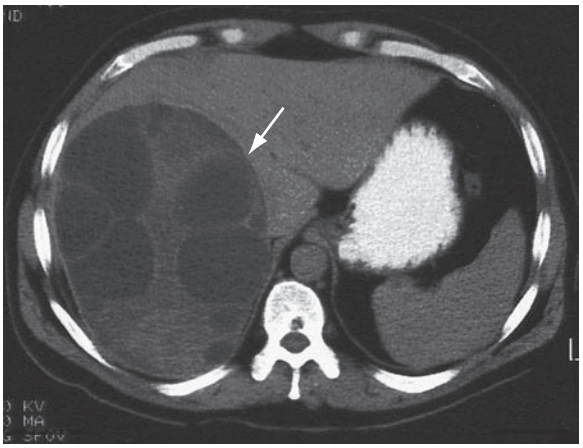


Fig. 7.27 **CT scan of liver showing a large hydatid cyst (arrow) with 'daughter' cysts lying within it.**

peripheral eosinophilia in 30% of cases and usually a positive hydatid complement fixation test or haemagglutination (85%). Plain abdominal X-ray may show calcification of the outer coat of the cyst. Ultrasound and CT scan demonstrate cysts and may show diagnostic daughter cysts within the parent cyst (Fig. 7.27).

Medical treatment (e.g. with albendazole 10 mg/kg, which penetrates into large cysts) results in cysts becoming smaller. Puncture, aspiration, injection, reaspiration (PAIR) has been used since the 1980s. Fine-needle aspiration is undertaken under ultrasound control with chemotherapeutic cover. Surgery can be performed with removal of the cyst intact if possible after first sterilizing the cyst with alcohol, saline or cetrimide. Chronic calcified cysts can be left. There have been no well-designed clinical trials for any modality.

Complications include rupture into the biliary tree, other organs or intraperitoneally, with spread of infection. The prognosis without any complications is good, although there is always the risk of rupture. Preventative measures include deworming of pet dogs and prevention of pets from eating infected carcasses, as well as veterinary control programmes.

Echinococcus multilocularis causes alveolar echinococcosis and is almost exclusively a hepatic disease, with a high mortality if not treated. Early diagnosis would enable radical surgery and then continued chemosuppression.

Acquired immunodeficiency syndrome (see also p. 187)

The liver is often involved and is a significant cause of morbidity or mortality. HIV itself is probably not the cause of the liver abnormalities. The following are seen:

- pre-existing/coincidental viral hepatitis – the hepatitis progresses more rapidly (HBV, HCV, HDV) and is a leading cause of death
- neoplasia: Kaposi's sarcoma and non-Hodgkin's lymphoma – increased risk of HCC
- opportunistic infection (e.g. *Mycobacterium tuberculosis, M. avium-intracellulare, Cryptococcus, Candida albicans*, toxoplasmosis)
- drug hepatotoxicity
- secondary sclerosing cholangitis (see p. 162)
- non-cirrhotic portal hypertension associated with antiretroviral therapy.

Clinical hepatomegaly is common (60% of patients).

LIVER DISEASE IN PREGNANCY

Liver function is not impaired in pregnancy. Any liver disease from whatever cause can occur incidentally and coincide with pregnancy. For example, viral hepatitis accounts for 40% of all cases of jaundice during pregnancy. Pregnancy does not necessarily exacerbate established liver disease, but it is uncommon for women with advanced liver disease to conceive.

The following changes take place:

- Plasma and blood volumes increase during pregnancy but the hepatic blood flow remains constant.
- The proportion of cardiac output delivered to the liver therefore falls from 35% to 29% in late pregnancy; drug metabolism can thus be affected.
- The size of the liver remains constant.
- Liver biochemistry remains unchanged apart from a rise in serum alkaline phosphatase from the placenta (up to three to four times) and a decrease in total protein owing to increased plasma volume.
- Triglycerides and cholesterol levels rise, and caeruloplasmin, transferrin, α_1-antitrypsin and fibrinogen levels are elevated owing to increased hepatic synthesis.
- Postpartum there is a tendency to hypercoagulability, and acute Budd–Chiari syndrome can occur.

There are a number of liver diseases that complicate pregnancy.

Hyperemesis gravidarum

Pathological vomiting during pregnancy can be associated with liver dysfunction and jaundice. Liver dysfunction resolves when vomiting subsides.

Intrahepatic cholestasis of pregnancy

This condition of unknown aetiology presents usually with pruritus alone in the third trimester. It has a familial tendency and there is a higher prevalence in Scandinavia, Chile and Bolivia.

Liver biochemistry shows a cholestatic picture with high serum ALP (up to four times normal) and raised aminotransferases which occasionally can be very high. The serum bilirubin is slightly raised with jaundice in 60% of cases. Liver biopsy is not indicated but would show centrilobular cholestasis.

Treatment is symptomatic with ursodeoxycholic acid 15 mg/kg daily. Prognosis is usually excellent for the mother but there is increased fetal loss and the condition resolves after delivery. Recurrent cholestasis may occur during subsequent pregnancies or with the ingestion of oestrogen-containing oral contraceptive pills.

Pre-eclampsia and eclampsia

Pre-eclampsia is characterized by hypertension, proteinuria and oedema occurring in the second or third trimester. Eclampsia is marked by seizures or coma in addition. Hepatic complications include subcapsular haematoma and infarction, and occasionally fulminant hepatic failure. The HELLP

syndrome – a combination of haemolysis, elevated liver enzymes and a low platelet count – can occur in association with severe pre-eclampsia. In the HELLP syndrome, there is epigastric pain, nausea and vomiting, with jaundice in 5% of patients. Delivery is the best treatment for eclampsia.

Acute fatty liver of pregnancy (AFLP)

This is a rare, serious condition of unknown aetiology. There is an association between acute fatty liver and long-chain 3-hydroxylacyl-CoA-dihydroxyl (LCHAD) deficiency. The mechanism is unclear, but abnormal fatty acid metabolites produced by the homozygous or heterozygous fetus enter the circulation and overcome maternal hepatic mitochondrial oxidation systems in a heterozygote mother. It presents in the last trimester with symptoms of fulminant hepatitis – jaundice, vomiting, abdominal pain, and occasionally haematemesis and coma.

Investigations show hepatocellular damage, hyperuricaemia, thrombocytopenia, and rarely DIC. CT scanning shows a low density of the liver owing to the high fat content. It can sometimes be difficult to differentiate from the HELLP syndrome and as LCHAD deficiency has also been shown in HELLP there is a view that there is a spectrum of HELLP to AFLP. Liver biopsy shows fine droplets of fat (microvesicular) in hepatocytes with little necrosis, but is not necessary for diagnosis.

Immediate delivery of the child may save both baby and mother. Early diagnosis and treatment has reduced the mortality to less than 20%. Treatment is as for acute liver failure.

LIVER TUMOURS

Secondary liver tumours

The most common liver tumour is a secondary (metastatic) tumour, particularly from the gastrointestinal tract (from the distribution of the portal blood supply), breast or bronchus. They are usually multiple.

Clinical features are variable but usually include weight loss, malaise, upper abdominal pain and hepatomegaly, with or without jaundice.

Diagnosis. Ultrasound is the primary investigation, with CT or MRI to define metastases and look for a primary. The serum alkaline phosphatase is almost invariably raised.

Treatment. This will depend on the site of the primary and the burden of liver metastases. The best results are obtained in colorectal cancer in patients with few hepatic metastases. If the primary tumour is removed and hepatic resection is performed, reasonable survival rates are possible. Chemotherapy is used, particularly with breast cancer (p. 488). Radiofrequency ablation of the metastases is an alternative to surgery.

PRIMARY MALIGNANT TUMOURS

Primary liver tumours may be benign or malignant, but the most common are malignant.

Hepatocellular carcinoma

Hepatocellular carcinoma (HCC) is the fifth most common cancer world-wide.

Aetiology

Carriers of HBV and HCV have an extremely high risk of developing HCC. In areas where HBV is prevalent, 90% of patients with this cancer are positive for the hepatitis B virus. Cirrhosis is present in approximately 80% of these patients. The development of HCC is related to the integration of viral HBV DNA into the genome of the host hepatocyte (see p. 336) and the degree of viral replication (>10 000 copies/mL). The risk of HCC in HCV is higher than in HBV (even higher with both HBV and HCV) despite no viral integration. Unlike HBV infection, cirrhosis is always present. Primary liver cancer is also associated with other forms of cirrhosis, such as alcoholic cirrhosis and haemochromatosis. Non-alcoholic fatty liver disease is associated with HCC, probably secondary to the presence of cirrhosis. Males are affected more than females. Other aetiological factors are aflatoxin (a metabolite of a fungus found in groundnuts) and androgenic steroids, and there is a weak association with the contraceptive pill.

Pathology

The tumour is either single or occurs as multiple nodules throughout the liver. Histologically it consists of cells resembling hepatocytes. It can metastasize via the hepatic or portal veins to the lymph nodes, bones and lungs.

Clinical features

The clinical features include weight loss, anorexia, fever, an ache in the right hypochondrium and ascites. The rapid development of these features in a cirrhotic patient is suggestive of HCC. On examination, an enlarged, irregular, tender liver may be felt. Increasingly, due to surveillance, HCC is found without symptoms in cirrhotics.

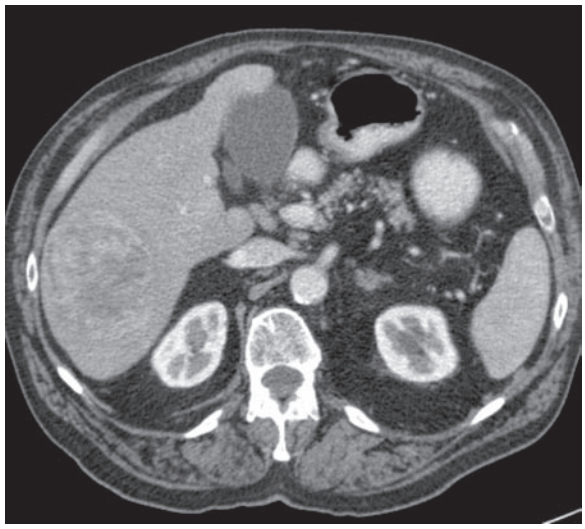
Investigations

Serum α-fetoprotein may be raised, but is normal in at least a third of patients. *Ultrasound* scans show filling defects in 90% of cases. Enhanced *CT* scans (Fig. 7.28) identify HCC but it is difficult to confirm the diagnosis in lesions <2 cm. An *MRI* can further help to delineate lesions. Tumour *biopsy*, particularly under ultrasonic guidance, may be performed for diagnosis, but is increasingly less used as imaging techniques show characteristic appearances (hypervascularity of the nodule and lack of portal vein wash out) and because seeding along the biopsy tract can occur.

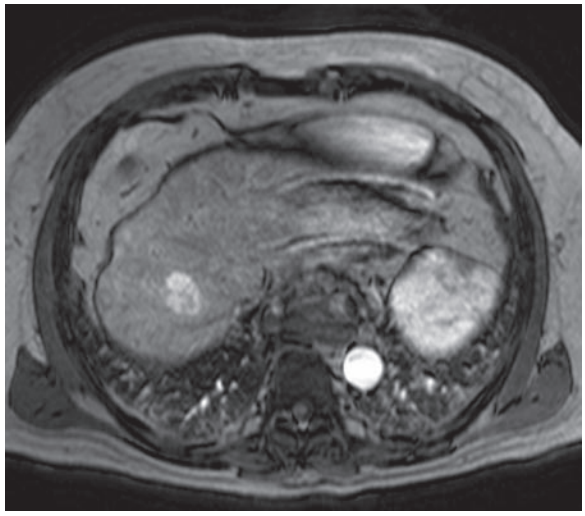
Treatment and prognosis

Surgical resection is occasionally possible. Conventional chemotherapy and radiotherapy are unhelpful, but transarterial embolization in selected patients prolongs survival. Radiofrequency ablation is also effective. Antiangiogenic compounds are being evaluated: sorafenib prolongs survival in patients with non-resectable tumours. Liver transplantation is curative in patients with smaller tumours (<3 nodules <3 cm diameter or a single tumour <5 cm).

In advanced cases survival is seldom more than 6 months.



(a)



(b)

Fig. 7.28 **Hepatocellular carcinoma. (a)** CT showing cirrhosis with a hepatocellular carcinoma. **(b)** T2-weighted MRI of the liver following gadolinium contrast showing a hepatocellular carcinoma.

Prevention

Persistent HBV infection, usually acquired after perinatal infection, is a high risk factor for HCC in many parts of the world, such as South East Asia. Widespread vaccination against HBV is being used and this has reduced the annual incidence of HCC in Taiwan.

Cholangiocarcinoma

Cholangiocarcinomas are increasing in incidence and can be extrahepatic (see p. 374) or intrahepatic. Intrahepatic adenocarcinomas arising from the bile ducts account for approximately 10% of primary tumours. They are not associated with cirrhosis or hepatitis B. In the Far East they may be associated with infestation with *Clonorchis sinensis* or *Opisthorchis viverrini*. The clinical features are similar to primary HCC except that jaundice is frequent with hilar tumours, and cholangitis is more frequent.

Surgical resection is rarely possible and patients usually die within 6 months. Transplantation is contraindicated, outside of specialized protocols.

FURTHER READING

Aljabiri MR, Lodato F, Burroughs AK. Surveillance and diagnosis for hepatocellular carcinoma. *Liver Transplantation* 2007; **13**(11 Suppl 2): S2–12.

Khan SA, Miras A, Pelling M et al. Cholangiocarcinoma and its management. *Gut* 2007; **56**(12): 1755–1756.

BENIGN TUMOURS

The most common benign tumour is a haemangioma. It is usually small and single but can be multiple and large. Haemangiomas are usually found incidentally on ultrasound, CT or MRI and have characteristic appearances. They require no treatment.

Hepatic adenomas are associated with oral contraceptives. They can present with abdominal pain or intraperitoneal bleeding. Resection is only required for symptomatic patients, those with tumours >5 cm diameter or in those in whom discontinuation of oral contraception does not result in shrinkage of the tumour.

MISCELLANEOUS CONDITIONS OF THE LIVER

Hepatic mitochondrial injury syndromes

These syndromes – in which there is mitochondrial damage with inhibition of β-oxidation of fatty acids – can be categorized as follows:

- *Genetic*, with abnormalities which include medium-chain acyl-coenzyme A dehydrogenase deficiency leading to microsteatosis.
- *Toxins* leading to liver failure include aflatoxin and cerulide (produced by *Bacillus cereus*) which causes food poisoning (see p. 133).
- *Drugs* (e.g. i.v. tetracycline, valproic acid and nucleoside reverse-transcriptase inhibitors) can produce a fatal microsteatosis.
- *Idiopathic*, the best known being fatty liver of pregnancy (p. 363) and Reye's syndrome. The latter, due to inhibition of β-oxidation and uncoupling of oxidative phosphorylation in mitochondria, leads in children to an acute encephalopathy and diffuse microvesicular fatty infiltration of the liver. Aspirin ingestion and viral infections have been implicated as precipitating agents. Mortality is about 50%, usually due to cerebral oedema.

Idiopathic adult ductopenia

This unexplained condition is characterized by pruritus and cholestatic jaundice. Histology of the liver shows a decrease in intrahepatic bile ducts in at least 50% of the portal tracts, together with the features of cholestasis and marked fibrosis or cirrhosis. In most, the disease is progressive and the only treatment is liver transplantation.

Indian childhood cirrhosis

This condition of children is seen in the Indian subcontinent. The cause is unknown. Eventually there is development of a micronodular cirrhosis with excess copper in the liver.

Hepatic porphyrias

These are dealt with on page 1074.

Cystic fibrosis (see also p. 844)

This disease affects mainly the lung and pancreas, but patients can develop fatty liver, cholestasis and cirrhosis. The aetiology of the liver involvement is unclear.

Coeliac disease (see also p. 277)

Abnormal liver biochemical tests are common. A tissue transglutaminase should be performed when hepatic causes are not found when investigating abnormal liver biochemistry. These return to normal with a gluten-free diet.

DRUGS AND THE LIVER

Drug metabolism

The liver is the major site of drug metabolism. Drugs are converted from fat-soluble to water-soluble substances that can be excreted in the urine or bile. This metabolism of drugs is mediated by a group of mixed-function enzymes (p. 925).

Drug hepatotoxicity

Many drugs impair liver function. When abnormal liver biochemical tests are found, drugs should always be considered as a cause, particularly when other causes have been excluded. Damage to the liver by drugs is usually classified as being either predictable (or dose-related) or non-predictable (not dose-related) (see p. 928). However, there is considerable overlap and at least six mechanisms may be involved in the production of damage:

1. disruption of intracellular calcium homeostasis
2. disruption of bile canalicular transport mechanisms
3. formation of non-functioning adducts (enzyme–drug) which may then
4. present on the surface of the hepatocyte as new immunogens (attacked by T cells)
5. induction of apoptosis
6. inhibition of mitochondrial function, which prevents fatty acid metabolism, and accumulation of both lactate and reactive oxygen species.

The predominant mechanism or combination of mechanisms determines the type of liver injury, i.e. hepatitic, cholestatic or immunological (skin rashes, fever and arthralgia, i.e. serum-sickness syndrome). Eosinophilia and circulating immune complexes and antibodies may occasionally be detected.

When a small amount of hepatotoxic drug whose effect is dose-dependent (e.g. paracetamol) is ingested, a large proportion of it undergoes conjugation with glucuronide and sulphate, whilst the remainder is metabolized by microsomal enzymes to produce toxic derivatives that are immediately detoxified by conjugation with glutathione. If larger doses are ingested, the former pathway becomes saturated and the toxic derivative is produced at a faster rate. Once the hepatic glutathione is depleted, large amounts of the toxic metabolite accumulate and produce damage (p. 947).

The 'predictability' of drugs to produce damage can, however, be affected by metabolic events preceding their ingestion. For example, chronic alcohol abusers may become more susceptible to liver damage because of the enzyme-inducing effects of alcohol, or ill or starving patients may become susceptible because of the depletion of hepatic glutathione produced by starvation. Many other factors such as environmental or genetic effects may be involved in determining the 'susceptibility' of certain patients to certain drugs.

The incidence of drug hepatotoxicity is 14 per 100 000 population with a 6% mortality. It is the most common cause of acute liver failure in the USA. Liver transplantation is used.

Hepatitic damage

The type of damage produced by various drugs is shown in Table 7.16. The diagnosis of these conditions is usually by exclusion of other causes. Most reactions occur within 3 months of starting the drug. Monitoring liver biochemistry in patients on long-term treatment, such as antituberculosis therapy, is **mandatory**. If a drug is suspected of causing hepatic damage it should be stopped immediately. Liver biopsy is of limited help in confirming the diagnosis, but occasionally hepatic eosinophilia or granulomas may be seen. Diagnostic challenge with subtherapeutic doses of the drug is sometimes required after the liver biochemistry has returned to normal, to confirm the diagnosis.

Individual drugs

Paracetamol. In high doses paracetamol produces liver cell necrosis (see above). The toxic metabolite binds irreversibly to liver cell membranes. Overdosage is discussed on page 947.

Halothane and other volatile liquid anaesthetics. Halothane, which is available in the UK on a named patient basis, produces a hepatitis in patients having repeated exposures. The mechanism is thought to be a hypersensitivity reaction. An unexplained fever occurs approximately 10 days after the second or subsequent halothane anaesthetic and is followed by jaundice, typically with a hepatitic picture. Most patients recover spontaneously but there is a high mortality in severe cases. There are no chronic sequelae. Both sevoflurane and isoflurane also cause hepatotoxicity in those sensitized to halogenated anaesthetics but the risk is smaller than with halothane and remote with desflurane.

Steroid compounds. Cholestasis is caused by natural and synthetic oestrogens as well as methyltestosterone. These agents interfere with canalicular biliary flow by blocking MRP2 and MDR3 (see Fig. 7.4) and cause a pure cholestasis. Cholestasis is rare with the contraceptive pill because of the low dosage used. However, the contraceptive pill is associated with an increased incidence of gallstones, hepatic adenomas (rarely HCCs), the Budd–Chiari syndrome and peliosis hepatis. The latter condition, which also occurs with anabolic steroids, consists of dilatation of the hepatic sinusoids to form blood-filled lakes.

Phenothiazines. Phenothiazines (e.g. chlorpromazine) can produce a cholestatic picture owing to a hypersensitivity reaction. It occurs in 1% of patients, usually within 4 weeks of starting the drug. Typically it is associated with a fever and eosinophilia. Recovery occurs on stopping the drug.

Table 7.16 Liver damage produced by some drugs

Types of liver damage	Drugs
Zone 3 necrosis	Carbon tetrachloride *Amanita* mushrooms Paracetamol Salicylates Piroxicam Cocaine
Zone 1 necrosis	Ferrous sulphate
Microvesicular fat	Sodium valproate Tetracyclines
Steatohepatitis	Amiodarone Synthetic oestrogens Nifedipine
Fibrosis	Methotrexate Other cytotoxic agents Arsenic Vitamin A Retinoids
Vascular	
Sinusoidal dilatation	Contraceptive drugs Anabolic steroids
Peliosis hepatis	Azathioprine Oral contraceptives Anabolic steroids, e.g. danazol Azathioprine
Veno-occlusive	Pyrrolizidine alkaloids (*Senecio* in bush tea) Cytotoxics – cyclophosphamide
Acute hepatitis	Isoniazid Rifampicin Methyldopa Atenolol Enalapril Verapamil Ketoconazole Cytotoxic drugs Clonazepam Disulfiram Niacin Volatile liquid anaesthetics, e.g. halothane
Chronic hepatitis	Methyldopa Nitrofurantoin Fenofibrate Isoniazid
General hypersensitivity	Sulphonamides, e.g. Sulfasalazine Co-trimoxazole Fansidar Penicillins, e.g. Flucloxacillin Ampicillin Amoxicillin Co-amoxiclav NSAIDs, e.g. Salicylates Diclofenac Allopurinol Antithyroid. e.g. Propylthiouracil Carbimazole Quinine, e.g. Quinidine Diltiazem Anticonvulsants, e.g. Phenytoin
Canalicular cholestasis	Sex hormones Ciclosporin Chlorpromazine Haloperidol Erythromycin Cimetidine/ranitidine Nitrofurantoin Imipramine Azathioprine Oral hypoglycaemics
Biliary sludge	Ceftriaxone
Sclerosing cholangitis	Hepatic arterial infusion of 5-fluorouracil
Hepatic tumours	Pills with high hormone content (adenomas)
Hepatocellular carcinoma	Contraceptive pill Danazol

NSAID, non-steroidal anti-inflammatory drug.
NB. Anti-HIV drugs, e.g. maraviroc, cause hepatic dysfunction.

Antituberculous chemotherapy. *Isoniazid* produces elevated aminotransferases in 10–20% of patients. Hepatic necrosis with jaundice occurs in a smaller percentage. The hepatotoxicity of isoniazid is related to its metabolites and is dependent on acetylator status. *Rifampicin* produces a hepatitis, usually within 3 weeks of starting the drug, particularly in patients on high doses. *Pyrazinamide* produces abnormal liver biochemical tests and, rarely, liver cell necrosis.

Amiodarone. This leads to a steatohepatitis histologically and liver failure if the drug is not stopped in time.

Sodium valproate. This causes mitochondrial injury with microvesicular steatosis. Intravenous carnitine should be used as an antidote.

FURTHER READING

Lee WM. Drug-induced hepatotoxicity. *New England Journal of Medicine* 2003; **349**: 474–485.

Drug prescribing for patients with liver disease

The metabolism of drugs is impaired in severe liver disease (with jaundice and ascites) as the removal of many drugs depends on liver blood flow and the integrity of the hepatocyte. In general, therefore, the effect of drugs is prolonged by liver disease and also by cholestasis. This is further accentuated by portosystemic shunting, which diminishes the first-pass extraction of drugs. With hypoproteinaemia there is decreased protein binding of some drugs, and bilirubin competes with many drugs for the binding sites on serum albumin. In patients with portosystemic encephalopathy, care must be taken in prescribing drugs with a central depressant action, e.g. narcotics, including codeine and anxiolytics. Other common drugs to be avoided in cirrhosis include ACE inhibitors (cause hepatorenal failure) and NSAIDs (bleeding).

GALL BLADDER AND BILIARY SYSTEM

The structure, formation and function of bile is discussed on page 321.

GALLSTONES

Prevalence of gallstones

Gallstones may be present at any age but are unusual before the third decade. The prevalence of gallstones is strongly influenced both by age and sex. There is a progressive increase in the presence of gallstones with age but the prevalence is two to three times higher in women than in men, although this difference is less marked in the sixth and seventh decade. At this age the prevalence ranges between 25% and 30%. The increase in life expectancy is reflected in an increased burden of symptomatic gallstone disease. There are considerable racial differences, gallstones being more common in Scandinavia, South America and Native North Americans but less common in Asian and African groups.

Types of gallstones

Two principal types of gallstone disease occur. In the western world 80% of gallstones contain cholesterol. The second, less frequent, type of gallstone is 'pigment stones', being predominantly composed of calcium bilirubinate or polymer-like complexes with calcium, copper and some cholesterol.

Cholesterol gallstones

The formation of cholesterol stones is the consequence of cholesterol crystallization from gall bladder bile. This is dependent upon three factors:

- cholesterol supersaturation of bile
- crystallization-promoting factors within bile
- motility of the gall bladder.

The majority of cholesterol is derived from hepatic uptake from dietary sources. However, hepatic biosynthesis may account for up to 20%. The rate-limiting step in cholesterol synthesis is β-hydroxy-β methyl glutaryl Co-A (HMG-CoA) reductase, which catalyses the first step, i.e. the conversion of acetate to mevalonate (see Fig. 7.4). The cholesterol formed is co-secreted with phospholipids into the biliary canaliculus as unilamellar vesicles. Cholesterol will only crystallize into stones when the bile is supersaturated with cholesterol relative to the bile salt and phospholipid content. This can occur as a consequence of excess cholesterol secretion into bile which, in some instances, has been shown to be due to an increase in HMG-CoA reductase activity. Recently, leptin (p. 229) has been shown to increase cholesterol secretion into bile. Elevated levels of leptin during rapid weight loss may account for the increased incidence of cholesterol gallstones. An alternative mechanism of supersaturation is a decreased *bile salt content* which may occur as a consequence of bile salt loss (e.g. terminal ileal resection or involvement with Crohn's disease).

The *composition* of the bile salt pool may also influence the ability to maintain cholesterol in solution. There is evidence that an increased proportion of the secondary hydrophobic bile acid (deoxycholic acid) in the bile acid pool may predispose to cholesterol stone formation. This has been linked with slow colonic transit during which the primary bile acid cholic acid may undergo microbial enzyme metabolism yielding deoxycholic acid which is then absorbed back into the bile salt pool (see Fig. 7.4).

Whilst cholesterol supersaturation of bile is essential for cholesterol stone formation, many individuals in whom such supersaturation occurs will never develop stones. It is the balance between cholesterol crystallizing and solubilizing factors that determines whether cholesterol will crystallize out of solution. A number of lipoproteins have been reported as putative crystallizing factors.

There is increasing evidence from epidemiological, family and twin studies of the importance of genetic factors in gallstone formation. A number of lithogenic genes have been identified which may interact with environmental factors. However, monogenic susceptibility appears uncommon. There are rare cases in which a single mutation leading to ABC (ATP-binding cassette) transporter deficiencies is associated with recurrent intra- and extrahepatic cholelithiasis in young patients.

Gall bladder motility represents a further factor that may influence the cholesterol crystallization from supersaturated bile. There is evidence from animal models that gall bladder stasis leads to cholesterol crystallization mediated by hypersecretion of mucin.

Abnormalities of gall bladder motility have been suggested as factors in such circumstances as pregnancy, multiparity and diabetes as well as octreotide-related gall bladder stones (p. 980). Recognized risk factors for gallstones are shown in Table 7.17.

Bile pigment stones

The pathogenesis of pigment stones is entirely independent of cholesterol gallstones. There are two main types of pigment gallstones, black and brown.

Table 7.17 Risk factors for cholesterol gallstones

Increasing age
Sex (F > M)
Family history
Multiparity
Obesity ± metabolic syndrome
Rapid weight loss
Diet (e.g. high in animal fat/low in fibre)
Drugs (e.g. contraceptive pill)
Ileal disease or resection
Diabetes mellitus
Acromegaly treated with octreotide
Liver cirrhosis

FURTHER READING

Portincasa P, Moschetta A, Palasciano G. Cholesterol gallstone disease. *Lancet* 2006; **368**: 230–239.

Black pigment gallstones are composed of calcium bilirubinate and a network of mucin glycoproteins that interlace with salts such as calcium carbonate and/or calcium phosphate. These stones range in colour from deep black to very dark brown and have a glass-like cross-sectional surface on fracturing. Black pigment stones are seen in 40–60% of patients with haemolytic conditions such as sickle cell disease and hereditary spherocytosis in which there is chronic excess bilirubin production. However, the majority of pigment stones occur without haemolysis. There is recent evidence that bile salt loss into the colon (consequent upon ileal resection or ileal disease) promotes solubilization and colonic reabsorption of bilirubin. This enhances the enterohepatic circulation and biliary secretion of bilirubin with the formation of gallstones. Pigment stones have also been linked with bacterial colonization of the biliary tree. Some pigment stones have been shown to contain bacteria, many of which produce glucuronidase and phospholipase, factors that are known to facilitate stone formation. It is speculated that this subclinical bacterial colonization of the bile duct is responsible for the pigment stone formation.

Brown stones are usually of a muddy hue and on cross-section seem to have alternating brown and tan layers. These stones are composed of calcium salts of fatty acids as well as calcium bilirubinate. They are almost always found in the presence of bile stasis and/or biliary infection. Brown stones are a common cause of recurrent bile duct stones following cholecystectomy and may also be found in the intrahepatic ducts in circumstances of duct disease such as Caroli's syndrome and primary sclerosing cholangitis. In the Far East such brown stones are identified within both the intra- and extrahepatic biliary tree and have been linked with chronic parasitic infection.

Clinical presentation of gallstones

(Fig. 7.29)

The majority of gallstones are asymptomatic and remain so during a person's lifetime. Gallstones are increasingly detected as an incidental finding at the time of either abdominal radiography or ultrasound scanning. Over a 10- to 15-year period approximately 20% of these stones will be the cause of symptoms with 10% having severe complications. Once gallstones have become symptomatic there is a strong trend towards recurrent complications, often of increasing severity. Gallstones do not cause dyspepsia, fat intolerance, flatulence or other vague upper abdominal symptoms. The chances of a fair, fat, female of 40 having gallstones are the same as in the general population.

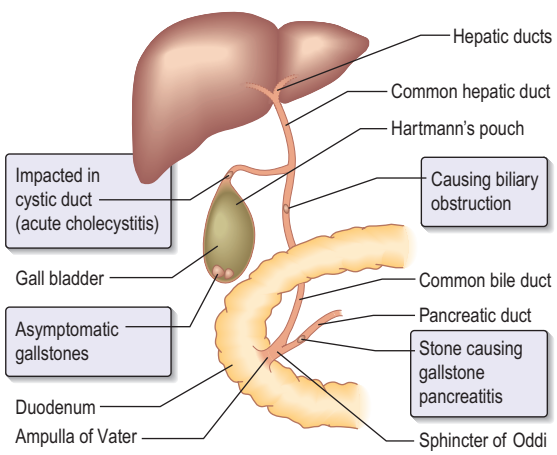


Fig. 7.29 **Clinical presentation of gallstones.**

Biliary or gallstone colic

Biliary colic is the term used for the pain associated with the temporary obstruction of the cystic or common bile duct by a stone usually migrating from the gall bladder. Despite the term 'colic', the pain of stone-induced ductular obstruction is severe but constant and has a crescendo characteristic. Many sufferers can relate the symptoms to over-indulgence with food, particularly when this has a high fat content. The most common time of day for such an episode is in the mid-evening and lasting until the early hours of the morning.

The initial site of pain is usually in the epigastrium but there may be a right upper quadrant component. Radiation may occur over the right shoulder and right subscapular region.

Nausea and vomiting frequently accompany the more severe attacks. The cessation of the attack may be spontaneous after a number of hours or terminated by the administration of opiate analgesia. More protracted pain, particularly when associated with fevers and rigors, suggests secondary complications such as cholecystitis, cholangitis or gallstone-related pancreatitis (see below).

Acute cholecystitis

The initial event in acute cholecystitis is obstruction to gall bladder emptying. In 95% of cases a gall bladder stone can be identified as the cause. Such obstruction results in an increase of gall bladder glandular secretion leading to progressive distension which in turn may compromise the vascular supply to the gall bladder.

There is also an inflammatory response secondary to retained bile within the gall bladder. Infection is a secondary phenomenon following the vascular and inflammatory events described above.

The initial clinical features of an episode of cholecystitis are similar to those of biliary colic described above. However, over a number of hours there is progression with severe localized right upper quadrant abdominal pain corresponding to parietal peritoneal involvement in the inflammatory process. The pain is associated with tenderness and muscle guarding or rigidity. Occasionally the gall bladder can become distended by pus (an empyema) and rarely an acute gangrenous cholecystitis develops which can perforate, with generalized peritonitis.

Investigations

Biliary colic as a consequence of a stone in the neck of the gall bladder or cystic duct is unlikely to be associated with significant abnormality of laboratory tests. Acute cholecystitis is usually associated with a moderate leucocytosis and raised inflammatory markers (e.g. C reactive protein).

- **The serum bilirubin, alkaline phosphatase and aminotransferase levels** may be marginally elevated in the presence of cholecystitis alone, even in the absence of bile duct obstruction. More significant elevation of the

bilirubin and alkaline phosphatase is in keeping with bile duct obstruction.

- **An abdominal ultrasound scan** is the single most useful investigation for the diagnosis of gallstone-related disease (Fig. 7.30). Look for:
  (a) *gallstones* within the gall bladder, particularly when these are obstructing the gall bladder neck or cystic duct
  (b) *focal tenderness* over the underlying gall bladder
  (c) *thickening of the gall bladder wall*. This may also be seen with hypoalbuminaemia, portal hypertension and acute viral hepatitis.

  Gallstones are a common finding in an ageing population, and in the absence of specific symptoms great care should be taken when determining whether the gallstones are responsible for the symptoms.
- **Biliary scintigraphy using technetium derivatives of iminodiacetate.** These isotopes are taken up by hepatocytes and excreted into bile. They delineate the extrahepatic biliary tree. The absence of cystic duct and gall bladder filling provides evidence towards acute cholecystitis although the findings must be closely correlated with the presenting symptoms.

Differential diagnosis

Typical cases of biliary colic are usually suspected by a careful clinical history. The differential diagnosis includes the irritable bowel syndrome (spasm of the hepatic flexure), carcinoma of the right side of the colon, atypical peptic ulcer disease, renal colic and pancreatitis.

The differential diagnosis of acute cholecystitis includes a number of other conditions marked by severe right upper quadrant pain and fever, e.g. acute episodes of pancreatitis, perforated peptic ulceration or an intrahepatic abscess. Conditions above the right diaphragm such as basal pneumonia as well as myocardial infarction may on occasions mimic the clinical picture.

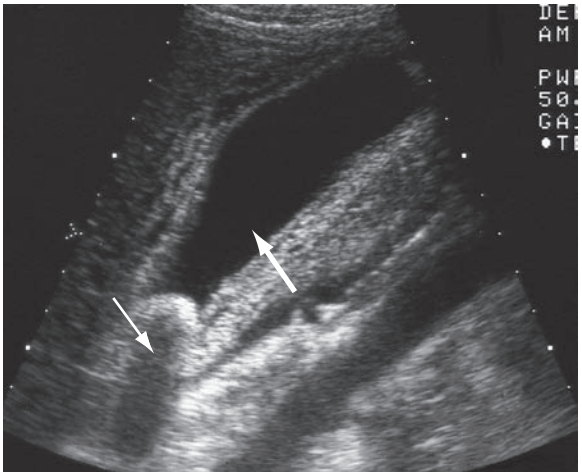


Fig. 7.30 **Ultrasound scan in a patient with acute cholecystitis.** There is a stone (casting an acoustic shadow – thin arrow) impacted in the gall bladder neck, with a distended gall bladder (thick arrow) and thickening and oedema of the gall bladder wall.

MANAGEMENT OF GALL BLADDER STONES

Cholecystectomy

Cholecystectomy is the treatment of choice for virtually all patients with *symptomatic* gall bladder stones. In patients admitted with specific gallstone-related complications (see below) cholecystectomy should be carried out during the period of that admission to prevent the risk of recurrence. For those presenting with pain alone an elective procedure can be planned but the waiting time should be minimized to avoid the high risk of recurrent symptoms (approximately 30% over 4 months) and the need for hospital admission.

Cholecystectomy should not be performed in the absence of typical symptoms just because stones are found on investigation. There is an ongoing debate as to whether prophylactic cholecystectomy is justified in young patients found to have small stones. Such patients have a long period over which they may develop symptomatic disease and small stones are an independent risk factor for the potentially serious complication of gallstone pancreatitis. Each case should be discussed on an individual risk benefit basis.

The laparotomy approach to cholecystectomy has now been largely replaced by the laparoscopic technique. Postoperative pain is minimized with only a short period of ileus and the early ability to mobilize the patient. Laparoscopic cholecystectomy can be safely carried out on a day-care basis in otherwise fit patients (although in most cases an overnight stay remains the norm). This has considerable cost benefits over open cholecystectomy, which is now reserved for a small proportion of patients with contraindications such as extensive previous upper abdominal surgery, ongoing bile duct obstruction or portal hypertension.

In approximately 5% of cases a laparoscopic cholecystectomy is converted to an open operation because of technical difficulties, in particular adhesions in the right upper quadrant or difficulty in identifying the biliary anatomy.

Acute cholecystitis. The initial management is conservative, consisting of nil by mouth, intravenous fluids, opiate analgesia and intravenous antibiotics (dictated by local policy with options including extended spectrum cephalosporins, fluoroquinolones or piperacillin/tazobactam).

Cholecystectomy is usually delayed for a few days to allow the symptoms to settle but can then be carried out quite safely in the majority of cases.

When the clinical situation fails to respond to this conservative management, particularly if there is increasing pain and fever, an empyema or gangrene of the gall bladder may have occurred. Urgent ultrasound is performed and urgent surgery will be required if these complications have developed.

Specific complications of cholecystectomy include a biliary leak either from the cystic duct or from the gall bladder bed. Injury to the bile duct itself occurs in up to 0.5% of laparoscopic operations and may have serious long-term sequelae in the form of biliary sepsis and secondary biliary liver injury. There is an overall mortality of 0.2%.

Stone dissolution and shock wave lithotripsy

These non-surgical techniques for the management of gall bladder stones are infrequently used but may still have a role in a few highly selected patients who may not be fit for laparoscopic cholecystectomy.

Dissolution. Pure or near-pure cholesterol stones can be solubilized by increasing the bile salt content of bile. Chenodeoxycholic acid and ursodeoxycholic acid were used but now laparoscopic surgery has made this treatment redundant. In any event, its efficacy was limited and the recurrence rate of gallstones was high. These limitations have restricted the application of this approach. Further benefit may be attained by the addition of an HMG-CoA reductase inhibitor (e.g. simvastatin) to a regimen of ursodeoxycholic acid. This will then combine reduced cholesterol content of bile as well an enhanced bile salt concentration but uses are still limited.

Extracorporeal shock wave lithotripsy. A shock wave can be directed either radiologically or by ultrasound on to gall bladder stones. This technique was highly successful but in only a restricted patient population.

The post-cholecystectomy syndrome

This refers to right upper quadrant pain, often biliary in type, which occurs a few months after the cholecystectomy but may be delayed for a number of years. The patients often comment that the pain is identical to that for which the original operation was carried out. In many cases this syndrome is related to functional large bowel disease with colonic spasm at the hepatic flexure (hepatic flexure syndrome). In a small proportion of patients the pain is a result of a retained common bile duct stone. In a further minority of patients, hypertension of the sphincter of Oddi is a potential cause. This is most likely in patients with abnormal liver biochemistry and dilatation of the common bile duct (on ultrasound) during episodes of pain (and in the absence of a documented retained stone). The diagnosis is confirmed by pressure measurements of the sphincter of Oddi, and the condition can be successfully managed by endoscopic sphincterotomy (see below).

COMMON BILE DUCT STONES

The classical features of common bile duct (CBD) stones are biliary colic, fever and jaundice (acute cholangitis). This triad is only present in the minority of patients. Abdominal pain is the most common symptom and has the typical features of biliary colic (see above). Jaundice is a variable accompaniment and is almost always preceded by abdominal pain. A patient with bile duct stones may experience sequential episodes of pain, only some of which are accompanied by jaundice. In contrast to malignant bile duct obstruction, the level of jaundice associated with CBD stones characteristically tends to fluctuate.

Fever is only present in a minority of cases but indicates biliary sepsis and sometimes an associated septicaemia. The presence of such biliary sepsis is a significant adverse prognostic factor.

A minority of patients with bile duct stones are discovered incidentally during imaging for gall bladder disease. Fifteen per cent of patients undergoing cholecystectomy will have stones within the bile duct only detected at the time of operative cholangiography. The frequency of asymptomatic bile duct stones resulting in complications is not well documented. It is likely that many such stones will pass into the duodenum without causing symptoms. However, the potential for serious complication is well recognized and in most circumstances incidentally identified bile duct stones are removed (see below).

Physical examination

If the patient is examined between episodes there may be no abnormal physical finding. During a symptomatic episode the patient may be jaundiced with a fever and associated tachycardia. There is tenderness in the right upper quadrant varying from mild to extremely severe.

More widespread abdominal tenderness extending from the epigastrium to the left upper quadrant, associated with distension, may indicate associated stone-related pancreatitis (see below).

Investigations

Laboratory tests

- *Full blood count* is usually normal in the presence of uncomplicated bile duct stones.
- *An elevated neutrophil count* and raised inflammatory markers (ESR and CRP) are frequent accompaniments of cholangitis.
- *The raised serum bilirubin* tends to be mild and often transient. Very high concentrations of bilirubin (≥ 200 μmol/L) almost always reflect complete bile duct obstruction.
- *Serum alkaline phosphatase and γ-glutamyl transpeptidase* are similarly elevated in proportion to the degree of hyperbilirubinaemia.
- *Aminotransferase levels* are usually mildly elevated but with complete bile duct obstruction there may be very marked rises to 10–15 times the normal value. These high levels may lead to an initial misdiagnosis of an hepatitic process.
- Serum amylase levels are often mildly elevated in the presence of bile duct obstruction but are markedly so if stone-related pancreatitis has occurred.
- Prothrombin time may be prolonged if bile duct obstruction has occurred; this reflects decreased absorption of vitamin K.

Transabdominal ultrasound

This is the initial imaging technique of choice and in many cases the only imaging technique required.

Bile duct obstruction is characterized by dilatation of intrahepatic biliary radicles, which are usually easily detected by the ultrasound scan. It may, however, not be possible to identify the cause of obstruction. Stones situated in the distal common bile duct are poorly visualized by transabdominal ultrasound and up to 50% are missed. The detection of stones within the gall bladder is poorly predictive as to the cause of bile duct obstruction. Asymptomatic gallstones are common (up to 15%) in patients who are 65 years and older. Conversely, in 5–10% of patients with bile duct stones, no calculi can be seen within the gall bladder.

Other imaging techniques

Magnetic resonance cholangiography (MRC) delineates the fluid column within the biliary tree and is a sensitive technique for the detection of common bile duct stones in the presence of a dilated duct. The technique may be less accurate in the absence of duct dilatation (see the role of endoscopic ultrasound, below) (Fig. 7.31).

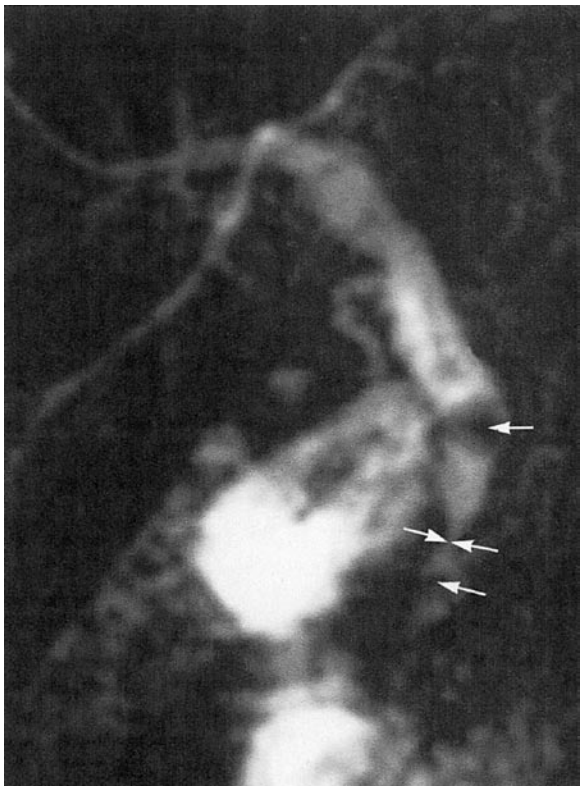


Fig. 7.31 A magnetic resonance cholangiogram in a patient presenting with abdominal pain and jaundice. This shows evidence of a distal common bile duct stricture (arrows) with a large gallstone proximal to this in the mid common bile duct (top arrow).

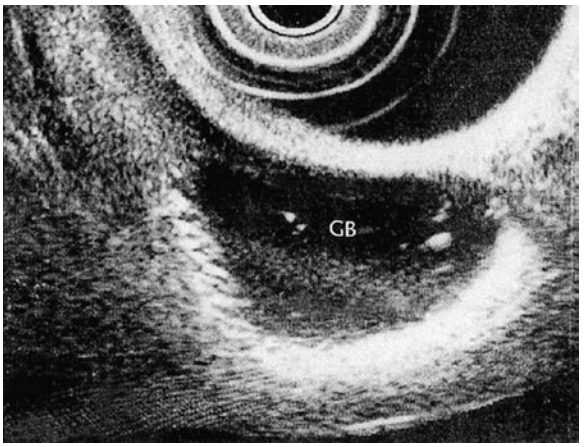


Fig. 7.32 An endoscopic ultrasound scan with the probe within the duodenum (upper margin of the figure) clearly demonstrating the gall bladder (GB) with multiple small stones within.

Spiral CT scanning is an alternative way to detect bile duct dilatation. Opaque stones are more readily identifiable within the bile duct than radiolucent cholesterol stones. CT scanning also provides a means of excluding other causes of bile duct obstruction such as carcinoma of the head of the pancreas.

Endoscopic ultrasound scanning (Fig. 7.32) has enabled high-resolution imaging of the common bile duct, gall bladder and pancreas, although unlike the preceding imaging techniques it is an invasive procedure. The endoscopic ultrasound probe in the duodenum is in close proximity with the distal common bile duct and hence can identify the majority of stones at this level.

This technique may be particularly useful for identifying small calculi (microcalculi), particularly in a non-dilated common bile duct.

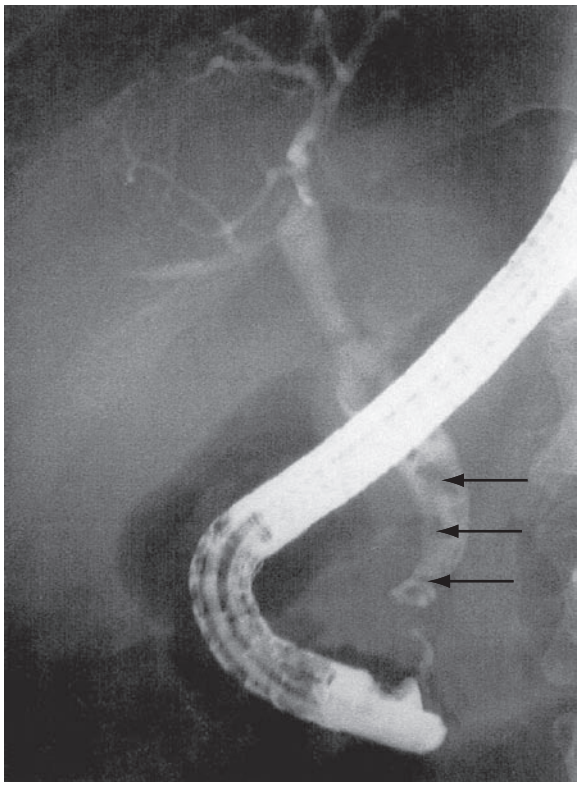


Fig. 7.33 An ERCP in a patient presenting with abdominal pain, jaundice and fever. Several stones can be seen within the dilated common bile duct (arrows).

Endoscopic retrograde cholangiography (ERC)

The endoscopic technique of ERC (p. 327) enables good visualization of the common bile duct. In experienced hands this will be successful in 98% of cases, providing good documentation of bile duct stones (Fig. 7.33).

However, microcalculi can still be missed. ERC gives the therapeutic opportunity for sphincterotomy and stone extraction (see below).

Differential diagnosis

Cholangitis may occur independently of gallstones in conditions such as primary sclerosing cholangitis and Caroli's syndrome. On occasions cholangitis may accompany malignant bile duct obstruction. Jaundice may also be a feature of acute cholecystitis in the absence of bile duct stones. A stone impacted in the cystic duct may compress and obstruct the bile duct (Mirizzi's syndrome). Common bile duct stones may produce pain, but in the absence of jaundice, the differential diagnosis is that of biliary colic (see above).

Management

Acute cholangitis has a high morbidity and mortality, particularly in the elderly. Successful management depends on

intravenous antibiotics (see as for acute cholecystitis), and urgent bile duct drainage by an endoscopic retrograde approach (Fig. 7.33). Access to the bile duct is achieved by sphincterotomy, and thereafter the stones can be removed either by balloon or basket catheters. In the severely ill patient a piece of plastic tubing (termed a stent) can be inserted into the bile duct to maintain bile drainage without the need to remove the stones, hence minimizing the time period to complete the procedure. The residual stones can then be removed endoscopically when the patient has recovered from the episode. In the presence of acute cholangitis surgical drainage has been associated with a high mortality and is now limited to those very few cases which cannot be managed by the endoscopic approach.

Urgent endoscopic bile duct clearance is also indicated in some patients with acute gallstone pancreatitis (see below). Patients who have retained common bile duct stones after a previous cholecystectomy are also optimally managed by endoscopic clearance. Patients shown to have common bile duct stones as well as gall bladder stones may be treated by two different approaches:

- *Laparoscopic cholecystectomy* which can also include exploration of the CBD via the cystic duct or by direct choledochotomy. By using these techniques the laparoscopist can extract stones from the common bile duct. This, however, prolongs the procedure, particularly in the presence of large stones or biliary sepsis.
- *Endoscopic approach* either immediately before or after the cholecystectomy. Removal of CBD stones by this method is the preferred way in the UK.

COMPLICATIONS OF GALLSTONES

- Acute cholecystitis and acute cholangitis have been discussed (p. 368).
- Gallstone-related pancreatitis is discussed on page 379.
- Gallstones can occasionally erode through the wall of the gall bladder into the intestine giving rise to a biliary enteric fistula. Passage of a gallstone through into the small bowel can give rise to an ileus or true obstruction.
- There is little evidence that gallstones are associated with an increased risk of adenocarcinoma of the gall bladder (p. 374).

MISCELLANEOUS CONDITIONS OF THE BILIARY TRACT

GALL BLADDER

There are a number of non-calculous conditions of the gall bladder, some of which have been associated with symptoms.

Non-calculous cholecystitis

Almost 10% of gall bladders removed for biliary symptoms are shown to have chronic inflammation within the wall but an absence of gallstones. Such cases are described as non-calculous cholecystitis. In many instances the gall bladder inflammation is minor and of doubtful significance. In a minority of cases non-calculous cholecystitis is characterized by severe inflammation frequently associated with gall bladder perforation. This condition is characteristically found in an elderly and critically ill group of patients. Chemical inflammation of the gall bladder may also occur from reflux of pancreatic enzymes back into the biliary tree, usually through the common channel at the ampulla of Vater. Bacterial infection of the gall bladder has occasionally been recognized as a cause of chronic inflammation.

The decision to carry out cholecystectomy in the absence of defined gall bladder stones should be guided by the specific features of the history and whether there is evidence of a diseased gall bladder wall on ultrasound scanning.

Cholesterolosis of the gall bladder

In cholesterolosis, cholesterol and other lipids are deposited in macrophages within the lamina propria of the gall bladder. These may be diffusely situated, giving a granular appearance to the gall bladder wall, or on occasions more discrete, giving a polypoid appearance (see below). Cholesterolosis of the gall bladder may coexist with gallstones but occurs independently. Some degree of cholesterolosis may be found in up to 25% of autopsies in an elderly population. It is doubtful whether this is a cause of symptoms.

Adenomyomatosis of the gall bladder

Adenomyomatosis is a gall bladder abnormality characterized by hyperplasia of the mucosa, thickening of the muscle wall and multiple intramural diverticula (the so-called 'Rokitansky–Aschoff sinuses').

The condition is usually detected as an incidental finding during investigation for possible gall bladder disease. It has been suggested that this condition is secondary to increased intraluminal gall bladder pressure but this is not proven. Gallstones may frequently coexist but there is no evidence to support a direct relationship.

It is unlikely that adenomyomatosis alone is a cause of biliary symptoms.

Chronic cholecystitis

There are no symptoms or signs that can conclusively be shown to be due to chronic cholecystitis. Symptoms attributed to this condition are vague, such as indigestion, upper abdominal discomfort or distension. There is no doubt that gall bladders studied histologically can show signs of chronic inflammation, and occasionally a small, shrunken gall bladder is found either radiologically or on ultrasound examination. However, these findings can be seen in asymptomatic people and therefore this clinical diagnosis should not be made. Most patients with chronic right hypochondrial pain suffer from functional bowel disease (p. 309).

EXTRAHEPATIC BILIARY TRACT

Primary sclerosing cholangitis

Primary sclerosing cholangitis (PSC) is a chronic cholestatic liver disease characterized by fibrosing inflammatory destruction of both the intra- and extrahepatic bile ducts. In 75% of

patients PSC is associated with inflammatory bowel disease (usually ulcerative colitis) and it is not unusual for the PSC to predate the onset of the inflammatory bowel disease. The causes are unknown but genetic susceptibility to PSC is associated with the HLA A1-B8-DR3 haplotype. The autoantibody pANCA (anti-neutrophil cytoplasmic antibody) is found in the serum of 60% of cases. Seventy per cent of patients are men and the average age of onset is approximately 40 years. Secondary PSC is seen in patients with HIV and cryptosporidium (p. 162).

Clinical features

With increasing screening of patients with inflammatory bowel disease PSC is detected at an asymptomatic phase with abnormal liver biochemistry, usually a raised serum alkaline phosphatase. Symptomatic presentation is usually with fluctuating pruritus, jaundice and cholangitis.

Diagnosis

The typical biliary changes associated with PSC may be identified by MRC scanning. This technique may fail to identify minor, but still clinically significant, intrahepatic duct abnormalities and this may require endoscopic retrograde cholangiography (ERC). The cholangiogram characteristically shows irregularity of calibre of both intra- and extrahepatic ducts, although either may be involved alone (see Fig. 7.34).

Management

Confirmation of the diagnosis comes from liver histology, which shows inflammation of the intrahepatic biliary radicles with associated scar tissue classically described as being onion skin in appearance. The histological changes may range from minor inflammatory infiltrates to the level of established cirrhosis. The presence of cirrhosis has prognostic implications.

PSC is a slowly progressive lesion (symptoms and biochemical tests may fluctuate), ultimately leading to liver cirrhosis and associated decompensation. Cholangiocarcinoma occurs in up to 15% of patients.

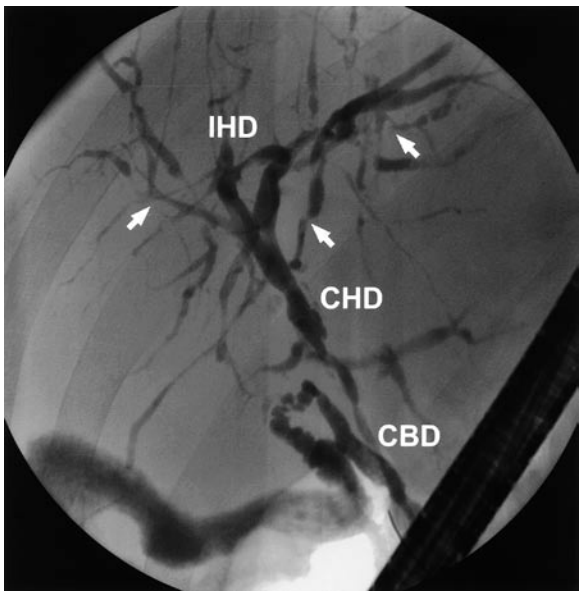


Fig. 7.34 **Primary sclerosing cholangitis.** An endoscopic cholangiogram showing the typical features of primary sclerosing cholangitis. There are calibre irregularities of the intrahepatic ducts (IHD). There is also minor stricturing of the extrahepatic ducts at the confluence between the common bile duct (CBD) and the common hepatic duct (CHD).

The only proven treatment is liver transplantation. The bile acid ursodeoxycholic acid has been evaluated extensively in the treatment of PSC, with well-documented improvement in liver function tests but no conclusive evidence of symptomatic, histological or survival benefit. There is some evidence that high dose therapy (30 mg/kg) may offer some benefit. In a small minority of patients with PSC the dominant lesion is of the extrahepatic ducts. Such lesions may be amenable to endoscopic biliary intervention with balloon dilatation and temporary stent placement.

Choledochal cyst

Congenital cystic disease of the bile ducts may occur at all levels of the biliary tree, although most cysts are extrahepatic. The dilatation may be saccular, diverticular or of fusiform configuration. In many cases there is associated pancreatobiliary malunion with the pancreatic duct draining directly into the common bile duct. The majority of symptomatic cases present in childhood with features of cholangitis. The formation of stones and sludge within the cystic segment may predispose to acute relapsing pancreatitis. In adult life choledochal cysts may be a differential diagnosis in patients presenting with symptoms suggestive of bile duct stones. The cyst must be fully resected to avoid the recurrent biliary complications as well as averting the risk (approximately 15%) of subsequent cholangiocarcinoma.

Haemobilia

Haemobilia is the term used to describe bleeding into the biliary tree. This may be as a consequence of liver trauma or as a complication of liver surgery. Biopsy of the liver is also a well-recognized cause. The end result is a fistula between a branch of the hepatic artery and an intrahepatic bile duct.

Haemobilia may be a cause of significant gastrointestinal blood loss and should be suspected when melaena is accompanied by right-sided upper abdominal pain and jaundice. However, the bleeding may occur without any overt biliary symptoms. If the diagnosis is suspected, bleeding may be managed by occlusion of the feeding artery by thrombosis performed radiologically.

Some patients will require surgery to control the bleeding point.

TUMOURS OF THE BILIARY TRACT

Gall bladder polyps

Polyps of the gall bladder are a common finding, being seen in approximately 4% of all patients referred for hepatobiliary ultrasonography. The vast majority of these are small (less than 5 mm), are non-neoplastic and are inflammatory in origin or composed of cholesterol deposits (see above). Adenomas are the most common benign neoplasm of the gall bladder. Only a proportion of these have a cancerous potential. The only reliable means of defining those at risk is polyp size. Cholecystectomy is recommended for any polyp approximating to 1 cm in diameter or larger.

FURTHER READING

Broome V, Bergquist A. Primary sclerosing cholangitis, inflammatory bowel disease and colon cancers. *Seminars in Liver Disease* 2006; **26**: 31–41.

Primary cancer of the gall bladder

Adenocarcinoma of the gall bladder represents 1% of all cancers. The mean age of occurrence is in the early sixties with a ratio of 3 women to 1 man. Gallstones have been suggested as an aetiological factor but this relationship remains unproven. Diffuse calcification of the gall bladder (porcelain gall bladder), considered to be the end stage of chronic cholecystitis, has also been associated with cancer of the gall bladder and is an indication for early cholecystectomy. Adenomatous polyps of the gall bladder in excess of 1 cm in diameter are also recognized as premalignant lesions (see above).

Carcinoma of the gall bladder is often detected at the time of planned cholecystectomy for gallstones and in such circumstances resection of an early lesion may be curative. Early lymphatic spread to the liver and adjacent biliary tract precludes curative resection in more advanced lesions. There are no proven chemotherapeutic agents for carcinoma of the gall bladder. A small proportion of cases are sensitive to radiotherapy but the overall 5-year survival is less than 5%.

Cholangiocarcinoma (see also p. 364)

Cancer of the biliary tree may be intra- or extrahepatic. These malignancies represent approximately 1% of all cancers. A number of associations have been identified such as that with choledochal cyst (see above), and chronic infection of the biliary tree with, for example, *Clonorchis sinensis*. There are also associations with autoimmune disease processes such as primary sclerosing cholangitis. The bile duct malignancy usually presents with jaundice and may be suspected by imaging, initially ultrasound and thereafter spiral CT and in particular magnetic resonance cholangiopancreatography (MRCP). The disease spread is usually by local lymphatics or local extension. Cholangiocarcinoma of the common bile duct may be resectable at presentation but local extension precludes such management in the majority of more proximal lesions. Localized disease justifies an aggressive surgical approach including partial hepatic resection. Chemoradiation has been used to treat localized small hilar cholangiocarcinoma and in a few cases has facilitated successful liver transplantation, but the carcinoma recurs.

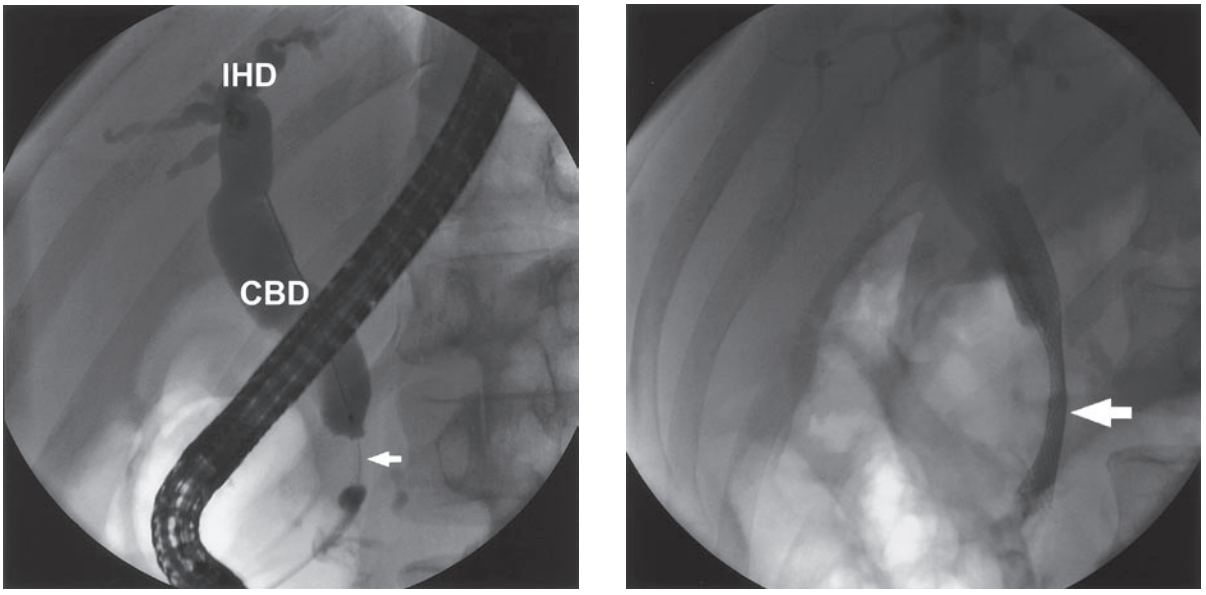
Secondary malignant involvement of the biliary tree

Carcinoma of the head of the pancreas frequently presents with common bile duct obstruction and jaundice. Metastases to the bile ducts from distant cancers are uncommon. Melanoma is the most frequent neoplasm to do so.

Other carcinomas that have caused bile duct metastases, in order of frequency, are those arising in the lung, breast and colon as well as those from the pancreas (metastatic as compared to direct infiltration). Infiltration of the bile duct is not uncommon in disseminated lymphomatous disease.

Palliation of malignant bile duct obstruction

A small proportion of cholangiocarcinomas are surgically resectable, more commonly those in the distal bile duct as compared to the hilar region. All patients must be fully screened for operability using the imaging techniques described above. However, in the greater proportion of patients the treatment is palliative. Relief of bile duct obstruction has been shown to improve quality of life considerably and with pain control is the major end point of palliation. In recent years endoscopic techniques have allowed the insertion of stents into the biliary tree to re-establish bile flow. The initial use of plastic stents has largely been replaced by self-expanding metal stents which have considerably longer periods of patency (Fig. 7.35). In the small proportion of patients in whom bile duct drainage is not possible endo-



(a) (b)

Fig. 7.35 **Stents. (a)** An endoscopic cholangiogram showing a distal stricture of the common bile duct (CBD) (arrow). This is secondary to a carcinoma of the head of pancreas. The CBD and intrahepatic ducts (IHD) are dilated. A guide wire has been passed across the stricture. **(b)** In the same patient a self-expanding metal stent (arrow) has been inserted re-establishing drainage of the biliary tree.

scopically, the percutaneous route offers an alternative method of stent placement. There is some evidence of benefit from the use of photodynamic therapy in those patients in whom biliary drainage has been achieved. This technique involves the use of a porphyrin derivative to sensitize the malignant cells prior to activation by an endoscopically placed laser probe. The aim is to provide local tumour destruction and maintain bile duct patency.

THE PANCREAS

STRUCTURE AND FUNCTION

Structure

The pancreas extends retroperitoneally across the posterior abdominal wall from the second part of the duodenum to the spleen. The head is encircled by the duodenum; the body, which forms the main bulk of the organ, ends in a tail that lies in contact with the spleen. The pancreas consists of exocrine and endocrine cells, the former making up 98% of the human pancreas.

The pancreatic acinar cells are grouped into lobules, forming the ductal system which eventually joins into the main pancreatic duct.

The main pancreatic duct has many tributary ductules and gradually tapers towards the tail of the pancreas. The main pancreatic duct itself usually joins the common bile duct to enter the duodenum as a short single duct at the ampulla of Vater.

Exocrine function

The pancreatic acinar cells are responsible for production of digestive enzymes. These include amylase, lipase, colipase, phospholipase and the proteases (trypsinogen and chymotrypsinogen). These enzymes are stored within the acinar cells in secretory granules and are released by exocytosis (Fig. 7.36).

After ingestion of a meal, pancreatic exocrine secretion is regulated by cephalic, gastric and intestinal stimuli. The cephalic phase is mediated by the central nervous system and is stimulated by behavioural cues related to the sight and smell of food. With ingestion of food, the gastric phase commences and in response to distension of the stomach a neural pathway involving the central nervous system stimulates pancreatic secretion. Both these phases are under vagal control. Finally, the presence of protein, fat and gastric acid within the small intestine further augments pancreatic secretion by both hormonal and neurotransmitter activity which produces local enteropancreatic control of secretion. Feedback regulatory events eventually terminate pancreatic secretion.

Cholecystokinin (CCK) is produced in specialized gut endocrine cells (I cells) of the mucosa of the small intestine and is secreted in response to intraluminal food. In animals, it exerts its biological activity by binding to specific G-protein-coupled receptors on target cells in the pancreas. Activated G-proteins lead to the activation of phospholipases. This in turn leads to calcium release from intracellular stores, which in turn results in the fusion of the digestive enzyme granules to the apical plasma membrane and enzyme release. There are no CCK receptors in pancreatic cells in humans, and CCK acts via receptors on vagal afferent fibres to stimulate pancreatic secretion. Of the enzymes produced by the pancreatic acinar cells, the proteases and colipase are secreted as inactive precursors and require duodenal enterokinase to initiate activity.

Secretin is also released from specialized enteroendocrine cells of the small intestine during a meal and in particular during duodenal acidification. Secretin has a direct effect on the pancreatic acinar cells as well as the ductal cells. There is also a vagal-mediated secretory response. Secretin action is mediated via G-coupled receptors and calcium-mediated enzyme release. Secretin results in a bicarbonate-rich pancreatic secretion.

Completion of the postprandial secretory phase involves both neural and hormonal control.

Central neuronal inhibition of pancreatic secretion acts through dopamine and somatostatin receptors mediated by noradrenergic nerves.

Pancreatic polypeptide from the islet cells is released from the pancreas in response to a meal and has an inhibitory effect upon acinar enzyme secretion both by a local effect and via central receptors.

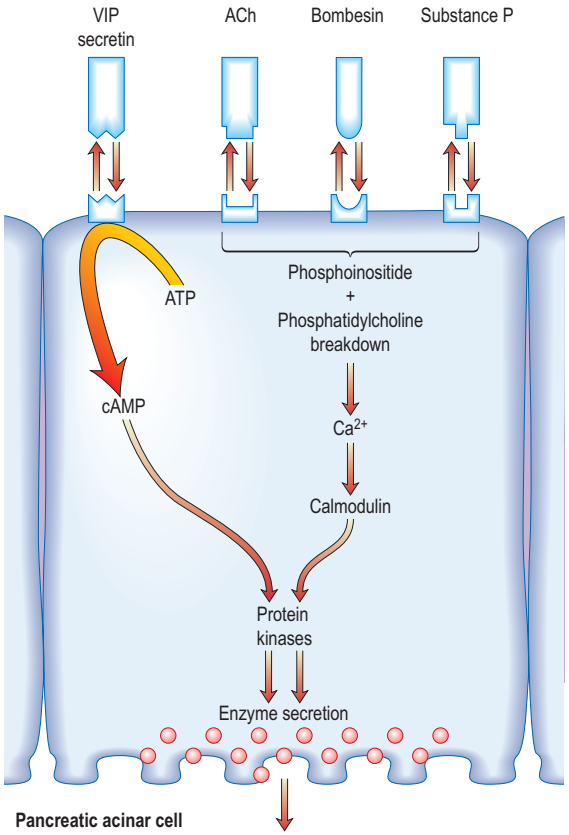


Fig. 7.36 **Diagram showing stimulus–secretion coupling of pancreatic cell protein secretion.** There is no CCK receptor in humans; stimulation is probably via neural fibres. VIP, vasoactive intestinal polypeptide; CCK, cholecystokinin; ACh, acetylcholine.

Ask the authors
www.studentconsult.com

Somatostatin, present within the pancreas, stomach and central nervous system, is released in response to food. Its effect is mediated both by direct pancreatic acinar inhibition and by a central nervous system effect.

Two other mechanisms of inhibition have been described. *Proteases* within the duodenal lumen have a negative feedback on acinar secretion. Secondly, nutrients within the ileum inhibit pancreatic secretion by means of local hormone release (peptide YY and glucagon-like peptide) acting on the acinar cells themselves as well as centrally.

There is increasing evidence that the gut-related peptides, leptin and ghrelin, as well as influencing appetite behaviour, are also regulatory factors in the exocrine function of the pancreas. This effect is believed to occur via hypothalamic centres.

The endocrine pancreas

This consists of hormone-producing cells arranged in nests or islets (islets of Langerhans). The hormones produced are secreted directly into the circulation and there is no access to the pancreatic ductular system. There are *five main types of islet cell* corresponding to different secretory components. The *beta cells* are the most common and are responsible for insulin production. The *alpha cells* produce glucagon. The *D cells* produce somatostatin, *PP cells* produce pancreatic polypeptide and *enterochromaffin cells* produce serotonin.

A number of other hormones have been identified within the endocrine pancreas including gastrin-releasing peptide, neuropeptide Y and galanin. These are believed to be neurotransmitters active in the neuro-gastrointestinal axis.

INVESTIGATION OF THE PANCREAS

Assessment of exocrine function

The assessment of pancreatic exocrine function is used in the investigation of patients with possible chronic pancreatic disease. Clinically evident fat malabsorption does not occur until there has been an 85–90% reduction in pancreatic lipase and is therefore a very late manifestation of pancreatic disease.

Direct tests of pancreatic function

These tests rely upon the analysis of a duodenal aspirate following pancreatic stimulation.

The original test involved the oral administration of a specified meal (Lundh meal). Pancreatic stimulation is now achieved by intravenous secretin and cholecystokinin.

The aspirate is assessed for pancreatic enzymes and bicarbonate production. The procedure is time-consuming and requires a meticulous technique. There is a good correlation with moderate to severe pancreatic function loss, but not for mild damage. These tests are not widely available.

The measurement of peak bicarbonate secretion following secretin stimulation has recently been described using an endoscopic technique for aspirate collection. This method offers similar levels of predictive accuracy as seen with the secretin–cholecystokinin stimulation test but does require a 30 minute endoscopic intubation.

Non-invasive indirect tests of pancreatic function

Faecal tests

- **Faecal fat estimation** (see p. 277).
- **Faecal chymotrypsin.** The test is not useful until severe impairment of pancreatic function is present.
- **Faecal elastase.** This pancreatic specific enzyme is not degraded in the intestine and has high concentrations within the faeces. Diminished levels may be detected in moderate as well as severe pancreatic insufficiency. This has replaced the faecal chymotrypsin test.

Oral pancreatic function tests

- *PABA test*. Oral *N*-benzoyl-L-tyrosyl-*p*-aminobenzoic acid is hydrolysed by chymotrypsin to release *p*-aminobenzoic acid (PABA), which is then absorbed, conjugated and excreted in the *urine* where it can be measured. The test is time-consuming but is specific for pancreatic insufficiency, with a 65–80% sensitivity.
- *Fluorescein dilaurate test*. Oral fluorescein dilaurate is digested by pancreatic esterase to release the fluorescein which is then absorbed and excreted in the *urine*. This test is relatively inexpensive and commercially available as the 'Pancreolauryl test'. This is highly sensitive and specific in severe pancreatic insufficiency but has only a 50% sensitivity in mild to moderate disease.

Clinical application of pancreatic function tests

Whilst the invasive duodenal aspiration tests represent the most sensitive and specific means of assessing pancreatic function, these are very rarely used outside specialized centres. The Pancreolauryl and PABA tests are non-invasive and are widely available but are only highly sensitive in the detection of severe pancreatic insufficiency. The faecal elastase test (in a commercially available form) provides similar sensitivity and specificity and is the test of choice as a screening tool for pancreatic insufficiency, but again detection of mild disease is problematic.

Pancreatic imaging (see p. 325)

Imaging has a pivotal role in the investigation and management of pancreatic disease, which covers the spectrum of acute, chronic and malignant conditions.

- *A plain abdominal radiograph* may show the calcification associated with chronic pancreatitis, particularly when alcohol is the aetiology.
- *Ultrasound* of the pancreas is a useful screening investigation for inflammation and neoplasia. Views may be limited by overlying bowel gas.
- *Spiral CT scan* with contrast enhancement and following a specific pancreatic protocol remains the gold standard imaging technique for the investigation of pancreatic disease.
- *MRI scanning* represents an alternative to CT. Magnetic resonance cholangiopancreatography (MRCP) gives clear definition of the pancreatic duct as well as the biliary tree. Gallstones (including microcalculi) may also be identified in the biliary tree using MRI/MRCP.
- *Endoscopic ultrasound* is very useful for identifying distal common bile duct stones which may be the

aetiology of an episode of acute pancreatitis. Endoscopic ultrasound can identify the early changes of chronic pancreatitis before these are evident on other imaging methods. There is also an increasing role for this technique to stage the operability of pancreatic adenocarcinoma, particularly with respect to vascular invasion. Endoscopic ultrasound is now considered the imaging technique of choice for investigating cystic lesions of the pancreas (see below). The technique allows fine-needle aspiration and histological sampling as well as the therapeutic option of cyst drainage. Endoscopic ultrasound is a sensitive means of detecting small pancreatic tumours, particularly those of neuroendocrine origin.

- *Endoscopic retrograde cholangiopancreatography* (ERCP) was considered the gold standard for diagnosing pancreatic disease. However, with the advent of MRCP and endoscopic ultrasound, ERCP is restricted to therapeutic intervention.

Summary. An initial transabdominal ultrasound supplemented by spiral CT provides sufficient diagnostic information for most inflammatory and neoplastic conditions of the pancreas. MRI and MRCP are now widely available and provide additional information, particularly with respect to pancreatic ductular and biliary anatomy. Endoscopic ultrasound is now widely available and has become a useful tool in the investigation of possible pancreatic malignancy, particularly for preoperative assessment.

PANCREATITIS

Classification

Pancreatitis is divided into acute and chronic. By definition acute pancreatitis is a process that occurs on the background of a previously normal pancreas and can return to normal after resolution of the episode. In chronic pancreatitis there is continuing inflammation with irreversible structural changes.

In practice the differentiation between acute and chronic pancreatitis may be extremely difficult. Any of the causes of acute pancreatitis if untreated may result in recurrent episodes classified as acute relapsing pancreatitis. In other cases the recurrent episodes of recurrent pancreatitis may represent exacerbations of an underlying chronic process.

Acute pancreatitis

The causes of acute pancreatitis are listed in Table 7.18. In the western world, gallstones and alcohol account for the vast majority of episodes. Alcohol also causes chronic pancreatitis (see below). The severity of the pancreatitis may range from mild and self-limiting to extremely severe with extensive pancreatic and peripancreatic necrosis as well as haemorrhage. In the most severe form of pancreatitis the mortality rises to between 40% and 50%.

Pathogenesis

Mechanisms by which pancreatic necrosis occurs remain speculative. Any theory must take into account how a very diverse group of aetiological factors can produce the same end point. There is some suggestion that the final common pathway is a marked elevation of intracellular calcium which in turn leads to activation of intracellular proteases. It is these activated enzymes which are responsible for cellular necrosis. In the case of gallstone-related pancreatitis it is believed that stones occlude the pancreatic drainage at the level of the ampulla leading to pancreatic ductular hypertension. Such ductular hypertension has been shown in animal models to increase cytosolic free ionized calcium. There is also evidence that alcohol interferes with calcium homeostasis in pancreatic acinar cells.

Table 7.18 Causes of pancreatitis

Acute	Chronic
Gallstones	Alcohol
Alcohol	Tropical
Infections (e.g. mumps, Coxsackie B)	Hereditary
Pancreatic tumours	Trypsinogen and inhibitory protein defects
Drugs (e.g. azathioprine, oestrogens, corticosteroids, didanosine)	Cystic fibrosis
Iatrogenic (e.g. post-surgical, post-ERCP)	Idiopathic
Hyperlipidaemias	Trauma
Miscellaneous	Hypercalcaemia
Trauma	
Scorpion bite	
Cardiac surgery	
Idiopathic	

Clinical features

Acute pancreatitis is a differential diagnosis in any patient with upper abdominal pain. The pain usually begins in the epigastrium accompanied by nausea and vomiting. As inflammation spreads throughout the peritoneal cavity the pain becomes more intense. Involvement of the retroperitoneum frequently leads to back pain.

The patient may give a history of previous similar episodes or be known to have gallstones. An attack may follow an alcoholic binge. However, in many cases there are no obvious aetiological factors.

Physical examination at the time of presentation may show little more than a patient in pain with some upper abdominal tenderness but no systemic abnormalities. In more severe disease the patient may have a tachycardia, hypotension and be oliguric. Abdominal examination may show widespread tenderness with guarding as well as reduced or absent bowel sounds. Specific clinical signs that support a diagnosis of severe necrotizing pancreatitis include periumbilical (Cullen's sign) and flank bruising (Grey Turner's sign). In patients with a gallstone aetiology the clinical picture may also include the features of jaundice or cholangitis.

Diagnosis

Blood tests

- *Serum amylase* is an extremely sensitive test if it is 3 times the upper limit of normal when measured within 24 hours of the onset of pain. A number of other conditions may occasionally cause a very elevated amylase (Table 7.19). Amylase levels gradually fall back towards normal over the next 3–5 days. With a late presentation the serum amylase level may give a false-negative result.
- *Urinary amylase* levels may be diagnostic as these remain elevated over a longer period of time.

Table 7.19 Elevation of serum amylase unrelated to pancreatitis

Leakage of upper gastrointestinal contents into the peritoneum
Upper gastrointestinal perforation
Biliary peritonitis
Intestinal infarction
Inherited abnormalities of amylase
Macroamylasaemia

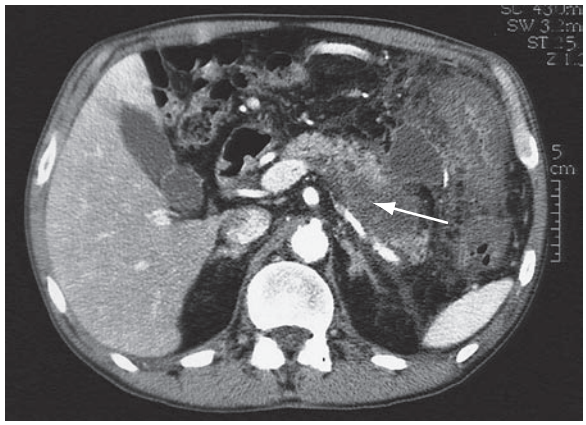


Fig. 7.37 **CT scan in patient with acute pancreatitis** showing necrosis of the pancreatic parenchyma (arrow) and a fluid collection extending outside the gland with inflammatory thickening of the colon.

- *Serum lipase* levels are also raised in acute pancreatitis and these remain elevated for a longer period of time than those of amylase. However, overall, the accuracy of serum lipase is not significantly greater than amylase and it is technically more difficult to measure.
- *C reactive protein level* is useful in assessing disease severity and prognosis.
- *Other baseline investigations* include a full blood count, urea and electrolytes, blood glucose, liver biochemistry, plasma calcium and arterial blood gases. These are documented at presentation and then repeated at 24 and 48 hours and provide a basis for assessing the severity of an attack (see below).

Radiology

- An erect *chest X-ray* is mandatory to exclude gastroduodenal perforation, which also raises the serum amylase (see Table 7.19). A supine abdominal film may show gallstones or pancreatic calcification.
- An *abdominal ultrasound scan* is used as a screening test to identify a possible biliary (gallstone) cause of pancreatitis. Gallstones are difficult to detect in the distal common bile duct but dilated intrahepatic ducts may be present in the presence of bile duct obstruction. Stones within the gall bladder are not sufficient to justify a diagnosis of gallstone-related pancreatitis. The ultrasound may also demonstrate pancreatic swelling and necrosis as well as peripancreatic fluid collections if present. In severe pancreatitis the pancreas may be difficult to visualize because of gas-filled loops of bowel.
- *Contrast-enhanced spiral CT scanning* is essential in all but the most mild attacks of pancreatitis (Fig. 7.37). It should be performed after 72 hours to assess the extent of pancreatic necrosis. CT provides very valuable prognostic information. Later, repeated CT scans can detect other complications including fluid collections, abscess formation and pseudocyst development.
- *MRI (MRCP)* assesses the degree of pancreatic damage and identifies gallstones within the biliary tree. MRI is particularly useful to differentiate between fluid and solid inflammatory masses.
- *ERCP* is used as a treatment measure to remove bile duct stones in the presence of gallstone-related pancreatitis (see below).

Table 7.20 Severe pancreatitis – factors during the first 48 hours that indicate severe pancreatitis and a poor prognosis (three or more factors present predict a severe episode)

Age	>55 years
WBC	$>15\times10^9$/L
Blood glucose	>10 mmol/L
Serum urea	>16 mmol/L
Serum albumin	<30 g/L
Serum aminotransferase	>200 U/L
Serum calcium	<2 mmol/L
Serum LDH	>600 U/L
P_aO_2	<8.0 kPa (60 mmHg)

LDH, lactate dehydrogenase

Assessment of disease severity

The majority of cases of acute pancreatitis are mild but approximately 25% run a more complicated course which may result in haemodynamic instability and multiple organ failure. The early prediction of such a severe attack allows appropriate monitoring and intensive care to be in place.

Early clinical assessment has been shown to have poor sensitivity for predicting a severe attack. Similarly, individual laboratory tests have very limited value. Elevations of CRP of >200 mg/L in the first 4 days have an 80% predictive value of a severe attack. Multiple factors are also used to develop scoring systems (Table 7.20). The **Ranson and Glasgow scoring systems** are based on such parameters and have been shown to have an 80% sensitivity for predicting a severe attack, although only after 48 hours following presentation. The **acute physiology and chronic health evaluation (APACHE) score** has been extensively adopted as a means of assessing the severity of a wide spectrum of illness. The APACHE scoring system is based on common physiological and laboratory values and adjusted for age as well as the presence or absence of a number of other chronic health problems (see Table 7.21). This scoring system appears to have a high sensitivity as early as 24 hours after onset of symptoms. There is evidence that obesity predicts the outcome from an episode of pancreatitis as the excessive adipose tissue is a substrate for activated enzyme activity. This will in turn generate an extensive inflammatory reaction. Even modest obesity (body mass index (BMI) between 25 and 30) has an adverse effect. This is incorporated as an adverse factor in the APACHE score (p. 920) for acute pancreatitis, and other variables have also been added.

Table 7.21 The APACHE II scoring system parameters

Physiological	Laboratory
Temperature	Oxygenation (P_aO_2)
Heart rate	Arterial pH
Respiratory rate	Serum:
Mean arterial pressure	sodium
Glasgow Coma Scale	potassium
	creatinine
	Haematocrit
	White blood cell count

Score 0–4 (normal–abnormal)
Adjust for age and severe organ insufficiency or for immunocompromised. **BMI** is an additional parameter

Treatment

The initial management of acute pancreatitis is similar, whatever the cause. A multiple factor scoring system (ideally APACHE II with a modification for obesity) should be used at the end of the first 24 hours after presentation to allow identification of the 25% of patients with a predicted severe attack. This should be repeated at 48 hours to identify a further subgroup who appear to be moving into the severe category. These patients should then be managed on a high-dependency or intensive care unit. Even patients outside the severe category may require considerable supportive care.

Early fluid losses in acute pancreatitis may be large, requiring well-maintained intravenous access as well as a central line and urinary catheter to monitor circulating volume and renal function.

- *Nasogastric suction* prevents abdominal distension and vomitus and hence the risk of aspiration pneumonia.
- *Baseline arterial blood gases* determine the need for continuous oxygen administration.
- *Prophylactic antibiotics*. Broad-spectrum antibiotics, e.g. cefuroxime, or aztreonam, along with an agent active against Gram-positive cocci, reduce the risk of infective complications and are given from the outset.
- *Analgesia requirements*. Pethidine and tramadol are the drugs of choice for immediate post-presentation pain control. Unless there is prompt resolution of pain, a patient-controlled system of administration is indicated to provide continuous and adequate pain relief. Fentanyl has been used widely for this application. There is a theoretical risk that morphine and diamorphine might exacerbate pancreatic ductular hypertension by causing sphincter of Oddi contraction and they are best avoided in acute pancreatitis.
- *Feeding*. In patients with a severe episode there is little likelihood of oral nutrition for a number of weeks. Total parenteral nutrition has been associated with a high risk of infection and has been replaced by enteral nutrition. In the absence of gastroparesis most patients will tolerate nasogastric administration of feed without exacerbation of pain. In those with gastroparesis or poorly tolerated nasogastric feeding (exacerbation of pain or precipitation of nausea and vomitus), postpyloric feeding should be instituted by the endoscopic placement of a nasojejunal tube.
- *Anticoagulation* with a low molecular weight heparin for DVT prophylaxis.

In a small proportion of patients, multiorgan failure will develop in the first few days after presentation, reflecting the extent of pancreatic necrosis. Such patients will require positive-pressure ventilation and often renal support. The mortality in this group is extremely high (in excess of 80%).

Gallstone-related pancreatitis

In patients with gallstone-related pancreatitis and associated bile duct obstruction (particularly when complicated by cholangitis), endoscopic intervention with sphincterotomy and stone extraction is the treatment of choice. In the absence of bile duct obstruction, sphincterotomy and stone extraction is only of proven benefit when the episode of pancreatitis is predicted as severe. In less severe cases of gallstone-related pancreatitis the presence of residual bile duct stones can be assessed by MRCP or endoscopic ultrasound after recovery from the episode of pancreatitis (usually a 6 week period is allowed) and, if present, removed at the time of ERCP. To prevent a recurrent episode of pancreatitis cholecystectomy should be carried out at a similar time interval.

Late complications of acute pancreatitis

Within the first 7 days the morbidity and mortality of acute pancreatitis reflect the systemic inflammatory response, which in turn results in multiple organ failure. After this initial period the prognosis thereafter is most closely related to the extent of pancreatic necrosis. This can be most accurately assessed by contrast-enhanced CT, which should be carried out in all patients with severe disease after the first week. Extensive necrosis (greater than 50% of the pancreas) is associated with high risk of further complicated disease, frequently requiring surgical intervention.

It is infection of the necrotic pancreas which is of most concern and which may rapidly lead to overwhelming sepsis. Prophylactic antibiotics have been used to prevent this but do not reliably do so. If there is evidence of incipient infection, monitored by a rising neutrophil count and CRP level, an aspirate of the necrotic pancreas is taken under ultrasound control and cultured. The vast majority of patients with a positive culture should be considered for surgical resection of the necrotic pancreas. In the most severe cases multiple operations are required to fully resect the areas of necrosis.

Peripancreatic fluid collections are common in the early stages of acute pancreatitis; the vast majority will resolve spontaneously. Some fluid collections will be surrounded by granulation tissue producing the so-called 'pseudocyst'. These by definition are not fully formed until 6 weeks after the onset of the illness. The smaller pseudocysts (less than 6 cm in diameter) frequently resolve on their own but others may persist in the long term, giving rise to potential complications such as infection, intraperitoneal bleeding and gastric outlet obstruction. Larger pseudocysts persisting for longer than 6 weeks are usually actively managed by endoscopic drainage under endoscopic ultrasound guidance. Surgical drainage may be indicated in cases in which the cyst anatomy is not compatible with endoscopic drainage.

Long-term outcome

The vast majority of patients with a mild to moderate episode of acute pancreatitis will make a full recovery with no long-term sequelae. Recurrent episodes of pancreatitis may occur, particularly if there has been any long-term pancreatic ductular damage. Patients with more severe acute pancreatitis may become pancreatic insufficient both with respect to exocrine (malabsorption) and endocrine function (diabetes).

Chronic pancreatitis

Aetiology

In developed countries by far the most common cause of chronic pancreatitis is alcohol, accounting for 60–80% of cases (see Table 7.18).

Tropical chronic pancreatitis is commonly found in children and young adults. The aetiology is unknown but there is some evidence of a genetic susceptibility with increased prevalence of mutations of the trypsin inhibitor gene *SPINK-1*. There is a strong likelihood of an interaction between genetic susceptibility and environmental triggers.

Hereditary chronic pancreatitis is an autosomal dominant condition with variable penetrance. The gene mutations have been mapped to a cluster of trypsinogen genes situated on the long arm of chromosome 7.

Autoimmune chronic pancreatitis also occurs. This is seen predominantly in middle-aged men and associated with raised titres of anti-nuclear factor and particularly raised levels of IgG_4. As well as pancreatic-related complications these patients often have other autoimmune phenomena such as a cholangiopathy, Sjögren's syndrome and renal disease.

Cystic fibrosis (p. 844). Almost all patients have established chronic pancreatitis, usually from birth. Cystic fibrosis gene mutations have also been identified in patients with chronic pancreatitis but in whom there were no other manifestations of cystic fibrosis.

Obstruction of the pancreatic duct because of either a benign or malignant process may result in chronic pancreatitis. Congenital abnormalities of the pancreatic duct, in particular pancreas divisum, have been implicated.

Pathogenesis

A possible common pathway for pancreatic damage is the inappropriate activation of enzymes within the pancreas. This has been well demonstrated in the case of hereditary pancreatitis where genetic abnormalities of cationic trypsinogen and its inhibitory proteins have led to unopposed trypsin activity within the pancreas itself.

Chronic alcohol intake is also believed to increase the level of trypsinogen relative to its inhibitor. Human trypsinogen has a propensity to autoactivate, and any relative impairment or deficiency of inhibitor proteins will lead to unopposed enzyme activity and possible pancreatic damage (Fig. 7.38).

It is believed that the intrapancreatic enzyme activity leads to the precipitation of proteins within the duct lumen in the form of plugs. These then form a nidus for calcification but are also the cause of ductal obstruction leading to ductal hypertension and further pancreatic damage. Cytokine activation and oxygen stress are thought to play a role in perpetuating this process.

Clinical features

Pain is the most common presentation of chronic pancreatitis. It is usually epigastric and often radiates through into the back. The pattern of pain may be episodic, with short periods of severe pain, or chronic unremitting. Exacerbations of the pain may follow further alcohol excess although this is not a uniform relationship.

During periods of abdominal pain anorexia is common and weight loss may be severe. This is particularly so in those patients with chronic unremitting symptoms. Exocrine and endocrine insufficiency may develop at any time, and occasionally malabsorption or diabetes is the presenting feature in the absence of abdominal pain. Jaundice secondary to obstruction of the common bile duct during its course through the fibrosed head of pancreas may also occur and may be a presenting feature in a small proportion of patients.

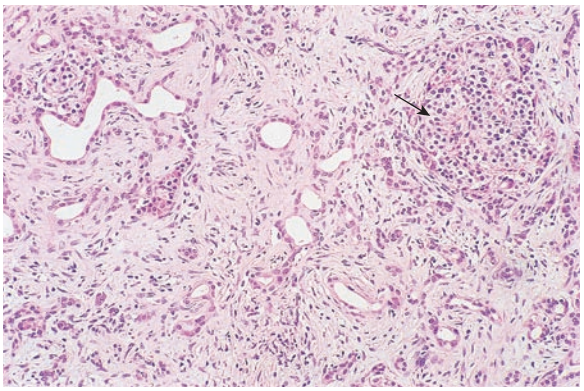


Fig. 7.38 **Histology of chronic pancreatitis.** There is considerable loss of acini and replacement by fibrosis. Inflammatory cells are relatively inconspicuous at this late stage. Islets of Langerhans (one is arrowed) sometimes escape destruction but their loss can result in diabetes mellitus. From Underwood JC (ed) *General and Systematic Pathology*, 4th edn. Edinburgh: Churchill Livingstone, 2004, (Fig. 16.25). Copyright © Elsevier 2004.

Investigations

The extent to which investigations are required is dependent upon the clinical setting. In a patient with known alcohol abuse and typical pain, few confirmatory tests are required.

- *Serum amylase and lipase* levels are rarely significantly elevated in established chronic pancreatitis.
- *Faecal elastase* level will be abnormal in the majority of patients with moderate to severe pancreatic disease.
- *PABA and Pancreolauryl tests* (see p. 376).
- *Transabdominal ultrasound scan* is used for initial assessment.
- *Contrast-enhanced spiral CT scan* provides a more detailed assessment. In the presence of pancreatic calcification and a dilated pancreatic duct the diagnosis of chronic pancreatitis can be easily established (Fig. 7.39). This may be much more difficult when these features are not present and in particular with an atypical presentation such as with steatorrhoea alone.
- *MRI with MRCP* is increasingly utilized to define more subtle abnormalities of the pancreatic duct which may be seen in non-dilated chronic pancreatitis.
- *Endoscopic ultrasound* is used in a small proportion of patients in whom the diagnosis is not confirmed with the above imaging or specifically for assessing complications of chronic pancreatitis including pseudocyst formation.
- *Diagnostic ERCP* has been replaced by MRCP.

Differential diagnosis

The differential diagnosis is that of pancreatic malignancy. Carcinoma of the pancreas can reproduce many of the symptoms and imaging abnormalities that are commonly seen with chronic pancreatitis. The diagnosis of malignancy

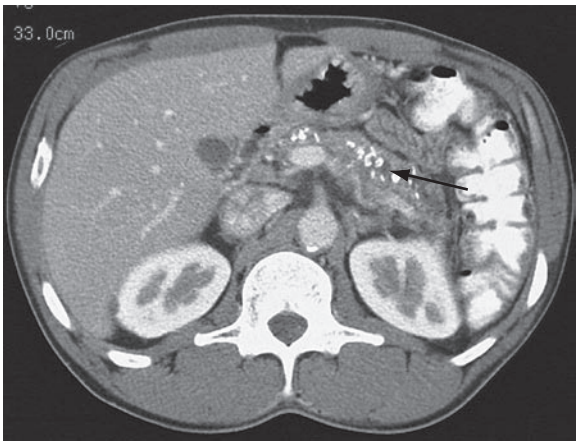


Fig. 7.39 **Contrast-enhanced CT scan demonstrating multiple calcific densities (arrow) along the line of the main pancreatic duct in a patient with chronic pancreatitis.**

should be considered in patients with a short history and in whom there is a localized ductular abnormality. Considerable difficulties may arise when a malignancy develops on the background of established chronic pancreatitis (the latter being a recognized premalignant lesion).

High-quality imaging is able to define malignant features with a localized mass lesion, local invasion and lymph node enlargement. Endoscopic ultrasound may provide the most accurate assessment of a potential mass lesion.

Treatment

In patients with alcohol-related chronic pancreatitis long-term abstinence is likely to be of benefit although this has been difficult to prove.

Abdominal pain. For short-term flare-ups of pain a combination of a non-steroidal anti-inflammatory drug and an opiate (tramadol) is usually sufficient for symptomatic relief. In patients with chronic unremitting pain this may be inadequate and also risks opiate addiction.

Tricyclic antidepressants (e.g. amitriptyline) are used for chronic pain and reduce the need for opiates. Antioxidants are not of proven benefit. Coeliac axis nerve block may produce good pain relief but is unreliable in its extent and duration of action. In the majority of patients some spontaneous improvement in pain control occurs with time. After a 6- to 10-year period some 60% of patients will become pain-free. For patients with recurrent severe or debilitating chronic pain, both endoscopic and surgical intervention has been used but with limited success. The endoscopic approach has centred upon improving duct drainage by removing intraductal stones and repeated stenting to maintain duct patency. Extracorporeal shock wave lithotripsy has been used to fragment stones within the head of pancreas. Surgical intervention usually involves a duct drainage procedure combined with partial resection of the diseased head of pancreas. There is recent evidence suggesting improved symptom control following surgical intervention as compared to the endoscopic approach. However, it is reasonable to attempt endoscopic therapy as a first measure.

Steatorrhoea. The steatorrhoea associated with pancreatic insufficiency may be high, with up to 30 mmol of fat lost per 24 hours. This will usually improve with pancreatic enzymes supplements. Current preparations are presented in the form of microspheres which reduce the problems of acid degradation in the stomach. An acid suppressor (H_2-receptor antagonist or proton pump inhibitor) is also given. Despite this, a proportion of patients continue to malabsorb, usually reflecting the inadequate mixing of the pancreatic supplements with the food as well as the low pH in the duodenum secondary to inadequate pancreatic bicarbonate production. There is no justification to reduce fat intake below the recommended levels of a normal diet as this may contribute to malnutrition seen in patients with chronic pancreatitis.

Diabetes associated with pancreatic endocrine failure may be difficult to control, with a rapid progression from oral hypoglycaemic agents to an insulin requirement. Brittle control is a common problem secondary to inadequate glucagon production from the damaged pancreas.

Complications

The most common structural complication of chronic pancreatitis is a pancreatic pseudocyst, a fluid collection surrounded by granulation tissue (see p. 379). These usually occur in relationship to a period of enhanced inflammatory activity within the pancreas giving abdominal pain but may develop silently during what would appear to be a stable phase. Intra- or retroperitoneal rupture, bleeding or cyst infection may occur. The larger cysts may occlude nearby structures including the duodenum and the bile duct. In pseudocysts less than 6 cm in diameter, spontaneous resolution can be anticipated. In larger cysts that have been present for a period in excess of 6 weeks, resolution is less common and a long-term complication rate of approximately 30% can be anticipated. Many pseudocysts are closely apposed to the posterior wall of the stomach or duodenum and can be successfully drained endoscopically using endoscopic ultrasound to identify the optimum drainage site. A direct fistula is created between the pseudocyst lumen and the gastric or duodenal lumen which is then kept patent by the insertion of plastic stents. This approach will be successful in approximately 75% of cases. Surgical drainage is required for failures of endoscopic therapy or in circumstances in which the pseudocyst anatomy does not allow endoscopic access.

Ascites and occasionally pleural effusions can be a direct consequence of chronic pancreatitis when there has been disruption of the main pancreatic duct. A high ascites or pleural fluid amylase will confirm the aetiology. Such disruptions of the main pancreatic duct require surgical intervention.

Cystic fibrosis (see pp. 46 and 844)

This common cause of pancreatic disease in childhood is inherited as an autosomal recessive. A specific gene mutation ΔF_{508} is present in 70% of cases. The gene(s) code for a membrane protein in epithelial cells which regulates chloride transport (the cystic fibrosis transmembrane regulator, CFTR). Defective chloride channel transport secondarily leads to a failure to hydrate pancreatic secretion. The increased viscosity of such secretions then leads to ductular obstruction and secondary pancreatic damage. Eighty-five per cent of patients with cystic fibrosis will have pancreatic failure, and in the majority of these this will develop in utero and be present from the perinatal period. It has recently been

FURTHER READING

Conwell DL, Banks PA. Chronic pancreatitis. *Current Opinion in Gastroenterology* 2008; **24**(5): 586–590.

Pannala R, Chari ST. Autoimmune pancreatitis. *Current Opinion in Gastroenterology* 2008; **24**(5):591–596.

recognized that a small proportion of patients with chronic pancreatitis of unknown aetiology (idiopathic) are homozygous for cystic fibrosis genes but have no other overt manifestation of cystic fibrosis. Furthermore there is accumulating evidence of patients with idiopathic pancreatitis who carry a single CF gene mutation. It is likely that other environmental or genetic factors combine with the single CF gene mutation to result in a diseased pancreas.

Clinical features and diagnosis

Clinical features and diagnostic tests are described on page 845; pancreatic function tests and imaging are described on page 376.

Treatment of cystic fibrosis

In the vast majority of patients with cystic fibrosis, pulmonary complications dominate the clinical picture. However, the management of pancreatic insufficiency is necessary to optimize the growth and overall nutrition. Pancreatic supplements are closely titrated against the level of steatorrhoea. Fat intake should be maintained to avoid nutritional deficit and enzyme intake adjusted to deal with this intake. To obtain optimal enzyme mixing with food, the supplements are taken throughout the meal. These are most commonly administered in an enteric-coated form to minimize degradation within the acid milieu of the stomach. A daily lipase intake of up to 10 000 units/kg bodyweight may be required. High-dose enteric-coated preparations are available but have been implicated in right-sided colon stricture formation in children with cystic fibrosis. The exact mechanism by which this occurred has not been fully defined but these preparations are no longer recommended in children.

Despite appropriate enzyme dosage, up to 20% of patients continue to have significant fat malabsorption. In many cases this is secondary to poor compliance. Other mechanisms include acid degradation of the enzyme preparations, and this can be ameliorated by acid suppression, either with H_2-receptor antagonists or proton pump inhibitors. There is documented evidence of an age-related deterioration in glucose tolerance in patients with cystic fibrosis. Ten per cent of patients will develop clinically significant diabetes mellitus. This must be detected at the earliest opportunity to prevent the adverse nutritional effects. In the large proportion of cases, management will be insulin dependent.

CARCINOMA OF THE PANCREAS

Incidence of pancreatic cancer in the West has been estimated at approximately 10 cases per 100 000 with no increase over the last 20 years. Pancreatic cancer is now the fifth most common cause of cancer death in the western world. The incidence increases with age and the majority of cases occur in patients over the age of 60. Approximately 60% of patients with this condition are male. Ninety-six per cent of pancreatic cancers are adenocarcinoma in type and the large majority are of ductal origin.

Risk factors. Smoking is associated with a twofold increase. Other possible environmental factors include the petroleum product, naphthalamine. Chronic pancreatitis is precancerous, particularly in patients with inherited trypsinogen gene mutations who have a 100 fold increased risk of malignancy.

There is increasing evidence that pancreatic cancer is a genetic disorder caused by germline and acquired somatic mutations associated with cancer development. Subtle changes in an oncogene and both mutations and deletions of tumour suppressor genes are the most common forms of genetic alteration. The best known among the oncogenes is K-*ras*, which is present in 90% of human ductal adenocarcinomas. Other oncogenes include β-catenin and tumour suppressor genes – *p16, p53* and *DPC4* (or SMAD4). The latter is a key transcription factor in growth regulation and is inactivated in over 50% of cases.

The accumulation of genetic precursors for pancreatic cancer is associated with progressive premalignant ductal histological changes. These are described as pancreatic intraepithelial neoplasias (PanINs). The changes are classified as 1A, 1B, -2 and -3, reflecting the progression from premalignancy to invasive cancer (Fig. 7.40).

A small percentage of pancreatic adenocarcinomas arise from cystic lesions including intraductal papillary mucinous tumours (IPMT) and mucinous cystic neoplasia (see below). There is evidence that these cystic neoplasms demonstrate a similar multistep genetic and histological progression to invasive adenocarcinoma.

Clinical picture

Pancreatic adenocarcinoma may be viewed clinically as two diseases – the lesions of the head and lesions of the body and tail.

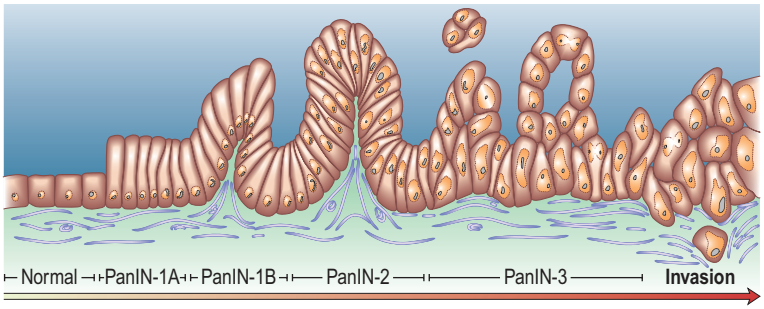


Fig. 7.40 **Genetic model for the development of pancreatic ductal adenocarcinoma.** The stages indicate histological progression with specific gene alterations. PanIN, pancreatic intraepithelial neoplasia. Modified from Maitra A, Adsay NV, Argani P et al. Multicomponent analysis of the pancreatic adenocarcinoma progression model using a pancreatic intraepithelial neoplasia tissue microarray. *Modern Pathology* 2003; **16**: 902–912 with permission from Nature Publishing Group.

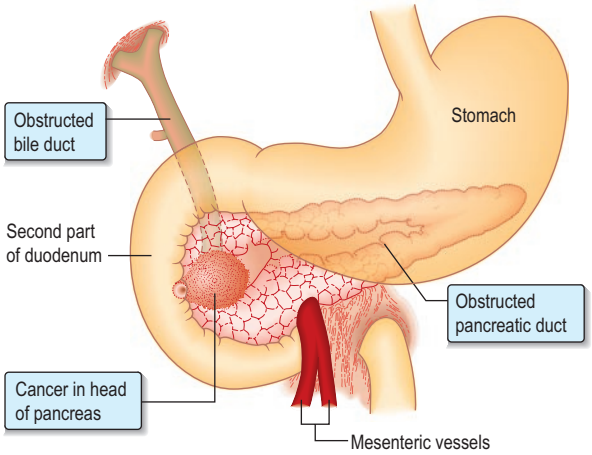


Fig. 7.41 **Carcinoma of the head of the pancreas.** A diagrammatic representation of the close relationship of a carcinoma of the head of pancreas to the surrounding structures.

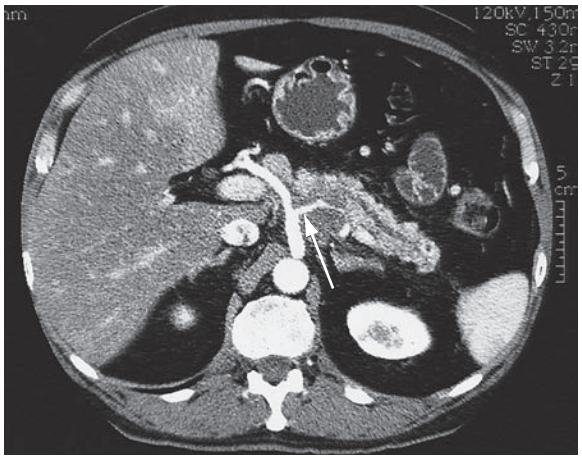


Fig. 7.42 **A contrast-enhanced CT scan showing a cancer of the body of the pancreas.** There is retroperitoneal tumour extension enclosing the branches of the coeliac axis (arrow).

Symptoms

Carcinoma of the head of pancreas or the ampulla of Vater tends to present earlier with obstruction to the bile duct as this passes through the head of pancreas giving jaundice (Fig. 7.41). These more localized lesions are usually painless, although pain may become a feature with tumour progression.

Carcinoma localized to the body or tail of the pancreas is much more likely to present with abdominal pain as well as non-specific symptoms such as anorexia and weight loss. The pain is often dull in character with radiation through into the back. A characteristic feature is partial relief of pain by sitting forward. Bile duct obstruction and jaundice may infrequently be late phenomena.

Physical signs

With *carcinoma of the head of pancreas* the patient is jaundiced with the characteristic scratch marks secondary to cholestasis. In a proportion of cases the gall bladder will be palpable (Courvoisier's sign). A central abdominal mass may be palpable as well as hepatomegaly if metastatic disease is present. With *carcinoma of the body and tail*, there are often no physical signs.

Other presenting physical signs include thrombo-embolic phenomena, polyarthritis and skin nodules. The latter are secondary to localized fat necrosis and associated inflammation. These manifestations, distant to the tumour itself, have not been fully explained but may precede the overt presentation of pancreatic cancer by months to years.

Investigations

- *Transabdominal ultrasound* is the initial investigation in the majority of patients. In the presence of bile duct obstruction this will confirm dilated intrahepatic bile ducts as well as a mass in the head of the pancreas. Ultrasound is less reliable when the cancer is found in the body and tail of the pancreas because of overlying bowel gas, with a sensitivity of detection of 60%.
- *Contrast-enhanced spiral CT scan* should confirm the presence of a mass lesion (Fig. 7.42). It is also necessary prior to possible surgical resection with contrast providing vascular definition to exclude tumour invasion as well as local lymph node involvement and distant metastases.
- *Laparoscopy* is also used for preoperative assessment.
- *ERCP* is usually restricted to palliative treatment but may provide a source of cytology to confirm the diagnosis when this is in question.
- *Percutaneous needle biopsy* is discouraged in potentially operable cases as this may be a source of tumour cell spread within the peritoneum. If palliative chemotherapy is considered, a histological diagnosis is essential prior to treatment.
- *MRI scanning and endoscopic ultrasound* are techniques that are useful in a small proportion of patients in whom the tumour has not been adequately defined.
- *Several tumour markers* have been evaluated for the diagnosis and monitoring of pancreatic cancer. The CA19-9 has a high sensitivity (80%) but a high false-positive rate. In individual patients single values of these tumour markers may be of little help but a progressive elevation over time is often diagnostic, and in such circumstances tumour marker levels can be used to monitor response to treatment.

Differential diagnosis

The diagnosis should not be difficult in the presence of painless jaundice or epigastric pain radiating into the back with progressive weight loss. Unfortunately many patients present with very minor symptoms including pain, change in bowel habit and weight loss. Imaging, particularly abdominal CT, should be performed if pancreatic cancer is suspected. Pancreatic cancer may rarely present with recurrent episodes of typical acute pancreatitis.

Management

The 5-year survival rate for carcinoma of the pancreas is approximately 2–5%, with surgical intervention representing the only chance of long-term survival. Approximately 20% of all cases have a localized tumour suitable for resection but in an elderly population many of these have co-morbid factors that preclude such major surgery. To optimize the percentage of patients undergoing possible surgical resec-

tion it is necessary to review each case in a multidisciplinary meeting. This approach also allows formulation of treatment strategies for those considered unsuitable for surgery.

In the majority of cases the management is palliative. Jaundice is a debilitating complication, often associated with severe pruritus but also the cause of non-specific malaise, lethargy and anorexia. Endoscopic placement of endoprostheses (stents) offers excellent palliation with a low associated procedural morbidity and mortality.

Palliative surgery has a role in duodenal obstruction (a complication seen in 10% of cases) but in advanced disease self-expanding metal stents can be placed across the duodenal obstruction with excellent short-term results.

The results of radiotherapy have been disappointing but there is now increasing evidence of benefit from a number of chemotherapeutic agents. 5-Fluorouracil and gemcitabine have been shown to improve short-term survival in advanced disease and have also demonstrated survival benefit as an adjuvant therapy to pancreatic resection.

With disease progression, abdominal pain is a frequent complicating factor which may prove extremely difficult to treat. This is best managed by experienced palliative care teams which offer a multidisciplinary approach. Endocrine and exocrine pancreatic failure occur and are managed as described on page 381.

CYSTIC TUMOURS OF THE PANCREAS

Cystic lesions of the pancreas are not uncommon. Seventy-five per cent of these lesions will be pseudocysts (see above) but of the remainder the majority are true cystic neoplasms.

Serous cyst adenomata are composed of multiple small cystic cavities lined by cuboidal glycogen-rich, mucin-poor cells. These lesions tend to occur in an elderly age group and are often an asymptomatic finding. Malignant transformation in a serous cystadenoma is extremely rare. Larger serous cystadenomata may cause local compressive complications (when over approximately 5 cm).

Mucinous cyst adenomata are almost exclusively found in women in the fifth and sixth decades and are sited in the pancreatic body and tail. Multilocular cysts are lined by tall mucin-synthesizing cells. Twenty per cent of these lesions are malignant at the time of presentation and the majority appear to have a malignant potential. As a consequence they are much more likely to produce symptoms.

Intraductal papillary mucinous tumour (IPMT) is a pancreatic cystic neoplasm that can arise from either the side branches or the main duct. The majority are found in men between the ages of 60 and 70. The presentation is usually with pancreatic pain but may be an incidental finding. The lesion is slowly progressive with a significant malignant potential; this appears to be more so when the main duct is the site of origin.

There is a high risk of malignancy in cystic lesions and therefore resection is desirable. The differentiation between pseudocysts and true cystic neoplasms is difficult, even with multiple imaging techniques. Patients with pseudocysts may have a history of pancreatitis. Endoscopic ultrasound scanning and fine-needle aspiration can be helpful in the differentiation. The measurement of cyst fluid CEA and CA19-9 may help to identify malignant change.

NEUROENDOCRINE TUMOURS OF THE PANCREAS

The islets of Langerhans (p. 376) have the capacity to synthesize more than one hormone. They also synthesize ectopic hormones that are not usually found in the pancreas such as gastrin, adrenocorticotrophin, vasoactive intestinal peptide and growth hormone. Whilst many pancreatic endocrine tumours are multihormonal, one peptide tends to predominate and is responsible for the clinical syndrome. Other tumours, whilst containing peptide hormone, are functionally inactive.

These tumours are rare with an incidence of less than 1 in 100 000 of the population. Insulinomas are the commonest variant (50%), gastrinomas account for 20% and the rarer functioning tumours 5%. The remaining 25% are non-functioning tumours. Approximately 25% of islet cell tumours are associated with a multiple endocrine syndrome (type I) (see p. 1026). The majority of the endocrine pancreatic tumours are malignant in their behaviour.

Presentation of pancreatic neuroendocrine tumours is most commonly related to the actions of ectopic hormone secretion. Identification of the primary and possibly metastatic lesions may be difficult despite multiple imaging techniques. Endoscopic ultrasound may be the most sensitive means of detecting small tumours. Many of these tumours have somatostatin receptors, and radiolabelled somatostatin analogue (such as octreotide) scanning provides a means of tumour localization.

Treatment options for pancreatic neuroendocrine tumours require a multidisciplinary approach and depend upon the presence or absence of metastatic (usually hepatic) disease. *Surgical resection* of the pancreatic lesion is the only potential curative approach. Aggressive surgical intervention including a resection of the primary lesion as well as liver resection for metastasis has been used in selected cases. Somatostatin analogues such as octreotide and lanreotide have been used specifically for the control of symptoms secondary to the hormonal secretion.

There is some evidence that the somatostatin analogues combined with interferon-alfa also control tumour proliferation. The chemotherapeutic agents streptozotocin, 5-fluorouracil and doxorubicin produce partial remission in approximately 40% of cases. Pancreatic neuroendocrine tumours show a very high degree of vascularization as well as abundant production and secretion of growth factors. There is preliminary evidence of benefit from antiangiogenesis therapy utilizing vascular endothelial growth factor (VEGF) antagonists. Antagonists to a number of oncogenic growth factors are currently under investigation.

In patients with extensive liver metastasis, occlusion of the arterial blood flow by hepatic arterial embolization may control hormone-related symptoms. In most cases the tumours are slowly progressive and may allow a reasonable quality of life for many years.

Clinical syndromes

Insulinoma is described on page 1059.

A *gastrinoma* accounts for approximately 1 in 1000 cases of duodenal ulcer disease. This results from hypersecretion of gastric acid secondary to ectopic gastrin secretion within the endocrine pancreas (Zollinger–Ellison syndrome). Recurrent severe duodenal ulceration occurs with only a partial response to acid suppression. The diagnosis is confirmed by an elevated gastrin level. High-dose proton pump inhibitors are used to suppress symptoms.

A *VIPoma* is an endocrine pancreatic tumour producing vasoactive intestinal polypeptide (VIP). This causes a severe secretory diarrhoea secondary to the stimulation of adenyl cyclase within the enterocyte (Verner–Morrison syndrome). The clinical syndrome is one of profuse watery diarrhoea, hypokalaemia and a metabolic acidosis. To produce the syndrome, the tumours are usually in excess of 3 cm in diameter. The clinical syndrome can be controlled by glucocorticoids, long-acting octreotide or lanreotide.

Glucagonomas are rare alpha-cell tumours which are responsible for the syndrome of migratory necrolytic dermatitis, weight loss, diabetes mellitus, deep vein thrombosis, anaemia and hypoalbuminaemia. The diagnosis is made by measuring pancreatic glucagon in the serum. Metastases are common at presentation but if the tumour is localized, pancreatic resection may be curative.

Somatostatinomas are rare malignant D cell tumours of the pancreas and 30% occur in the duodenum and small bowel. These tumours cause diabetes mellitus, gallstones and diarrhoea/steatorrhoea. They can be diagnosed by high serum somatostatin levels, CT/MR scans or octreotide scintiscans. Treatment is by resection, where possible, and octreotide therapy.

Non-functioning neuroendocrine tumours usually present by the local mass effect with pain, weight loss and the occasional bile duct obstruction. At the time of presentation the tumours are frequently large with distant metastases. Curative resection is possible in only a small proportion. Palliative surgery may be indicated to alleviate mass-related symptoms.

FURTHER READING

Frossard JL, Steer ML, Pastor CM. Acute pancreatitis. *Lancet* 2008; **371**:143–152.

Koorstra J-B, Hustinx SR, Offerhaus GJA et al. Pancreatic carcinogenesis. *Pancreatology* 2008; **8**: 110–125.

SIGNIFICANT WEBSITES

http://www.gastrohep.com
Resources for gastroenterology, hepatology and endoscopy

http://www.aasld.org (click on guidelines for management)
Viral hepatitis

http://www.emedicine.com/emerg/topic98.htm
Cholecystitis, cholelithiasis

This page intentionally left blank

8 Haematological disease

Introduction and general aspects 387
The red cell 389
Anaemia 392
Microcytic anaemia 393
Normocytic anaemia 397
Macrocytic anaemias 397
Megaloblastic anaemia 397
Macrocytosis without megaloblastic changes 401
Anaemia due to marrow failure (aplastic anaemia) 402
Haemolytic anaemias: an introduction 403
Inherited haemolytic anaemia 404
Red cell membrane defects 404
Haemoglobin abnormalities 406
The thalassaemias 407
Sickle syndromes 409
Metabolic disorders of the red cell 412
Acquired haemolytic anaemia 414
Autoimmune haemolytic anaemias 415
Drug-induced immune haemolytic anaemia 417
Alloimmune haemolytic anaemia 417
Non-immune haemolytic anaemia 418
Mechanical haemolytic anaemia 419
Myeloproliferative disorders 419
Polycythaemia 419
Myelofibrosis (myelosclerosis) 421
Myelodysplasia (MDS) 422
The spleen 422
Splenomegaly 423
Blood transfusion 423
Blood groups 424
Procedure for blood transfusion 424
Complications of blood transfusion 425
Blood, blood components and blood products 428
The white cell 430
Neutrophils 430
Eosinophils 431
Basophils 431
Monocytes 431
Lymphocytes 431
Haemostasis and thrombosis 431
Haemostasis 431
Vascular disorders 435
Platelet disorders 435
Inherited coagulation disorders 438
Acquired coagulation disorders 440
Thrombosis 442

INTRODUCTION AND GENERAL ASPECTS

Blood consists of:

- red cells
- white cells
- platelets
- plasma, in which the above elements are suspended.

Plasma is the liquid component of blood, which contains soluble fibrinogen. Serum is what remains after the formation of the fibrin clot.

The formation of blood cells (haemopoiesis)

The haemopoietic system includes the bone marrow, liver, spleen, lymph nodes and thymus. There is an enormous turnover of cells with the red cells surviving 120 days, platelets around 7 days but granulocytes only 7 hours. The production of as many as 10^{13} new myeloid cells (all blood cells except for lymphocytes) per day in the normal healthy state obviously requires to be tightly regulated according to the needs of the body.

Blood islands are formed in the yolk sac in the third week of gestation and produce primitive blood cells, which migrate to the liver and spleen. These organs are the chief sites of haemopoiesis from 6 weeks to 7 months, when the bone marrow becomes the main source of blood cells. However in childhood and adult life, the bone marrow is the only source of blood cells in a normal person.

At birth, haemopoiesis is present in the marrow of nearly every bone. As the child grows the active red marrow is gradually replaced by fat (yellow marrow) so that haemopoiesis in the adult becomes confined to the central skeleton and the proximal ends (trabecular area) of the long bones. Only if the demand for blood cells increases and persists do the areas of red marrow extend. Pathological processes interfering with normal haemopoiesis may result in resumption of haemopoietic activity in the liver and spleen, which is referred to as *extramedullary haemopoiesis*.

All blood cells are derived from pluripotent stem cells. These stem cells are supported by stromal cells (see below), which also influence haemopoiesis. The stem cell has two

properties – the first is *self-renewal*, i.e. the production of more stem cells, and the second is its proliferation and differentiation into progenitor cells, committed to one specific cell line.

There are two major ancestral cell lines derived from the pluripotential stem cell: lymphocytic and myeloid (non-lymphocytic) cells (Fig. 8.1). The former gives rise to T and B cells. The myeloid stem cell gives rise to the progenitor CFU-GEMM (colony-forming unit, granulocyte–erythrocyte–monocyte–megakaryocyte). The progenitor cells such as CFU-GEMM cannot be recognized in bone marrow biopsies but are recognized by their ability to form colonies when haemopoietic cells are immobilized in a soft gel matrix. The CFU-GEMM can go on to form CFU-*GM*, CFU-Eo, and CFU-*Meg*, each of which can produce a particular cell type (e.g. neutrophils, eosinophils and platelets) under appropriate growth conditions. Haemopoiesis is under the control of growth factors and inhibitors, and the microenvironment of the bone marrow also plays a role in its regulation.

Haemopoietic growth factors

Haemopoietic growth factors are glycoproteins, which regulate the differentiation and proliferation of haemopoietic progenitor cells and the function of mature blood cells. They act on the cytokine-receptor superfamily expressed on haemopoietic cells at various stages of development to maintain the haemopoietic progenitor cells and to stimulate increased production of one or more cell lines in response to stresses such as blood loss and infection (Fig. 8.1).

These haemopoietic growth factors including erythropoietin, interleukin 3 (IL-3), IL-6, -7, -11 -12, β-catenin, stem cell factor (SCF, Steel factor or C-kit ligand) and Fms-tyrosine kinase 3 (Flt3) act via their specific receptor on cell surfaces to stimulate the cytoplasmic janus kinase (JAK) (see p. 24). This major signal transducer activates tyrosine kinase causing gene activation in the cell nucleus. Colony-stimulating factors (CSFs, the prefix indicating the cell type, see Fig. 8.1), as well

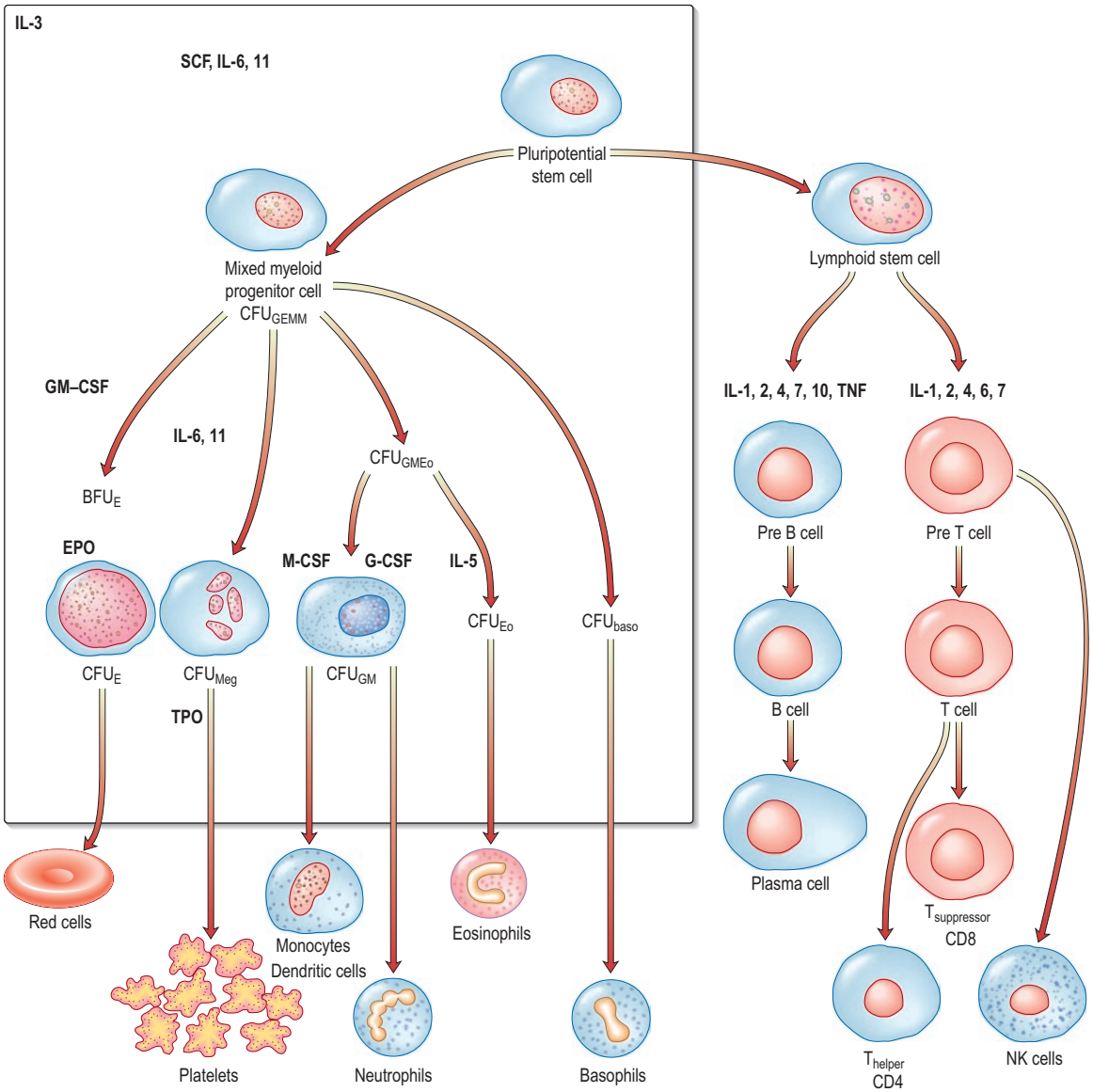


Fig. 8.1 **Role of growth factors in normal haemopoiesis.** Some of the multiple growth factors acting on stem cells and early progenitor cells are shown. baso, basophil; BFU, burst-forming unit; CFU, colony-forming unit; CSF, colony-stimulating factor; E, erythroid; Eo, eosinophil; EPO, erythropoietin; G, granulocyte; GEMM, mixed granulocyte, erythroid, monocyte, megakaryocyte; GM, granulocyte, monocyte; IL, interleukin; M, monocyte; Meg, megakaryocyte; SCF, stem cell (Steel) factor or C-kit ligand; TNF, tumour necrosis factor; TPO, thrombopoietin.

as interleukins and erythropoietin (EPO) regulate the lineage committed progenitor cells.

Thrombopoietin (TPO, which, like erythropoietin, is produced in the kidneys and the liver) controls platelet production, along with IL-6 and IL-11. In addition to these factors stimulating haemopoiesis, other factors inhibit the process and include tumour necrosis factor (TNF) and transforming growth factor β (TGF-β). Many of the growth factors are produced by activated T cells, monocytes and bone marrow stromal cells such as fibroblasts, endothelial cells and macrophages; these cells are also involved in inflammatory responses. Bone marrow stem cells can differentiate into other organ cell types, e.g. heart, liver, nerves, and this is called *stem cell plasticity.*

Uses in treatment

Many growth factors have been produced by recombinant DNA techniques and are being used clinically. Examples include granulocyte-colony-stimulating factor (G-CSF), which is used to accelerate haemopoietic recovery after chemotherapy and haemopoietic cell transplantation, and erythropoietin, which is used to treat anaemia in patients with chronic renal failure. Thrombopoietin is undergoing clinical trials in patients with immune thrombocytopenic purpura.

Stem cell diseases

The clonal proliferation of stem cells leads to diseases, e.g. leukaemia (see p. 467), polycythaemia vera (see p. 419), myelofibrosis (see p. 421), paroxysmal nocturnal haemoglobinuria (see p. 418). Failure of stem cell growth leads to aplastic anaemia (see p. 402). Viral infections, e.g. hepatitis, and chemotherapy damage stem cells leading to marrow failure.

Peripheral blood

Automated cell counters are used to measure the level of haemoglobin (Hb) and the number and size of red cells, white cells and platelets (Table 8.1). Other indices can be derived from these values. A carefully evaluated blood film is still an essential adjunct to the above, as definitive abnormalities of cells can be seen.

Table 8.1 Normal values for peripheral blood

	Male	Female
Hb (g/dL)	13.5–17.5	11.5–16
PCV (haematocrit; L/L)	0.4–0.54	0.37–0.47
RCC (10^{12}/L)	4.5–6.0	3.9–5.0
MCV (fL)	80–96	
MCH (pg)	27–32	
MCHC (g/dL)	32–36	
RDW (%)	11–15	
WBC (10^9/L)	4.0–11.0	
Platelets (10^9/L)	150–400	
ESR (mm/h)	< 20	
Reticulocytes	0.5–2.5% (50–100 × 10^9/L)	

ESR, erythrocyte sedimentation rate; Hb, haemoglobin; MCH, mean corpuscular haemoglobin; MCHC, mean corpuscular haemoglobin concentration; MCV, mean corpuscular volume of red cells; PCV, packed cell volume; RCC, red cell count; RDW, red blood cell distribution width; WBC, white blood count.

- *The mean corpuscular volume (MCV)* of red cells is the most useful of the indices and is used to classify anaemia (see p. 392).
- The *red cell distribution width (RDW)* is calculated by dividing the *standard deviation* of the red cell width by the *mean cell width* × 100. An elevated RDW suggests variation in red cell size, i.e. anisocytosis, and this is seen in iron deficiency. In β thalassaemia trait, the RDW is usually normal.
- *The white cell count* (WCC, or WBC, white blood count) gives the total number of circulating leucocytes, and many automated cell counters produce differential counts as well.
- *Reticulocytes* normally are less than 2% of the red cells (see p. 390). The reticulocyte count gives a guide to the erythroid activity in the bone marrow. An increased count is seen with increased marrow maturity, e.g. following haemorrhage or haemolysis, and during the response to treatment with a specific haematinic. A low count in the presence of anaemia indicates an inappropriate response by the bone marrow and may be seen in bone marrow failure (from whatever cause) or where there is a deficiency of a haematinic.
- *Erythrocyte sedimentation rate (ESR)* is the rate of fall of red cells in a column of blood and is a measure of the acute-phase response. The pathological process may be immunological, infective, ischaemic, malignant or traumatic. A raised ESR reflects an increase in the plasma concentration of large proteins, such as fibrinogen and immunoglobulins. These proteins cause rouleaux formation, with red cells clumping together and therefore falling more rapidly. The ESR increases with age, and is higher in females than in males. It is low in polycythaemia vera, owing to the high red cell concentration, and increased in patients with severe anaemia.
- *Plasma viscosity* is a measurement used instead of the ESR in some laboratories. It is also dependent on the concentration of large molecules such as fibrinogen and immunoglobulins. It is not affected by the level of Hb.
- *C-reactive protein (CRP)* is a pentraxin, one of the proteins produced in the acute-phase response. It is synthesized exclusively in the liver and rises within 6 hours of an acute event. It rises with fever (possibly triggered by IL-1, IL-6 and TNF-α and other cytokines), in inflammatory conditions and after trauma. It follows the clinical state of the patient much more rapidly than the ESR and is unaffected by the level of Hb, but it is less helpful than the ESR or plasma viscosity in monitoring chronic inflammatory diseases. The measurement of CRP is easy and quick to perform using an immunoassay that can be automated. High-sensitivity assays have shown that increased levels may predict future cardiovascular disease (see p. 747).

THE RED CELL

Erythropoiesis

Red cell precursors pass through several stages in the bone marrow. The earliest morphologically recognizable cells are pronormoblasts. Smaller normoblasts result from cell

divisions, and precursors at each stage progressively contain less RNA and more Hb in the cytoplasm. The nucleus becomes more condensed and is eventually lost from the late normoblast in the bone marrow, when the cell becomes a reticulocyte.

- *Reticulocytes* contain residual ribosomal RNA and are still able to synthesize Hb. They remain in the marrow for about 1–2 days and are released into the circulation, where they lose their RNA and become mature red cells (erythrocytes) after another 1–2 days. Mature red cells are non-nucleated biconcave discs.
- Nucleated red cells (*normoblasts*) are not normally present in peripheral blood, but are present if there is extramedullary haemopoiesis and in some marrow disorders (see leucoeryothroblastic anaemia, p. 430).
- About 10% of *erythroblasts* die in the bone marrow even during normal erythropoiesis. Such ineffective erythropoiesis is substantially increased in some anaemias such as thalassaemia major and megaloblastic anaemia.
- *Erythropoietin* is a hormone which controls erythropoiesis. The gene for erythropoietin is on chromosome 7 and codes for a heavily glycosylated polypeptide of 165 amino acids. Erythropoietin has a molecular weight of 30 400 and is produced in the peritubular cells in the kidneys (90%) and in the liver (10%). Its production is regulated mainly by tissue oxygen tension. Production is increased if there is hypoxia from whatever cause – for example, anaemia, or cardiac or pulmonary disease. The erythropoietin gene is one of a number of genes that is regulated by the hypoxic sensor pathway. The 3′-flanking region of the erythropoietin gene has a hypoxic response element, which is necessary for the induction of transcription of the gene in hypoxic cells. Hypoxia-inducible factor 1 (HIF-1) is a transcription factor, which binds to the hypoxia response element and acts as a master regulator of several genes that are responsive to hypoxia. Erythropoietin stimulates an increase in the proportion of bone marrow precursor cells committed to erythropoiesis, and CFU-E are stimulated to proliferate and differentiate. Increased ‘inappropriate’ production of erythropoietin is also seen in patients with renal disease and neoplasms in other sites, which result in polycythaemia (see Table 8.16).

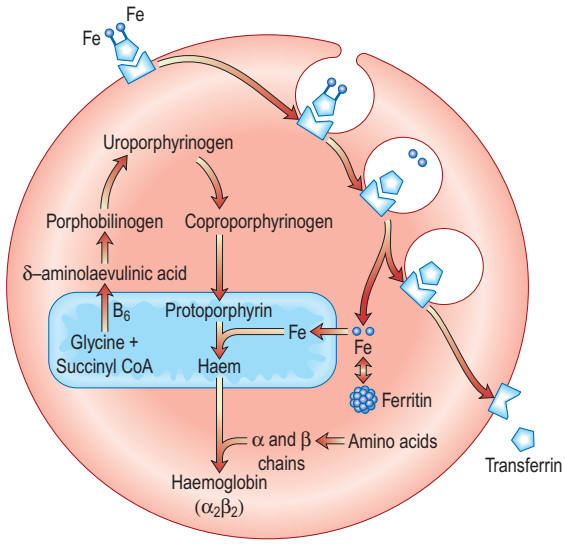


Fig. 8.2 **Haemoglobin synthesis.** Transferrin attaches to a surface receptor on developing red cells. Iron is released and transported to the mitochondria, where it combines with protoporphyrin to form haem. Protoporphyrin itself is manufactured from glycine and succinyl-CoA. Haem combines with α and β chains (formed on ribosomes) to make haemoglobin.

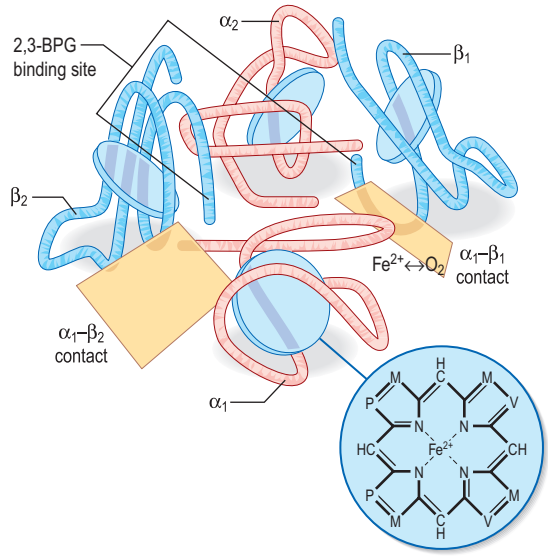


Fig. 8.3 **Model of the haemoglobin molecule showing α (pink) and β (blue) chains.** 2,3-BPG (bisphosphoglycerate) or diphosphoglycoate (DGP) binds in the centre of the molecule and stabilizes the deoxygenated form by cross-linking the β chains (also see Fig. 8.4). M, methyl; P, propionic acid; V, vinyl.

Haemoglobin synthesis

Haemoglobin performs the main functions of red cells – carrying O_2 to the tissues and returning CO_2 from the tissues to the lungs. Each normal adult Hb molecule (Hb A) has a molecular weight of 68 000 and consists of two α and two β globin polypeptide chains ($\alpha_2\beta_2$) which have 141 and 146 amino acids, respectively. HbA comprises about 97% of the Hb in adults. Two other types, Hb A_2 ($\alpha_2\delta_2$) and Hb F ($\alpha_2\gamma_2$), are found in adults in small amounts (1.5–3.2% and <1%, respectively) (see p. 406).

Haemoglobin synthesis occurs in the mitochondria of the developing red cell (Fig. 8.2). The major rate-limiting step is the conversion of glycine and succinic acid to δ-aminolaevulinic acid (ALA) by ALA synthase. Vitamin B_6 is a coenzyme for this reaction, which is inhibited by haem and stimulated by erythropoietin. Two molecules of δ-ALA condense to form a pyrrole ring (porphobilinogen). These rings are then grouped in fours to produce protoporphyrins. Finally, iron is inserted to form haem. Haem is then inserted into the globin chains to form Hb. The structure of Hb is shown in Figure 8.3.

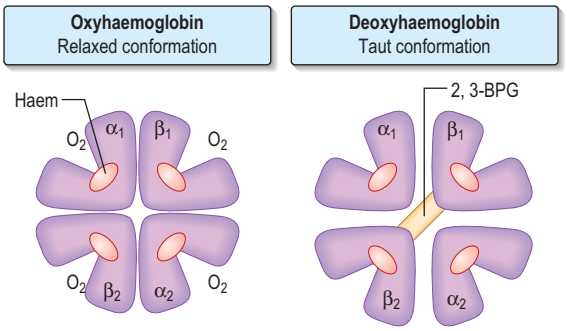


Fig. 8.4 **Oxygenated and deoxygenated haemoglobin molecule.** The haemoglobin molecule is predominantly stabilized by α–β chain bonds rather than α–α and β–β chain bonds. The structure of the molecule changes during O_2 uptake and release. When O_2 is released, the β chains rotate on the $\alpha_1\beta_2$ and $\alpha_2\beta_1$ contacts, allowing the entry of 2,3-BPG which causes a lower affinity of haemoglobin for O_2 and improved delivery of O_2 to the tissues. From Hoffbrand AV, Pettit JE. *Essential Haematology*, 3rd edn. Oxford: Blackwell Scientific Publications, 1993, with permission.

Haemoglobin function

The biconcave shape of red cells provides a large surface area for the uptake and release of oxygen and carbon dioxide. Haemoglobin becomes saturated with oxygen in the pulmonary capillaries where the partial pressure of oxygen is high and Hb has a high affinity for oxygen. Oxygen is released in the tissues where the partial pressure of oxygen is low and Hb has a low affinity for oxygen.

In adult haemoglobin (Hb A), a haem group is bound to each of the four globin chains; the haem group has a porphyrin ring with a ferrous atom which can reversibly bind one oxygen molecule. The haemoglobin molecule exists in two conformations, R and T. The T (taut) conformation of deoxyhaemoglobin is characterized by the globin units being held tightly together by electrostatic bonds (Fig. 8.4). These bonds are broken when oxygen binds to haemoglobin, resulting in the R (relaxed) conformation in which the remaining oxygen-binding sites are more exposed and have a much higher affinity for oxygen than in the T conformation. The binding of one oxygen molecule to deoxyhaemoglobin increases the oxygen affinity of the remaining binding sites – this property is known as 'cooperativity' and is the reason for the sigmoid shape of the oxygen dissociation curve. Haemoglobin is, therefore, an example of an allosteric protein. The binding of oxygen can be influenced by secondary effectors – hydrogen ions, carbon dioxide and red-cell 2,3-bisphosphoglycerate (2,3-BPG). Hydrogen ions and carbon dioxide added to blood cause a reduction in the oxygen-binding affinity of haemoglobin (the Bohr effect). Oxygenation of haemoglobin reduces its affinity for carbon dioxide (the Haldane effect). These effects help the exchange of carbon dioxide and oxygen in the tissues.

Red cell metabolism produces 2,3-BPG from glycolysis. 2,3-BPG accumulates because it is sequestered by binding to deoxyhaemoglobin. The binding of 2,3-BPG stabilizes the T conformation and reduces its affinity for oxygen. The P_{50} is the partial pressure of oxygen at which the haemoglobin is 50% saturated with oxygen. P_{50} increases with 2,3-BPG concentrations, which increase when oxygen availability is reduced in conditions such as hypoxia or anaemia. P_{50} also rises with increasing body temperature, which may be beneficial during prolonged exercise. Haemoglobin regulates oxygen transport as shown in the oxyhaemoglobin dissociation curve. When the primary limitation to oxygen transport is in the periphery, e.g. heavy exercise, anaemia, the P_{50} is increased to enhance oxygen unloading. When the primary limitation is in the lungs, e.g. lung disease, high altitude exposure, the P_{50} is reduced to enhance oxygen loading.

A summary of normal red cell production and destruction is given in Figure 8.5.

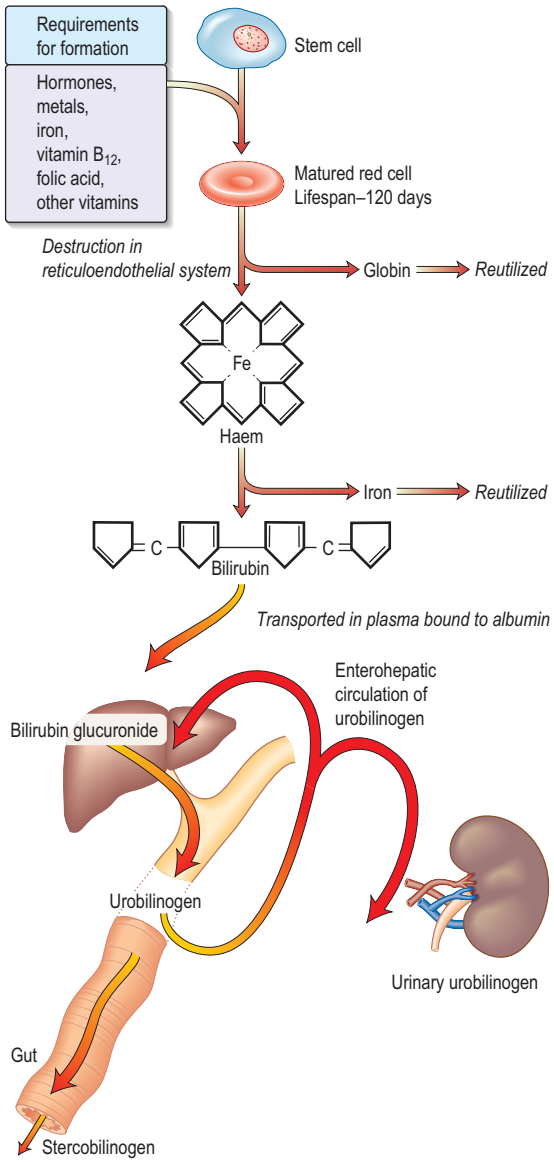


Fig. 8.5 **Red cell production and breakdown** (see p. 322).

Ask the authors
www.studentconsult.com

FURTHER READING

Kaushansky K. Lineage-specific hematopoietic growth factors. *New England Journal of Medicine* 2006**; 354:** 2034–2045.

Macdougall IC, Eckardt K-U. Novel strategies for stimulating erythropoiesis and potential new treatments for anaemia. *Lancet* 2006**; 368:** 947–953.

ANAEMIA

Anaemia is present when there is a decrease in the level of haemoglobin in the blood below the reference level for the age and sex of the individual (Table 8.1). Alterations in the level of Hb may occur as a result of changes in the plasma volume, as shown in Figure 8.6. A reduction in the plasma volume will lead to a spuriously high Hb – this is seen with dehydration and in the clinical condition of apparent polycythaemia (see p. 421). A raised plasma volume produces a spurious anaemia, even when combined with a small increase in red cell volume as occurs in pregnancy.

The various types of anaemia, classified by MCV, are shown in Figure 8.7. There are three major types:

- hypochromic microcytic with a low MCV
- normochromic normocytic with a normal MCV
- macrocytic with a high MCV.

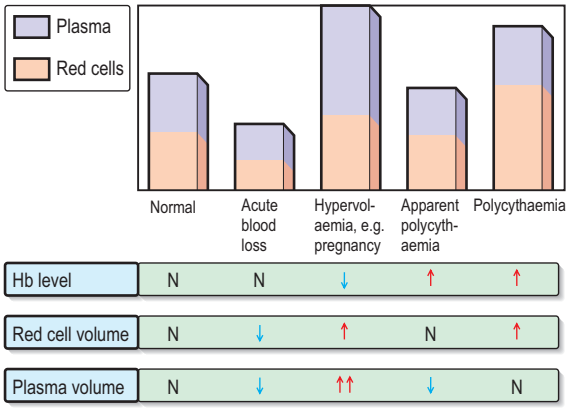


Fig. 8.6 **Alterations of haemoglobin in relation to plasma.**

Clinical features

Patients with anaemia may be asymptomatic. A slowly falling level of Hb allows for haemodynamic compensation and enhancement of the oxygen-carrying capacity of the blood. A rise in 2,3-BPG causes a shift of the oxygen dissociation curve to the right, so that oxygen is more readily given up to the tissues. Where blood loss is rapid, more severe symptoms will occur, particularly in elderly people.

Symptoms (all non-specific)

- Fatigue, headaches and faintness are all very common in the general population
- Breathlessness
- Angina
- Intermittent claudication
- Palpitations.

A common problem is that anaemia exacerbates cardiorespiratory problems especially in the elderly. For example, angina or intermittent claudication may be precipitated by anaemia. A good way to assess the effects of anaemia is to ask about breathlessness in relation to different levels of exercise (e.g. walking on the flat or climbing one flight of stairs).

Signs

- Pallor
- Tachycardia
- Systolic flow murmur
- Cardiac failure.

Specific signs seen in the different types of anaemia will be discussed in the appropriate sections. Examples include:

- koilonychia – spoon-shaped nails seen in iron deficiency anaemia
- jaundice – found in haemolytic anaemia
- bone deformities – found in thalassaemia major

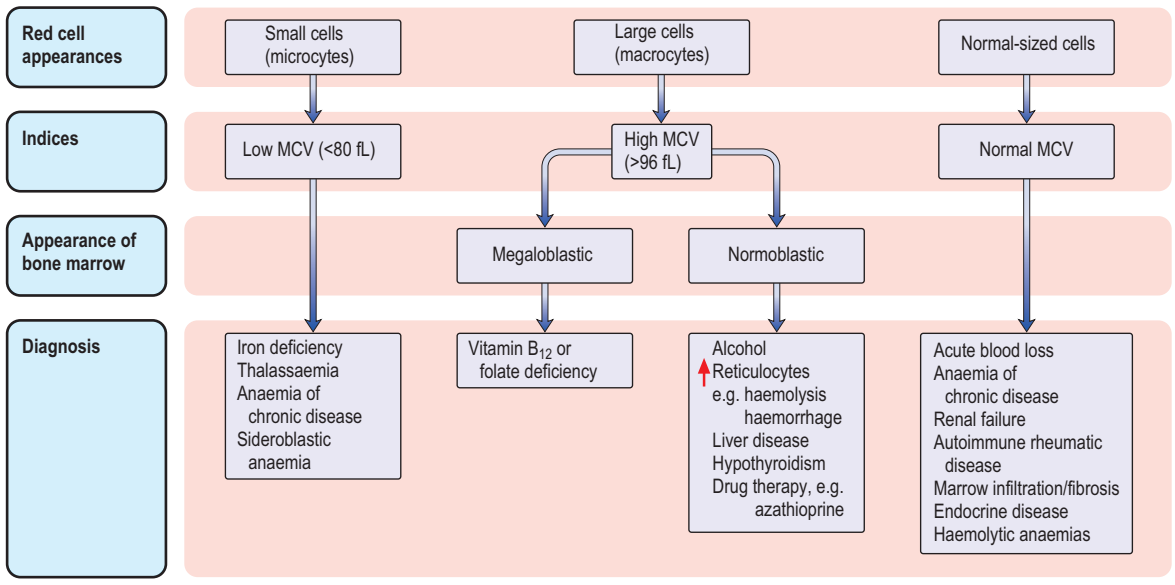


Fig. 8.7 **Classification of anaemia.** MCV, mean corpuscular volume.

- leg ulcers – occur in association with sickle cell disease.

It must be emphasized that anaemia is not a diagnosis, and a cause must be found.

Investigations

Peripheral blood

A low haemoglobin should always be considered in relation to:

- the white blood cell (WBC) count
- the platelet count
- the reticulocyte count (as this indicates marrow activity)
- the blood film, as abnormal red cell morphology (Fig. 8.9) may indicate the diagnosis. Where two populations of red cells are seen, the blood film is said to be dimorphic. This may, for example, be seen in patients with 'double deficiencies' (e.g. combined iron and folate deficiency in coeliac disease, or following treatment of anaemic patients with the appropriate haematinic).

Bone marrow

Examination of the bone marrow is performed to further investigate abnormalities found in the peripheral blood (Practical Box 8.1). Aspiration provides a film which can be examined by microscopy for the morphology of the developing haemopoietic cells. The trephine provides a core of bone which is processed as a histological specimen and allows an overall view of the bone marrow architecture, cellularity and presence/absence of abnormal infiltrates.

The following are assessed:

- cellularity of the marrow
- type of erythropoiesis (e.g. normoblastic or megaloblastic)
- cellularity of the various cell lines
- infiltration of the marrow
- iron stores.

Special tests may be performed for further diagnosis: cytogenetic, immunological, cytochemical markers, biochemical analyses (e.g. deoxyuridine suppression test), microbiological culture.

Practical Box 8.1 Techniques for obtaining bone marrow

The technique should be explained to the patient and consent obtained

Aspiration

Site – usually iliac crest
Give local anaesthetic injection
Use special bone marrow needle (e.g. Salah)
Aspirate marrow
Make smear with glass slide
Stain with:
  Romanowsky technique
  Perls' reaction (acid ferrocyanide) for iron

Trephine

Indications include:
  'Dry tap' obtained with aspiration
  Better assessment of cellularity, e.g. aplastic anaemia
  Better assessment of presence of infiltration or fibrosis

Technique

Site – usually posterior iliac crest
Give local anaesthetic injection
Use special needle (e.g. Jamshidi – longer and wider than for aspiration)
Obtain core of bone
Fix in formalin; decalcify – this takes a few days
Stain with:
  Haematoxylin and eosin
  Reticulin stain

MICROCYTIC ANAEMIA

Iron deficiency is the most common cause of anaemia in the world, affecting 30% of the world's population equivalent to 500 million people. This is because of the body's limited ability to absorb iron and the frequent loss of iron owing to haemorrhage. Although iron is abundant, most is in the insoluble ferric (Fe^{3+}) form, which has poor bioavailability. Ferrous (Fe^{2+}) is more readily absorbed. Free iron is toxic, and it is bound to various proteins for transport and storage.

The other causes of a microcytic hypochromic anaemia are anaemia of chronic disease, sideroblastic anaemia and thalassaemia. In thalassaemia (see p. 407) there is a defect in globin synthesis, in contrast to the other three causes of microcytic anaemia where the defect is in the synthesis of haem.

Iron

Dietary intake

The average daily diet in the UK contains 15–20 mg of iron, although normally only 10% of this is absorbed. Absorption may be increased to 20–30% in iron deficiency and pregnancy.

Non-haem iron is mainly derived from cereals, which are commonly fortified with iron; it forms the main part of dietary iron. Haem iron is derived from haemoglobin and myoglobin in red or organ meats. Haem iron is better absorbed than non-haem iron, whose availability is more affected by other dietary constituents.

Absorption (Fig. 8.8a,b)

Factors influencing iron and haem iron absorption are shown in Table 8.2

Dietary haem iron is more rapidly absorbed than non-haem iron derived from vegetables and grain. Most haem is absorbed in the proximal intestine, with absorptive capacity decreasing distally. The intestinal haem transporter HCP1 (haem carrier protein 1) has been identified and found to be highly expressed in the duodenum. It is upregulated by hypoxia and iron deficiency. Some haem iron may be reabsorbed intact into circulation via the cell by two exporter proteins – BCRP (breast cancer resistant protein) and FLVCR (feline leukaemia virus subgroup C).

Non-haem iron absorption occurs primarily in the duodenum. Non-haem iron is dissolved in the low pH of the stomach and reduced from the ferric to the ferrous form by a brush border ferrireductase. Cells in duodenal crypts are able to sense the body's iron requirements and retain this

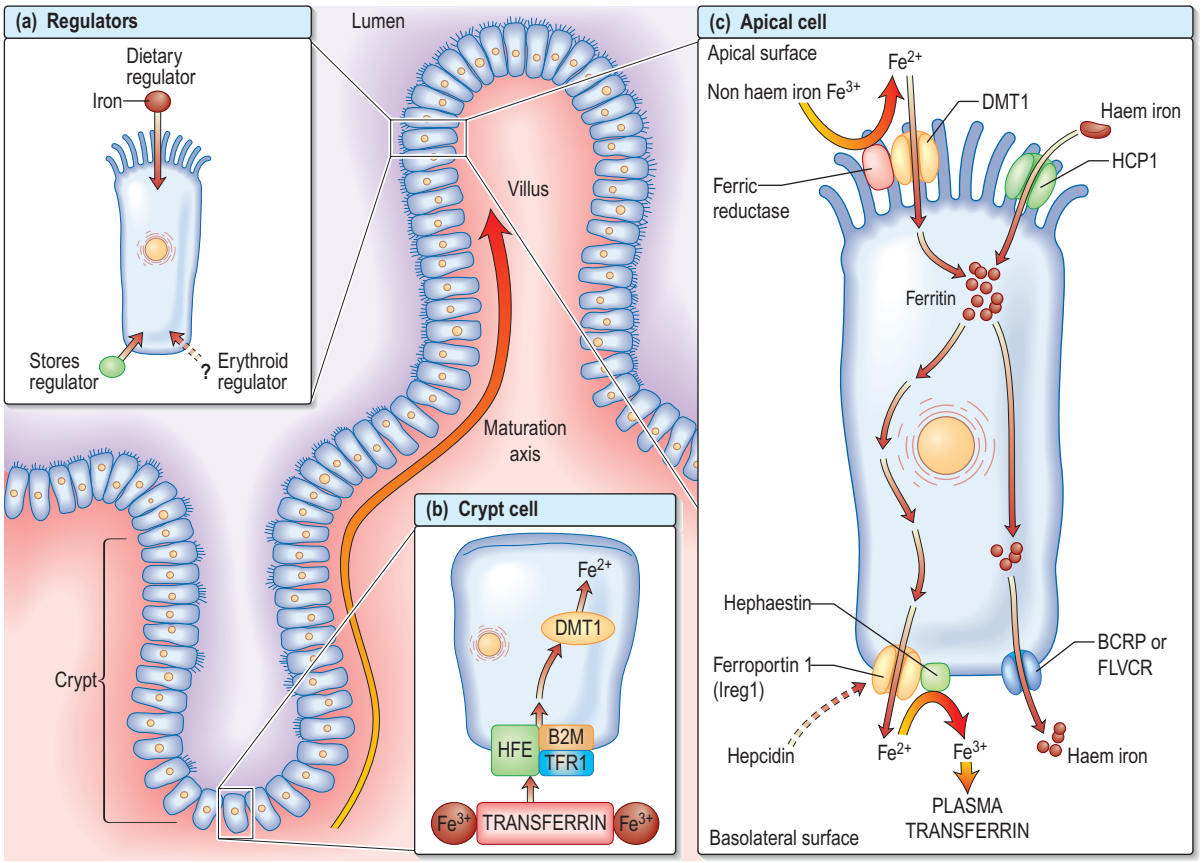


Fig. 8.8 **(a) Regulation of the absorption of intestinal iron.** The iron-absorbing cells of the duodenal epithelium originate in the intestinal crypts and migrate toward the tip of the villus as they differentiate (maturation axis). Absorption of intestinal iron is regulated by at least three independent mechanisms. First, iron absorption is influenced by recent dietary iron intake (dietary regulator). After a large dietary bolus, absorptive cells are resistant to iron uptake for several days. Second, iron absorption can be modulated considerably in response to body iron stores (stores regulator). Third, a signal communicates the state of bone marrow erythropoiesis to the intestine (erythroid regulator).
(b) Duodenal crypt cells sense body iron status through the binding of transferrin to the HFE/B_2M/TFR1 complex. Cytosolic enzymes change the oxidative state of iron from ferric (Fe^{3+}) to ferrous (Fe^{2+}). A decrease in crypt cell iron concentration upregulates the divalent metal transporter (DMT1). This increases as crypt cells migrate up the villus and become mature absorptive cells.
(c) Apical cell. Dietary iron is reduced from the ferric to the ferrous state by the brush border ferrireductase. DMT1 facilitates iron absorption from the intestinal lumen. The export proteins, e.g. ferroportin 1 (Ireg1) and hephaestin, transfer iron from the enterocyte into the circulation. BCRP, breast cancer resistant protein; B2M, β_2-microglobulin; FLVCR, feline leukaemia virus subgroup C; HCP-1, Haem carrier protein-1; HFE, hereditary haemochromatosis gene.

Table 8.2 Factors influencing iron absorption

Haem iron is absorbed better than non-haem iron
Ferrous iron is absorbed better than ferric iron
Gastric acidity helps to keep iron in the ferrous state and soluble in the upper gut
Formation of insoluble complexes with phytate or phosphate decreases iron absorption
Iron absorption is increased with low iron stores and increased erythropoietic activity, e.g. bleeding, haemolysis, high altitude
There is a decreased absorption in iron overload, except in hereditary haemochromatosis, where it is increased

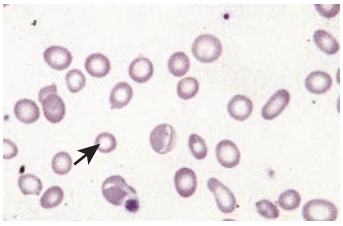


Fig. 8.9 **Hypochromic microcytic cells (arrow) on a blood film.** Poikilocytosis and anisocytosis are seen.

information as they mature into cells capable of absorbing iron at the tips of the villi. A protein, divalent metal transporter 1 (DMT1) or natural resistance-associated macrophage protein (NRAMP2), transports iron (and other metals) across the apical (luminal) surface of the mucosal cells in the small intestine. Haem iron is absorbed in a separate less-well-characterized process.

Once inside the mucosal cell, iron may be transferred across the cell to reach the plasma, or be stored as ferritin; the body's iron status at the time the absorptive cell developed from the crypt cell is probably the crucial

deciding factor. Iron stored as ferritin will be lost into the gut lumen when the mucosal cells are shed; this regulates iron balance. The mechanism of transport of iron across the basolateral surface of mucosal cells involves a transporter protein, ferroportin 1 (Ireg1). This transporter protein requires an accessory, multicopper protein, hephaestin (Fig. 8.8c).

The body iron content is closely regulated by the control of iron absorption but there is no physiological mechanism for eliminating excess iron from the body. The key molecule regulating iron absorption is hepcidin, a 25 amino acid peptide synthesized in the liver. Hepcidin acts by regulating the activity of the iron exporting protein ferroportin by binding to ferroportin causing its internalization and degradation, thereby decreasing iron efflux from iron exporting tissues into plasma. Therefore, high levels of hepcidin will destroy ferroportin and limit iron absorption, and vice versa low levels of hepcidin will encourage iron absorption. However, the importance of hepcidin goes further than its role in physiologic iron absorption. For example, most forms of haemochromatosis result from dysregulation of hepcidin or defects of hepcidin or ferroportin themselves.

A longstanding mystery is why anaemias characterized by ineffective erythropoiesis such as thalassaemia are associated with excessive and inappropriate iron absorption. Preliminary evidence again suggests that the increased iron absorption in beta-thalassaemia is mediated by downregulation of hepcidin and upregulation of ferroportin.

Transport in the blood

The normal serum iron level is about 13–32 μmol/L; there is a diurnal rhythm with higher levels in the morning. Iron is transported in the plasma bound to transferrin, a β-globulin that is synthesized in the liver. Each transferrin molecule binds two atoms of ferric iron and is normally one-third saturated. Most of the iron bound to transferrin comes from macrophages in the reticuloendothelial system and not from iron absorbed by the intestine. Transferrin-bound iron becomes attached by specific receptors to erythroblasts and reticulocytes in the marrow and the iron is removed (Fig. 8.2).

In an average adult male, 20 mg of iron, chiefly obtained from red cell breakdown in the macrophages of the reticuloendothelial system, is incorporated into Hb every day.

Iron stores

About two-thirds of the total body iron is in the circulation as haemoglobin (2.5–3 g in a normal adult man). Iron is stored in reticuloendothelial cells, hepatocytes and skeletal muscle cells (500–1500 mg). About two-thirds of this is stored as ferritin and one-third as haemosiderin in normal individuals. Small amounts of iron are also found in plasma (about 4 mg bound to transferrin), with some in myoglobin and enzymes.

Ferritin is a water-soluble complex of iron and protein. It is more easily mobilized than haemosiderin for Hb formation. It is present in small amounts in plasma.

Haemosiderin is an insoluble iron–protein complex found in macrophages in the bone marrow, liver and spleen. Unlike ferritin, it is visible by light microscopy in tissue sections and bone marrow films after staining by Perls' reaction.

Requirements

Each day 0.5–1.0 mg of iron is lost in the faeces, urine and sweat. Menstruating women lose 30–40 mL of blood per month, an average of about 0.5–0.7 mg of iron per day. Blood loss through menstruation in excess of 100 mL will usually result in iron deficiency as increased iron absorption from the gut cannot compensate for such losses of iron. The demand for iron also increases during growth (about 0.6 mg per day) and pregnancy (1–2 mg per day). In the normal adult the iron content of the body remains relatively fixed. Increases in the body iron content (haemochromatosis) are classified into:

- Hereditary haemochromatosis (see p. 356), where a mutation in the HFE gene causes increased iron absorption.
- Secondary haemochromatosis (transfusion siderosis; see p. 408). This is due to iron overload in conditions where repeated transfusion is the only therapy.

Iron deficiency

Iron deficiency anaemia develops when there is inadequate iron for haemoglobin synthesis. A normal level of Hb is maintained for as long as possible after the iron stores are depleted; latent iron deficiency is said to be present during this period.

- blood loss
- increased demands such as growth and pregnancy
- decreased absorption (e.g. postgastrectomy)
- poor intake.

Most iron deficiency is due to blood loss, usually from the uterus or gastrointestinal tract. Premenopausal women are in a state of precarious iron balance owing to menstruation. Iron deficiency affects more than a quarter of the world's population, but isolated nutritional iron deficiency is rare in developed countries. The most common cause of iron deficiency world-wide is blood loss from the gastrointestinal tract resulting from hookworm infestation. The poor quality of the diet, predominantly containing vegetables, also contributes to the high prevalence of iron deficiency in developing countries. Even in developed countries, iron deficiency is not uncommon in infancy where iron intake is insufficient for the demands of growth. It is more prevalent in infants born prematurely or where the introduction of mixed feeding is delayed.

Clinical features

The symptoms of anaemia are described on page 392. Other clinical features occur as a result of tissue iron deficiency. These are mainly epithelial changes induced by the effect of inadequate iron in the cells:

- brittle nails
- spoon-shaped nails (koilonychia)
- atrophy of the papillae of the tongue
- angular stomatitis
- brittle hair
- a syndrome of dysphagia and glossitis (Plummer–Vinson or Paterson–Brown–Kelly syndrome; see p. 255).

The diagnosis of iron deficiency anaemia relies on a clinical history with questions about dietary intake, regular self-

medication with non-steroidal anti-inflammatory drugs (which may give rise to gastrointestinal bleeding), and the presence of blood in the faeces (which may be a sign of haemorrhoids or carcinoma of the lower bowel). In women, a careful enquiry about the duration of periods, the occurrence of clots and the number of sanitary towels or tampons (normal 3–5/day) used should be made.

Investigations

- ***Blood count and film***. A characteristic blood film is shown in Figure 8.9. The red cells are microcytic (MCV < 80 fL) and hypochromic (MCH (mean corpuscular haemoglobin) < 27 pg). There is poikilocytosis (variation in shape) and anisocytosis (variation in size). Target cells are seen.
- ***Serum iron and iron-binding capacity***. The serum iron falls and the total iron-binding capacity (TIBC) rises in iron deficiency compared with normal. Iron deficiency is regularly present when the transferrin saturation (i.e. serum iron divided by TIBC) falls below 19% (Table 8.3).
- ***Serum ferritin***. The level of serum ferritin reflects the amount of stored iron. The normal values for serum ferritin are 30–300 μg/L (11.6–144 nmol/L) in males and 15–200 μg/L (5.8–96 nmol/L) in females. In simple iron deficiency, a low serum ferritin confirms the diagnosis. However, ferritin is an acute-phase reactant, and levels increase in the presence of inflammatory or malignant diseases. In these cases, measurement of serum iron/TIBC, serum ferritin and soluble transferrin receptors is used.
- ***Serum soluble transferrin receptors***. The number of transferrin receptors increases in iron deficiency. The results of this immunoassay compare well with results from bone marrow aspiration at estimating iron stores. This assay can help to distinguish between iron deficiency and anaemia of chronic disease (Table 8.3), and may avoid the need for bone marrow examination even in complex cases.
- ***Bone marrow***. Erythroid hyperplasia with ragged normoblasts is seen in the marrow in iron deficiency. Staining using Perls' reaction (acid ferrocyanide) does not show the characteristic Prussian-blue granules of stainable iron in the bone marrow fragments or in the erythroblasts. Examination of the bone marrow is not essential for the diagnosis of iron deficiency but it may be helpful in the investigation of complicated cases of anaemia, e.g. to determine if iron deficiency is present in a patient with anaemia of chronic disease.
- ***Other investigations***. These will be indicated by the clinical history and examination. Investigations of the gastrointestinal tract are often required to determine the cause of the iron deficiency (see p. 270).

Differential diagnosis

The presence of anaemia with microcytosis and hypochromia does not necessarily indicate iron deficiency. The most common other causes are thalassaemia, sideroblastic anaemia and anaemia of chronic disease, and in these disorders the iron stores are normal or increased. The differential diagnosis of microcytic anaemia is shown in Table 8.3.

Treatment

Iron deficiency is not a diagnosis per se. The correct management of iron deficiency is to find and treat the underlying cause, and to give iron to correct the anaemia and replace iron stores. The response to iron therapy can be monitored using the reticulocyte count and Hb level, with an expected rise in haemoglobin of 1 g/dL per week.

Oral iron is all that is required in most cases. The best preparation is ferrous sulphate (200 mg three times daily, a total of 180 mg ferrous iron), which is absorbed best when the patient is fasting. If the patient has side-effects such as nausea, diarrhoea or constipation, taking the tablets with food or reducing the dose using a preparation with less iron such as ferrous gluconate (300 mg twice daily, only 70 mg ferrous iron) is all that is usually required to reduce the symptoms. The use of expensive iron compounds, particularly the slow-release ones, which release iron beyond its main sites of absorption, is unnecessary.

In developing countries, distribution of iron tablets and fortification of food are the main approaches for the alleviation of iron deficiency. However, iron supplementation programmes have been ineffective, probably mainly because of poor compliance.

Oral iron should be given for long enough to correct the Hb level and to replenish the iron stores. This can take 6 months. The commonest causes of failure of response to oral iron are:

- lack of compliance
- continuing haemorrhage
- incorrect diagnosis, e.g. thalassaemia trait.

Table 8.3 Microcytic anaemia: the differential diagnosis

	Iron deficiency	Anaemia of chronic disease	Thalassaemia trait (α or β)	Sideroblastic anaemia
MCV	Reduced	Low normal or normal	Very low for degree of anaemia	Low in inherited type but often raised in acquired type
Serum iron	Reduced	Reduced	Normal	Raised
Serum TIBC	Raised	Reduced	Normal	Normal
Serum ferritin	Reduced	Normal or raised	Normal	Raised
Serum soluble transfer receptors	Increased	Normal	Normal or raised	Normal or raised
Iron in marrow	Absent	Present	Present	Present
Iron in erythroblasts	Absent	Absent or reduced	Present	Ring forms

TIBC, total iron binding capacity.

These possibilities should be considered before parenteral iron is used. However, parenteral iron is required by occasional patients, e.g. intolerant to oral preparation, severe malabsorption, chronic disease (e.g. inflammatory bowel disease). Iron stores are replaced much faster with parenteral iron than with oral iron, but the haematological response is no quicker. Parenteral iron can be given as repeated deep intramuscular injections of iron–sorbitol (1.5 mg of iron per kg bodyweight) or by intravenous infusion of low molecular weight iron dextran (test dose required) or iron sucrose.

Anaemia of chronic disease

One of the most common types of anaemia, particularly in hospital patients, is the anaemia of chronic disease, occurring in patients with chronic infections such as tuberculosis or chronic inflammatory disease such as Crohn's disease, rheumatoid arthritis, systemic lupus erythematosus (SLE), polymyalgia rheumatica and malignant disease. There is decreased release of iron from the bone marrow to developing erythroblasts, an inadequate erythropoietin response to the anaemia, and decreased red cell survival.

The exact mechanisms responsible for these effects are not clear but it now seems likely that high levels of hepcidin expression play a key role (see iron absorption).

The serum iron and the TIBC are low, and the serum ferritin is normal or raised because of the inflammatory process. The serum soluble transferrin receptor level is normal (Table 8.3). Stainable iron is present in the bone marrow, but iron is not seen in the developing erythroblasts. Patients do not respond to iron therapy, and treatment is, in general, that of the underlying disorder. Recombinant erythropoietin therapy is used in the anaemia of renal disease (see p. 633), and occasionally in inflammatory disease (rheumatoid arthritis, inflammatory bowel disease).

Sideroblastic anaemia

Sideroblastic anaemias are inherited or acquired disorders characterized by a refractory anaemia, a variable number of hypochromic cells in the peripheral blood, and excess iron and ring sideroblasts in the bone marrow. The presence of ring sideroblasts is the diagnostic feature of sideroblastic anaemia. There is accumulation of iron in the mitochondria of erythroblasts owing to disordered haem synthesis forming a ring of iron granules around the nucleus that can be seen with Perls' reaction. The blood film is often dimorphic; ineffective haem synthesis is responsible for the microcytic hypochromic cells. Sideroblastic anaemias are classified as shown in Table 8.4. A structural defect in δ-aminolaevulinic acid (ALA) synthase, the pyridoxine-dependent enzyme responsible for the first step in haem synthesis (Fig. 8.2), has been identified in one form of inherited sideroblastic anaemia. Primary acquired sideroblastic anaemia is one of the myelodysplastic syndromes (see p. 422). Lead toxicity is described on page 944.

Table 8.4 Classification of sideroblastic anaemia

Inherited
X-linked disease – transmitted by females
Acquired
Myelodysplasia
Myeloproliferative disorders
Myeloid leukaemia
Drugs, e.g. isoniazid
Alcohol abuse
Lead toxicity
Other disorders, e.g. rheumatoid arthritis, carcinoma, megaloblastic and haemolytic anaemias

Treatment

Some patients respond when drugs or alcohol are withdrawn, if these are the causative agents. In occasional cases, there is a response to pyridoxine. Treatment with folic acid may be required to treat accompanying folate deficiency.

NORMOCYTIC ANAEMIA

Normocytic, normochromic anaemia is seen in anaemia of chronic disease, in some endocrine disorders (e.g. hypopituitarism, hypothyroidism and hypoadrenalism) and in some haematological disorders (e.g. aplastic anaemia and some haemolytic anaemias) (Fig. 8.7). In addition, this type of anaemia is seen acutely following blood loss.

MACROCYTIC ANAEMIAS

These can be divided into megaloblastic and non-megaloblastic types, depending on bone marrow findings.

MEGALOBLASTIC ANAEMIA

Megaloblastic anaemia is characterized by the presence in the bone marrow of erythroblasts with delayed nuclear maturation because of defective DNA synthesis (megaloblasts). Megaloblasts are large and have large immature nuclei. The nuclear chromatin is more finely dispersed than normal and has an open stippled appearance (Fig. 8.10). In addition, giant metamyelocytes are frequently seen in megaloblastic anaemia. These cells are about twice the size of normal cells and often have twisted nuclei. Megaloblastic changes occur in:

- vitamin B_{12} deficiency or abnormal vitamin B_{12} metabolism
- folic acid deficiency or abnormal folate metabolism
- other defects of DNA synthesis, such as congenital enzyme deficiencies in DNA synthesis (e.g. orotic aciduria), or resulting from therapy with drugs interfering with DNA synthesis (e.g. hydroxycarbamide (hydroxyurea), azathioprine, zidovudine – AZT)
- myelodysplasia due to dyserythropoiesis.

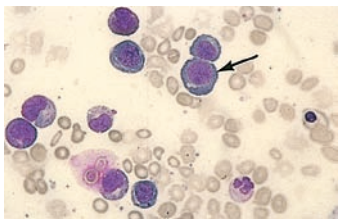


Fig. 8.10 **Megaloblasts (arrowed) in the bone marrow.**

FURTHER READING

Fleming RE, Bacon BR. Orchestration of iron homeostasis. *New England Journal of Medicine* 2005; **352:** 1741–1744.

Weiss G, Goodnough LT. Anemia of chronic disease. *New England Journal of Medicine* 2005; **352:** 1011–1023.

More online
www.studentconsult.com

Haematological findings

- Anaemia may be present. The MCV is characteristically >96 fL unless there is a coexisting cause of microcytosis when there may be a dimorphic picture with a normal/low average MCV.
- The peripheral blood film shows oval macrocytes with hypersegmented polymorphs with six or more lobes in the nucleus (Fig. 8.11).
- If severe, there may be leucopenia and thrombocytopenia.

Biochemical basis of megaloblastic anaemia

The key biochemical problem common to both vitamin B_{12} and folate deficiency is a block in DNA synthesis owing to an inability to methylate deoxyuridine monophosphate to deoxythymidine monophosphate, which is then used to build DNA (Fig. 8.12). The methyl group is supplied by the folate coenzyme, methylene tetrahydrofolate.

Deficiency of folate reduces the supply of this coenzyme; deficiency of vitamin B_{12} also reduces its supply by slowing the demethylation of methyltetrahydrofolate (methyl THF) and preventing cells receiving tetrahydrofolate for synthesis of methylene tetrahydrofolate polyglutamate.

Other congenital and acquired forms of megaloblastic anaemia are due to interference with purine or pyrimidine synthesis causing an inhibition in DNA synthesis.

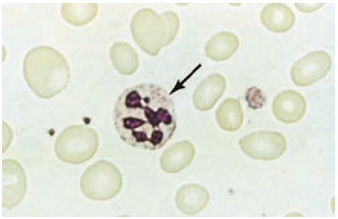


Fig. 8.11 **Macrocytes and a hypersegmented neutrophil (arrowed) on a peripheral blood film.**

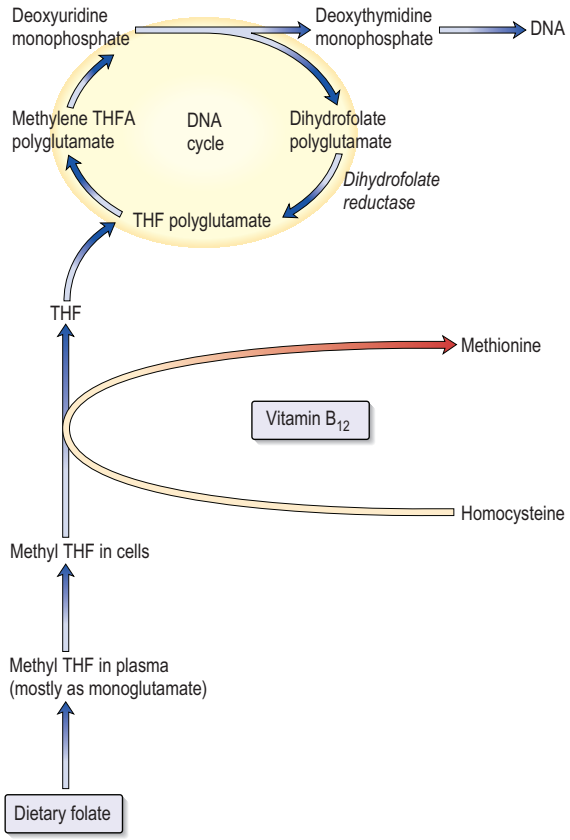


Fig. 8.12 **Biochemical basis of megaloblastic anaemia.** The metabolic relationship between vitamin B_{12} and folate and their role in DNA synthesis. THF, tetrahydrofolate.

Deoxyuridine suppression test

This is a useful method for rapidly determining the nature and severity of the vitamin B_{12} or folate deficiency in severe or complex cases of megaloblastic anaemia.

Tritiated thymidine is added to the patient's bone marrow in vitro. In a normoblastic marrow, the thymidine requirement is supplied by the methylation of deoxyuridine and this 'suppresses' the requirement for preformed tritiated thymidine to less than 5%. In a megaloblastic marrow, however, much more tritiated thymidine is used (5–50%). If the addition of B_{12} corrects the abnormality, it suggests that B_{12} is the cause of the deficiency. The addition of folate corrects the abnormality in both vitamin B_{12} and folate deficiency.

Vitamin B_{12}

Vitamin B_{12} is synthesized by certain microorganisms, and humans are ultimately dependent on animal sources. It is found in meat, fish, eggs and milk, but not in plants. Vitamin B_{12} is not usually destroyed by cooking. The average daily diet contains 5–30 µg of vitamin B_{12}, of which 2–3 µg is absorbed. The average adult stores some 2–3 mg, mainly in the liver, and it may take 2 years or more after absorptive failure before B_{12} deficiency develops, as the daily losses are small (1–2 µg).

Structure and function

Cobalamins consist of a planar group with a central cobalt atom (corrin ring) and a nucleotide set at right-angles (Fig. 8.13). Vitamin B_{12} was first crystallized as cyanocobalamin, but the main natural cobalamins have deoxyadenosyl-, methyl- and hydroxocobalamin groups attached to the cobalt atom.

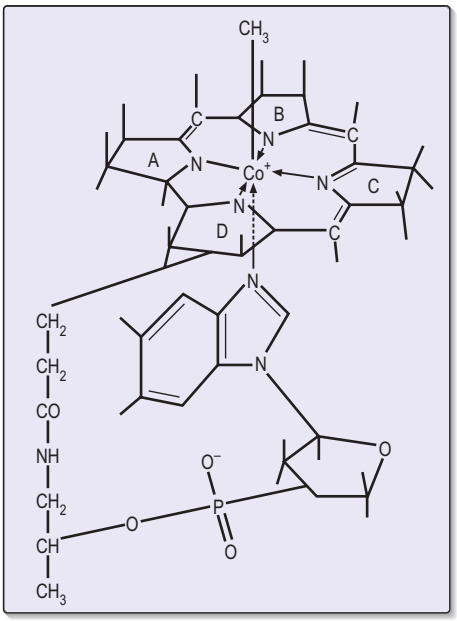


Fig. 8.13 **Methylcobalamin structure.** This is the main form of vitamin B_{12} in the plasma.

The main function of B_{12} is the methylation of homocysteine to methionine with the demethylation of methyl THF polyglutamate to THF. THF is a substrate for folate polyglutamate synthesis.

Deoxyadenosylcobalamin is a coenzyme for the conversion of methylmalonyl CoA to succinyl CoA. Measurement of methylmalonic acid in urine was used as a test for vitamin B_{12} deficiency but it is no longer carried out routinely.

Absorption and transport

Vitamin B_{12} is liberated from protein complexes in food by gastric enzymes and then binds to a vitamin B_{12}-binding protein ('R' binder) which is related to *plasma transcobalamin I* (TCI) and is derived from saliva. Vitamin B_{12} is released from the 'R' binder by pancreatic enzymes and then becomes bound to intrinsic factor.

Intrinsic factor is a glycoprotein with a molecular weight of 45 000. It is secreted by gastric parietal cells along with H^+ ions. It combines with vitamin B_{12} and carries it to specific receptors on the surface of the mucosa of the ileum. Vitamin B_{12} enters the ileal cells and intrinsic factor remains in the lumen and is excreted. Vitamin B_{12} is transported from the enterocytes to the bone marrow and other tissues by the *glycoprotein transcobalamin II* (TCII). Although TCII is the essential carrier protein for vitamin B_{12}, the amount of B_{12} on TCII is low. However, it has a rapid clearance and is able to deliver cobalamin to all cells of the body. Vitamin B_{12} in plasma is mainly bound to TCI (70–90%), but the functional role of this protein is unknown. About 1% of an oral dose of B_{12} is absorbed 'passively' without the need for intrinsic factor.

Vitamin B_{12} deficiency

There are a number of causes of B_{12} deficiency and abnormal B_{12} metabolism (Table 8.5). The most common cause of vitamin B_{12} deficiency in adults is pernicious anaemia. Malabsorption of vitamin B_{12} because of pancreatitis, coeliac disease or treatment with metformin is mild and does not usually result in significant vitamin B_{12} deficiency.

Pernicious anaemia

Pernicious anaemia (PA) is an autoimmune disorder in which there is atrophic gastritis with loss of parietal cells in the gastric mucosa with consequent failure of intrinsic factor production and vitamin B_{12} malabsorption.

Table 8.5 Vitamin B_{12} deficiency and abnormal B_{12} utilization: further causes (see text)

Low dietary intake	Abnormal utilization
Vegans	Congenital transcobalamin II deficiency
	Nitrous oxide (inactivates B_{12})
Impaired absorption	
Stomach	
Pernicious anaemia	
Gastrectomy	
Congenital deficiency of intrinsic factor	
Small bowel	
Ileal disease or resection	
Bacterial overgrowth	
Tropical sprue	
Fish tapeworm (*Diphyllobothrium latum*)	

Pathogenesis of pernicious anaemia

This disease is common in the elderly, with 1 in 8000 of the population aged over 60 years being affected in the UK. It can be seen in all races, but occurs more frequently in fair-haired and blue-eyed individuals, and those who have the blood group A. It is more common in females than males.

There is an association with other autoimmune diseases, particularly thyroid disease, Addison's disease and vitiligo. Approximately one-half of all patients with PA have thyroid antibodies. There is a higher incidence of gastric carcinoma with PA (1–3%) than in the general population.

Parietal cell antibodies are present in the serum in 90% of patients with PA – and also in many older patients with gastric atrophy. Conversely, intrinsic factor antibodies, although found in only 50% of patients with PA, are specific for this diagnosis. Two types of intrinsic factor antibodies are found: a blocking antibody, which inhibits binding of intrinsic factor to B_{12}, and a precipitating antibody, which inhibits the binding of the B_{12}–intrinsic factor complex to its receptor site in the ileum.

B_{12} deficiency may rarely occur in children from a congenital deficiency or abnormality of intrinsic factor, or as a result of early onset of the adult autoimmune type.

Pathology

Autoimmune gastritis (see p. 263) affecting the fundus is present with plasma cell and lymphoid infiltration. The parietal and chief cells are replaced by mucin-secreting cells. There is achlorhydria and absent secretion of intrinsic factor. The histological abnormality can be improved by corticosteroid therapy, which supports an autoimmune basis for the disease.

Clinical features

The onset of PA is insidious, with progressively increasing symptoms of anaemia. Patients are sometimes said to have a lemon-yellow colour owing to a combination of pallor and mild jaundice caused by excess breakdown of haemoglobin. A red sore tongue (glossitis) and angular stomatitis are sometimes present.

The neurological changes, if left untreated for a long time, can be irreversible. These neurological abnormalities occur only with very low levels of serum B_{12} (less than 60 ng/L or 50 pmol/L) and occasionally occur in patients who are not clinically anaemic. The classical neurological features are those of a polyneuropathy progressively involving the peripheral nerves and the posterior and eventually the lateral columns of the spinal cord (subacute combined degeneration; see p. 1175). Patients present with symmetrical paraesthesiae in the fingers and toes, early loss of vibration sense and proprioception, and progressive weakness and ataxia. Paraplegia may result. Dementia, psychiatric problems, hallucinations, delusions, and optic atrophy may occur from vitamin B_{12} deficiency.

Investigations

- **Haematological findings** show the features of a megaloblastic anaemia as described on page 398.
- **Bone marrow** shows the typical features of megaloblastic erythropoiesis (Fig. 8.10), although

it is frequently not performed in cases of straightforward macrocytic anaemia and a low serum vitamin B_{12}.

- **Serum bilirubin** may be raised as a result of ineffective erythropoiesis. Normally a minor fraction of serum bilirubin results from premature breakdown of newly formed red cells in the bone marrow. In many megaloblastic anaemias, where the destruction of developing red cells is much increased, the serum bilirubin can be increased. LDH can also be raised due to haemolysis.
- **Serum methylmalonic acid (MMA) and homocysteine (HC)** are raised in B_{12} deficiency. They are only useful in cases where the B_{12} and folate levels are not conclusive with only HC raised in folate deficiency.
- **Serum vitamin B_{12}** is usually well below 160 ng/L, which is the lower end of the normal range. Serum vitamin B_{12} can be assayed using radioisotope dilution or immunological assays.
- **Serum folate level** is normal or high, and the red cell folate is normal or reduced owing to inhibition of normal folate synthesis.

Absorption tests

Vitamin B_{12} absorption tests are performed only occasionally when the underlying cause of the B_{12} deficiency is not obvious. They cannot be performed in the UK as radioactive B_{12} is not available. However, the principle of the absorption test is useful.

Schilling test. Radioactive B_{12} is given orally followed by an i.m. injection of non-radioactive B_{12} to saturate B_{12} binding proteins and to flush out ^{58}Co-B_{12}. The urine is collected for 24 hours and >10% of the oral dose would be excreted in a normal person. If this is abnormal, the test is repeated with the addition of oral intrinsic factor capsules. If the excretion is now normal, the diagnosis is pernicious anaemia or gastrectomy. If the excretion is still abnormal, the lesion must be in the terminal ileum or there may be bacterial overgrowth. The latter could be confirmed by repeating the test after a course of antibiotics.

Gastrointestinal investigations

In PA there is achlorhydria. Intubation studies can be performed to confirm this but are rarely carried out in routine practice. Endoscopy or barium meal examination of the stomach is performed only if gastric symptoms are present.

Differential diagnosis

Vitamin B_{12} deficiency must be differentiated from other causes of megaloblastic anaemia, principally folate deficiency, but usually this is quite clear from the blood level of these two vitamins.

Pernicious anaemia should be distinguished from other causes of vitamin B_{12} deficiency (Table 8.5). Any disease involving the terminal ileum or bacterial overgrowth in the small bowel can produce vitamin B_{12} deficiency (see p. 280). Gastrectomy can lead, in the long term, to vitamin B_{12} deficiency. Vegans are strict vegetarians and eat no meat or animal products and prophylactic B_{12} should be given.

Treatment

See page 401.

Folic acid

Folic acid monoglutamate is not itself present in nature but occurs as polyglutamates (extra glutamic acid residues). Folates are present in food as polyglutamates in the reduced dihydrofolate or tetrahydrofolate (THF) forms (Fig. 8.14), with methyl (CH_3), formyl (CHO) or methylene (CH_2) groups attached to the pteridine part of the molecule. Polyglutamates are broken down to monoglutamates in the upper gastrointestinal tract, and during the absorptive process these are converted to methyl THF monoglutamate, which is the main form in the serum. The methylation of homocysteine to methionine requires both methylcobalamin and methyl THF as coenzymes. This reaction is the first step in which methyl THF entering cells from the plasma is converted into folate polyglutamates. Intracellular polyglutamates are the active forms of folate and act as coenzymes in the transfer of single carbon units in amino acid metabolism and DNA synthesis (Fig. 8.12).

Dietary intake

Folate is found in green vegetables such as spinach and broccoli, and offal, such as liver and kidney. Cooking causes a loss of 60–90% of the folate. The minimal daily requirement is about 100 μg.

Folate deficiency

The causes of folate deficiency are shown in Table 8.6. The main cause is poor intake, which may occur alone or in combination with excessive utilization or malabsorption. The

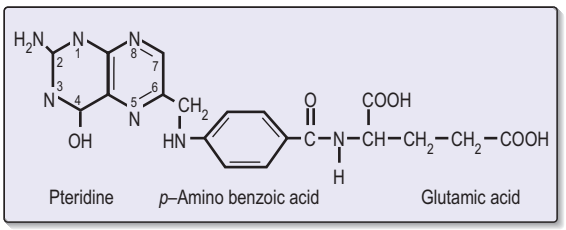


Fig. 8.14 **Folic acid structure.** This is formed from three building blocks as shown. Tetrahydrofolate has additional hydrogen atoms at positions 5, 6, 7 and 8.

Table 8.6 Causes of folate deficiency

Nutritional (major cause)

Poor intake
Old age
Poor social conditions
Starvation
Alcohol excess (also causes impaired utilization)

Poor intake due to anorexia
Gastrointestinal disease, e.g. partial gastrectomy, coeliac disease, Crohn's disease
Cancer

Antifolate drugs
Anticonvulsants:
Phenytoin
Primidone
Methotrexate
Pyrimethamine
Trimethoprim

Excess utilization

Physiological
Pregnancy
Lactation
Prematurity

Pathological
Haematological disease with excess red cell production, e.g. haemolysis
Malignant disease with increased cell turnover
Inflammatory disease
Metabolic disease, e.g. homocystinuria
Haemodialysis or peritoneal dialysis

Malabsorption
Occurs in small bowel disease, but the effect is minor compared with that of anorexia

body's reserves of folate, unlike vitamin B_{12}, are low (about 10 mg). On a deficient diet, folate deficiency develops over the course of about 4 months, but folate deficiency may develop rapidly in patients who have both a poor intake and excess utilization of folate (e.g. patients in intensive care units).

There is no simple relationship between maternal folate status and fetal abnormalities but folic acid supplements at the time of conception and in the first 12 weeks of pregnancy reduce the incidence of neural tube defects. Despite worldwide public health campaigns recommending periconceptional daily supplementation of synthetic folic acid to reduce the risk of neural tube defects, many women are not following these recommendations. In the USA and Canada mandatory fortification of grain products, e.g. bread, flour and rice, has substantially improved folate status and has been associated with a significant fall in neural tube defects.

Clinical features

Patients with folate deficiency may be asymptomatic or present with symptoms of anaemia or of the underlying cause. Glossitis can occur. Unlike with B_{12} deficiency, neuropathy does not occur.

Investigations

The haematological findings are those of a megaloblastic anaemia as discussed on page 398.

Blood measurements

Serum and red cell folate are assayed by radioisotope dilution or immunological methods. Normal levels of serum folate are 4–18 μg/L (5–63 nmol/L). The amount of folate in the red cells is a better measure of tissue folate; the normal range is 160–640 μg/mL.

Further investigations

In many cases of folate deficiency the cause is not obvious from the clinical picture or dietary history. Occult gastrointestinal disease should then be suspected and appropriate investigations, such as small bowel biopsy, should be performed.

Treatment and prevention of megaloblastic anaemia

Treatment depends on the type of deficiency. Blood transfusion is not indicated in chronic anaemia; indeed, it is dangerous to transfuse elderly patients, as heart failure may be precipitated. Folic acid may produce a haematological response in vitamin B_{12} deficiency but may aggravate the neuropathy. Large doses of folic acid alone should not be used to treat megaloblastic anaemia unless the serum vitamin B_{12} level is known to be normal. In severely ill patients, it may be necessary to treat with both folic acid and vitamin B_{12} while awaiting serum levels.

Treatment of vitamin B_{12} deficiency

Hydroxocobalamin 1000 μg can be given intramuscularly to a total of 5–6 mg over the course of 3 weeks; 1000 μg is then necessary every 3 months for the rest of the patient's life. Alternatively, it is now recommended that oral B_{12} 2 mg per day is given, as 1–2% of an oral dose is absorbed by diffusion and therefore does not require intrinsic factor.

Compliance with an oral daily regimen may be a problem, particularly in elderly patients. The use of sublingual nuggets of B_{12} (2×1000 μg daily) has been suggested to be an effective and more convenient option.

Clinical improvement may occur within 48 hours and a reticulocytosis can be seen some 2–3 days after starting therapy, peaking at 5–7 days. Improvement of the polyneuropathy may occur over 6–12 months, but longstanding spinal cord damage is irreversible. Hypokalaemia can occur and, if severe, supplements should be given. Iron deficiency often develops in the first few weeks of therapy. Hyperuricaemia also occurs but clinical gout is uncommon. In patients who have had a total gastrectomy or an ileal resection, vitamin B_{12} should be monitored; if low levels occur, prophylactic vitamin B_{12} should be given.

Treatment of folate deficiency

Folate deficiency can be corrected by giving 5 mg of folic acid daily; the same haematological response occurs as seen after treatment of vitamin B_{12} deficiency. Treatment should be given for about 4 months to replace body stores. Any underlying cause, e.g. coeliac disease, should be treated.

Prophylactic folic acid (400 μg daily) is recommended for all women planning a pregnancy to reduce neural tube defects. Many authorities also recommend prophylactic administration of folate throughout pregnancy. Whether this can be achieved by increased consumption of foods with a high folate content or whether women should take folate supplements is under debate.

Women who have had a child with a neural tube defect should take 5 mg folic acid daily before and during a subsequent pregnancy.

Prophylactic folic acid is also given in chronic haematological disorders where there is rapid cell turnover. A dose of 5 mg each week is probably sufficient.

MACROCYTOSIS WITHOUT MEGALOBLASTIC CHANGES

A raised MCV with macrocytosis on the peripheral blood film can occur with a normoblastic rather than a megaloblastic bone marrow.

A common *physiological* cause of macrocytosis is pregnancy. Macrocytosis may also occur in the newborn.

Common *pathological* causes are:

- alcohol excess
- liver disease
- reticulocytosis
- hypothyroidism
- some haematological disorders (e.g. aplastic anaemia, sideroblastic anaemia, pure red cell aplasia)
- drugs (e.g. cytotoxics – azathioprine)
- spurious (agglutinated red cells measured on red cell counters)
- cold agglutinins due to autoagglutination of red cells (see p. 416) (the MCV decreases to normal with warming of the sample to 37°C).

In all these conditions, normal serum levels of vitamin B_{12} and folate will be found. The exact mechanisms in each case are uncertain, but in some there is increased lipid deposition in the red cell membrane.

An increased number of reticulocytes leads to a raised MCV because they are large cells.

Alcohol is a frequent cause of a raised MCV in an otherwise normal individual. A megaloblastic anaemia can also occur in people who abuse alcohol; this is due to a toxic effect of alcohol on erythropoiesis or to dietary folate deficiency.

ANAEMIA DUE TO MARROW FAILURE (APLASTIC ANAEMIA)

Aplastic anaemia is defined as pancytopenia with hypocellularity (aplasia) of the bone marrow; there are no leukaemic, cancerous or other abnormal cells in the peripheral blood or bone marrow. It is an uncommon but serious condition that may be inherited but is more commonly acquired.

Aplastic anaemia is due to a reduction in the number of pluripotential stem cells (Fig. 8.1) together with a fault in those remaining or an immune reaction against them so that they are unable to repopulate the bone marrow. Failure of only one cell line may also occur, resulting in isolated deficiencies such as the absence of red cell precursors in pure red cell aplasia. Evolution to myelodysplasia, paroxysmal nocturnal haemoglobinuria (PNH) or acute myeloblastic leukaemia occurs in some cases, probably owing to the emergence of an abnormal clone of haemopoietic cells.

Causes

A list of causes of aplasia is given in Table 8.7. Immune mechanisms are probably responsible for most cases of idiopathic acquired aplastic anaemia and play a part in at least the persistence of many secondary cases. Activated cytotoxic T cells in blood and bone marrow are responsible for the bone marrow failure.

Many drugs may cause marrow aplasia, including cytotoxic drugs such as busulfan and doxorubicin, which are expected to cause transient aplasia as a consequence of their therapeutic use. However, some individuals develop aplasia due to sensitivity to non-cytotoxic drugs such as chloramphenicol, gold, carbimazole, chlorpromazine, phenytoin, ribovirin, tolbutamide, non-steroidal anti-inflammatory agents, and many others which have been reported to cause occasional cases of aplasia.

Congenital aplastic anaemias are rare. Gene mutations are being identified, e.g. the telomerase RNA component, and have been seen in one-third of aplastic anaemias. Fanconi's anaemia is inherited as an autosomal recessive and is associated with skeletal, renal and central nervous system abnormalities. It usually presents between the ages of 5 and 10 years.

Clinical features

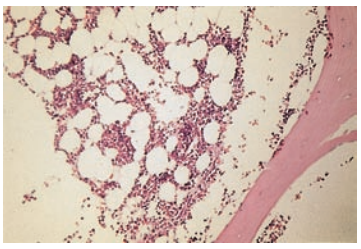
The clinical manifestations of marrow failure from any cause are anaemia, bleeding and infection. Bleeding is often the predominant initial presentation of aplastic anaemia with bruising with minimal trauma or blood blisters in the mouth. Physical findings include ecchymoses, bleeding gums and epistaxis. Mouth infections are common. Lymphadenopathy and hepatosplenomegaly are rare in aplastic anaemia.

Investigations

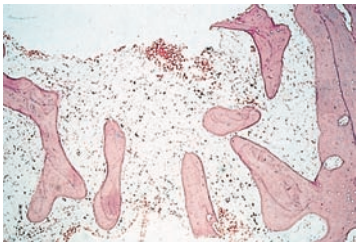
- Pancytopenia
- The virtual absence of reticulocytes
- A hypocellular or aplastic bone marrow with increased fat spaces (Fig. 8.15).

Differential diagnosis

This is from other causes of pancytopenia (Table 8.8). A bone marrow trephine is essential for assessment of the bone marrow cellularity.



(a)



(b)

Fig. 8.15 **Bone marrow trephine biopsies in low-power view. (a)** Normal cellularity. **(b)** Hypocellularity in aplastic anaemia.

Table 8.7 Causes of aplastic anaemia

Primary
Congenital, e.g. Fanconi's anaemia
Idiopathic acquired (67% of cases)

Secondary
Chemicals, e.g. benzene, toluene, glue sniffing
Drugs:
- e.g. chemotherapeutic (idiosyncratic reactions)
- Antibiotics, e.g. chloramphenicol, gold, penicillamine, phenytoin, carbamazepine, carbimazole

Insecticides
Ionizing radiation
Infections:
- Viral, e.g. hepatitis, EBV, HIV,erythrovirus
- Other, e.g. tuberculosis

Paroxysmal nocturnal haemoglobinuria
Miscellaneous, e.g. pregnancy

Table 8.8 Causes of pancytopenia

Aplastic anaemia (see Table 8.7)
Drugs
Megaloblastic anaemia
Bone marrow infiltration or replacement:
- Hodgkin's and non-Hodgkin's lymphoma
- Acute leukaemia
- Myeloma
- Secondary carcinoma:
  - Myelofibrosis

Hypersplenism
Systemic lupus erythematosus
Disseminated tuberculosis
Paroxysmal nocturnal haemoglobinuria
Overwhelming sepsis

FURTHER READING

British Committee for Standards in Haematology. Guidelines for the diagnosis and management of acquired aplastic anaemia. *British Journal of Haematology* 2003; **123:** 782–801.

Brodsky RA, Jones RJ. Aplastic anaemia. *Lancet* 2005; **365:** 1647–1656.

Treatment and prognosis

The treatment of aplastic anaemia depends on providing supportive care while awaiting bone marrow recovery and specific treatment to accelerate marrow recovery.

The main danger is infection and stringent measures should be undertaken to avoid this (see also p. 461). Any suspicion of infection in a severely neutropenic patient should lead to immediate institution of broad-spectrum parenteral antibiotics. Supportive care including transfusions of red cells and platelets should be given as necessary. The cause of the aplastic anaemia must be eliminated if possible.

The course of aplastic anaemia can be variable, ranging from a rapid spontaneous remission to a persistent increasingly severe pancytopenia, which may lead to death through haemorrhage or infection. The most reliable determinants for the prognosis are the number of neutrophils, reticulocytes, platelets, and the cellularity of the bone marrow.

A bad prognosis (i.e. severe aplastic anaemia) is associated with the presence of two of the following three features:

- neutrophil count of $<0.5 \times 10^9$/L
- platelet count of $<20 \times 10^9$/L
- reticulocyte count of $<40 \times 10^9$/L.

In severe aplastic anaemia, there is a very poor outcome without treatment.

Bone marrow transplantation is the treatment of choice for some patients. For the patients *under the age of 40* who fail to respond to immunosuppression, bone marrow transplantation using unrelated donors is an option, but the results are poor (5-year survival of only 30%) owing to a high incidence of graft rejection, graft-versus-host disease and viral infections. For those with an HLA-identical sibling donor, it gives a 75–90% chance of long-term survival and restoring the blood count to normal. Patients *over the age of 40* are not eligible for bone marrow transplantation whether an HLA-identical donor is available or not, because of the high risk of graft-versus-host disease as a complication of bone marrow transplantation. Immunosuppressive therapy is used for patients without HLA-matched siblings and those over the age of 40 years; antilymphocyte globulin (ALG) and ciclosporin in combination give a response rate of 60–80%.

Levels of haemopoietic growth factors (Fig. 8.1) are normal or increased in most patients with aplastic anaemia, and are ineffective as primary treatment.

Androgens (e.g. oxymetholone) are sometimes useful in patients not responding to immunosuppression and in patients with moderately severe aplastic anaemia.

Steroids have little activity in severe aplastic anaemia but are used for serum sickness due to ALG. They are also used to treat children with congenital pure red cell aplasia (Diamond–Blackfan syndrome). Adult pure red cell aplasia is associated with a thymoma in 30% of cases and thymectomy may induce a remission. It may also be associated with autoimmune disease or may be idiopathic. Steroids and ciclosporin are effective treatment in some cases.

HAEMOLYTIC ANAEMIAS: AN INTRODUCTION

Haemolytic anaemias are caused by increased destruction of red cells. The red cell normally survives about 120 days, but in haemolytic anaemias the red cell survival times are considerably shortened.

Breakdown of normal red cells occurs in the macrophages of the bone marrow, liver and spleen (Fig. 8.5).

Consequences of haemolysis (Fig. 8.16)

Shortening of red cell survival does not always cause anaemia as there is a compensatory increase in red cell production by the bone marrow. If the red cell loss can be contained within the marrow's capacity for increased output, then a haemolytic state can exist without anaemia (*compensated haemolytic disease*). The bone marrow can increase its output by six to eight times by increasing the proportion of cells committed to erythropoiesis (*erythroid hyperplasia*) and by expanding the volume of active marrow. In addition, immature red cells (*reticulocytes*) are released prematurely. These cells are larger than mature cells and stain with a light blue tinge on a peripheral blood film (the description of this appearance on the blood film is *polychromasia*). Reticulocytes may be counted accurately as a percentage of all red cells on a blood film using a supravital stain for residual RNA (e.g. new methylene blue).

Sites of haemolysis

Extravascular haemolysis

In most haemolytic conditions red cell destruction is extravascular. The red cells are removed from the circulation by macrophages in the reticuloendothelial system, particularly the spleen.

Intravascular haemolysis

When red cells are rapidly destroyed within the circulation, haemoglobin is liberated (Fig. 8.17). This is initially bound to plasma haptoglobins but these soon become saturated.

Excess free plasma Hb is filtered by the renal glomerulus and enters the urine, although small amounts are reabsorbed by the renal tubules. In the renal tubular cell, Hb is broken down and becomes deposited in the cells as *haemosiderin*. This can be detected in the spun sediment of urine using Perls' reaction. Some of the free plasma Hb is oxidized to *methaemoglobin*, which dissociates into *ferrihaem* and globin. *Plasma haemopexin* binds ferrihaem; but if its binding capacity is exceeded, ferrihaem becomes attached to albumin, forming *methaemalbumin*. On spectrophotometry of the plasma, methaemalbumin forms a characteristic band; this is the basis of *Schumm's test*.

The *liver* plays a major role in removing Hb bound to haptoglobin and haemopexin and any remaining free Hb.

Evidence for haemolysis

Increased red cell breakdown is accompanied by increased red cell production. This is shown in Figure 8.16.

Demonstration of shortened red cell lifespan

Red cell survival can be estimated from ^{51}Cr-labelled red cells given intravenously but is rarely performed.

Intravascular haemolysis

This is suggested by raised levels of plasma Hb, haemosiderinuria, very low or absent haptoglobins, and the presence of methaemalbumin (positive Schumm's test).

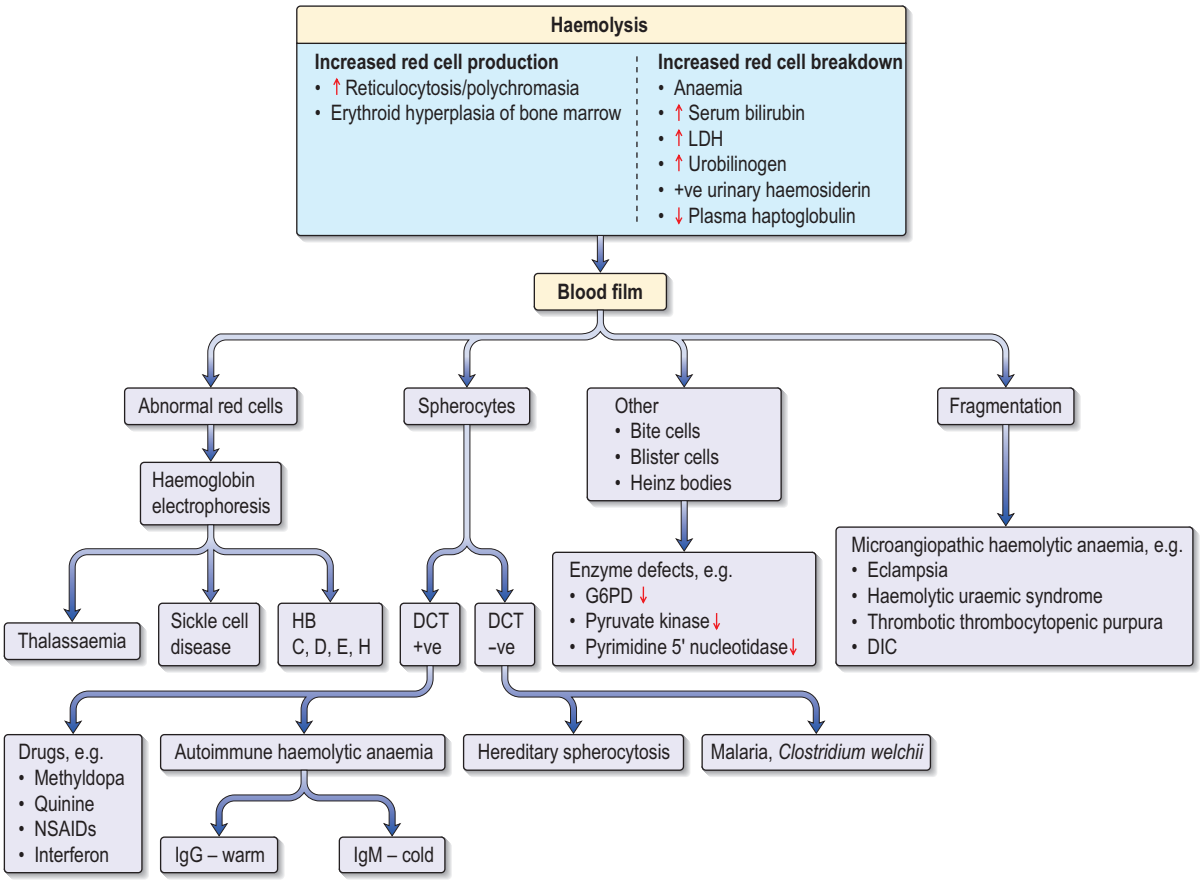


Fig. 8.16 **Haemolysis.** DCT, direct Coombs' test; G6PD, glucose-6-phosphate dehydrogenase; LDH, lactate dehydrogenase.

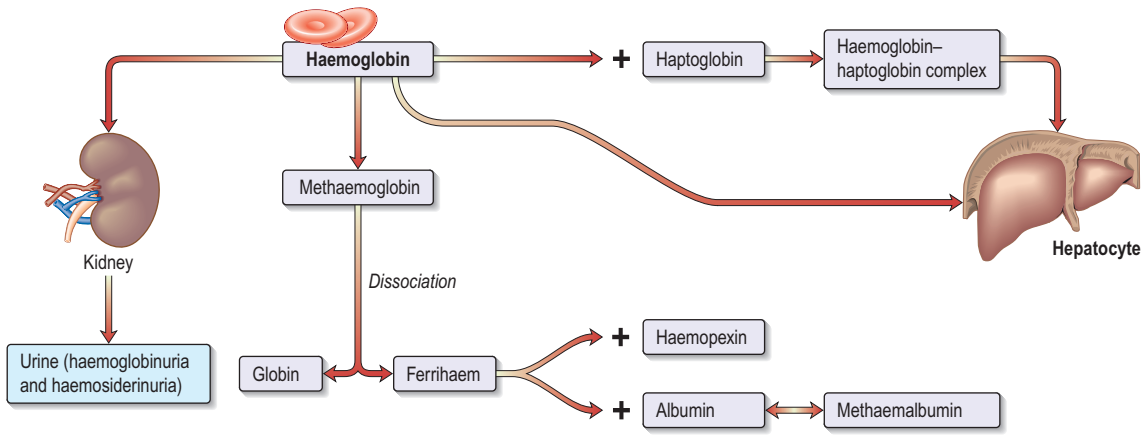


Fig. 8.17 **The fate of haemoglobin in the plasma following haemolysis.**

Various laboratory studies will be necessary to determine the exact type of haemolytic anaemia present. The causes of haemolytic anaemias are shown in Table 8.9.

INHERITED HAEMOLYTIC ANAEMIA

RED CELL MEMBRANE DEFECTS

The normal red cell membrane consists of a lipid bilayer crossed by integral proteins with an underlying lattice of proteins (or cytoskeleton), including spectrin, actin, ankyrin and protein 4.1, attached to the integral proteins (Fig. 8.18).

Hereditary spherocytosis (HS)

HS is the most common inherited haemolytic anaemia in northern Europeans, affecting 1 in 5000. It is inherited in an autosomal dominant manner, but in 25% of patients neither parent is affected and it is presumed that HS has occurred by spontaneous mutation. HS is due to defects in the red cell membrane, resulting in the cells losing part of the cell membrane as they pass through the spleen, possibly because the lipid bilayer is inadequately supported by the membrane

skeleton. The best-characterized defect is a deficiency in the structural protein spectrin, but quantitative defects in other membrane proteins have been identified (Fig. 8.18), with ankyrin defects being the most common. The abnormal red cell membrane in HS is associated functionally with an increased permeability to sodium, and this requires an increased rate of active transport of sodium out of the cells which is dependent on ATP produced by glycolysis. The surface-to-volume ratio decreases, and the cells become spherocytic. Spherocytes are more rigid and less deformable than normal red cells. They are unable to pass through the splenic microcirculation so they have a shortened lifespan.

Clinical features

The condition may present with jaundice at birth. However, the onset of jaundice can be delayed for many years and some patients may go through life with no symptoms and are detected only during family studies. The patient may eventually develop anaemia, splenomegaly and ulcers on the leg. As in many haemolytic anaemias, the course of the disease may be interrupted by aplastic, haemolytic and megaloblastic crises. Aplastic anaemia usually occurs after infections, particularly with erythro(parvo)virus, whereas megaloblastic anaemia is the result of folate depletion owing to the hyperactivity of the bone marrow. Chronic haemolysis leads to the formation of pigment gallstones (see p. 367).

Table 8.9 Causes of haemolytic anaemia

Inherited

Red cell membrane defect
Hereditary spherocytosis
Hereditary elliptocytosis

Haemoglobin abnormalities
Thalassaemia
Sickle cell disease

Metabolic defects
Glucose-6-phosphate dehydrogenase deficiency
Pyruvate kinase deficiency
Pyrimidine kinase deficiency

Miscellaneous
Infections, e.g. malaria, mycoplasma
Clostridium welchii, generalized sepsis
Drugs and chemicals causing damage to the red cell membrane or oxidative haemolysis
Hypersplenism
Burns

Acquired

Immune
Autoimmune (see Table 8.15)
- Warm
- Cold

Alloimmune
- Haemolytic transfusion reactions
- Haemolytic disease of the newborn
- After allogeneic bone marrow or organ transplantation

Drug-induced

Non-immune
Acquired membrane defects
- Paroxysmal nocturnal haemoglobinuria

Mechanical
- Microangiopathic haemolytic anaemia
- Valve prosthesis
- March haemoglobinuria

Secondary to systemic disease
- Renal and liver failure

Investigations

- **Anaemia**. This is usually mild, but occasionally can be severe.
- **Blood film**. This shows spherocytes (Fig. 8.19) and reticulocytes.
- **Haemolysis** is evident (e.g. the serum bilirubin and urinary urobilinogen will be raised).
- **Osmotic fragility**. When red cells are placed in solutions of increasing hypotonicity, they take in water, swell, and eventually lyse. Spherocytes tolerate hypotonic solutions less well than do normal biconcave red cells. Osmotic fragility tests are infrequently carried out in routine practice, but may be useful to confirm a suspicion of spherocytosis on a blood film.
- **Direct antiglobulin (Coombs') test** is negative in hereditary spherocytosis, virtually ruling out autoimmune haemolytic anaemia where spherocytes are also commonly present.

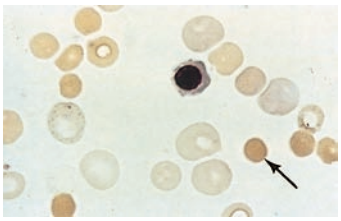


Fig. 8.19 **Spherocytes (arrowed). This blood film also shows reticulocytes, polychromasia and a nucleated erythroblast.**

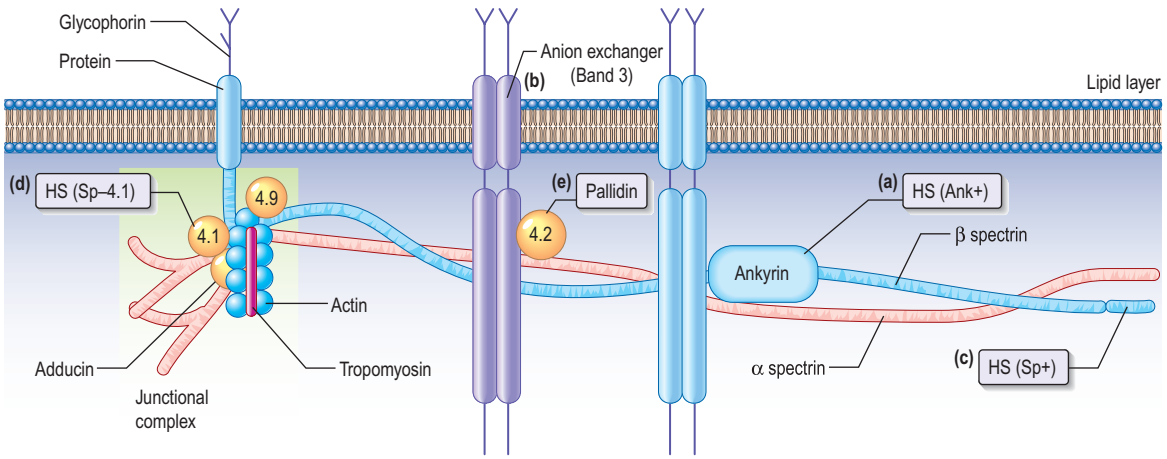


Fig. 8.18 **Hereditary spherocytosis (HS) and hereditary elliptocytosis (HE): red cell membrane showing the sites (purple) of the principal defects.** Vertical interactions producing HS: **(a)** ankyrin mutation, HS (Ank+) producing deficiency (45% of cases); **(b)** HS band 3 deficiency (20%); **(c)** β spectrin deficiency, HS (Sp+) (<20%); **(d)** abnormal spectrin/protein 4.1 binding, HS (Sp–4.1); **(e)** protein 4.2 (pallidin) deficiency (Japanese). These produce various autosomal dominant and recessive forms of the disease. Horizontal interactions producing HE: α spectrin (80%), protein 4.1 (15%), β spectrin (5%).

Treatment

The spleen, which is the site of cell destruction, should be removed in all but the mildest cases. The decision about splenectomy in symptomless patients is difficult, but a raised bilirubin and especially the presence of gallstones should encourage splenectomy.

It is best to postpone splenectomy until after childhood, as sudden overwhelming fatal infections, usually due to encapsulated organisms such as pneumococci, may occur (see p. 423). Splenectomy should be preceded by appropriate immunization and followed by lifelong penicillin prophylaxis (see Box 8.3).

Following splenectomy, the spherocytes persist but the Hb level usually returns to normal as the red cells are no longer destroyed.

Folate deficiency often occurs in chronic haemolysis with rapid cell turnover. Folate levels should be monitored, or folic acid can be given prophylactically.

Hereditary elliptocytosis

This disorder of the red cell membrane is inherited in an autosomal dominant manner and has a prevalence of 1 in 2500 in Caucasians. The red cells are elliptical due to deficiencies of protein 4.1 or the spectrin/actin/4.1 complex which leads to weakness of the horizontal protein interaction and to the membrane defect (Fig. 8.18). Clinically it is a similar condition to HS but milder. Only a minority of patients have anaemia and only occasional patients require splenectomy.

Rarely, hereditary spherocytosis or elliptocytosis may be inherited in a homozygous fashion giving rise to a severe haemolytic anaemia sometimes necessitating splenectomy in early childhood.

Hereditary stomatocytosis

Stomatocytes are red cells in which the pale central area appears slit-like. Their presence in large numbers may occur in a hereditary haemolytic anaemia associated with a membrane defect, but excess alcohol intake is also a common cause. Although these hereditary conditions are very rare, a correct diagnosis is required since splenectomy is surprisingly contraindicated as it may result in fatal thromboembolic events.

HAEMOGLOBIN ABNORMALITIES

In early embryonic life, haemoglobins Gower 1, Gower 2 and Portland predominate (Fig. 8.20). Later, fetal haemoglobin (Hb F), which has two α and two γ chains, is produced. There is increasing synthesis of β chains from 13 weeks of gestation and at term there is 80% Hb F and 20% Hb A. The switch from Hb F to Hb A occurs after birth when the genes for γ chain production are further suppressed and there is rapid increase in the synthesis of β chains. The exact mechanism responsible for the switch remains unknown. There is little Hb F produced (normally less than 1%) from 6 months after birth. The δ chain is synthesized just before birth and Hb A_2 ($\alpha_2\beta_2$) remains at a level of about 2% throughout adult life (Table 8.10).

Globin chains are synthesized in the same way as any protein (see p. 31). Four globin chain genes are required to control α-chain production (Fig. 8.20). Two are present on each haploid genome (genes derived from one parent). These are situated close together on chromosome 16. The genes controlling the production of ε, γ, δ and β chains are close together on chromosome 11. The globin genes are arranged on chromosomes 16 and 11 in the order in which they are expressed and combine to give different haemoglobins. Normal haemoglobin synthesis is discussed on page 390.

Abnormal haemoglobins

Abnormalities occur in:

- globin chain production (e.g. thalassaemia)
- structure of the globin chain (e.g. sickle cell disease)
- combined defects of globin chain production and structure, e.g. sickle cell β-thalassaemia.

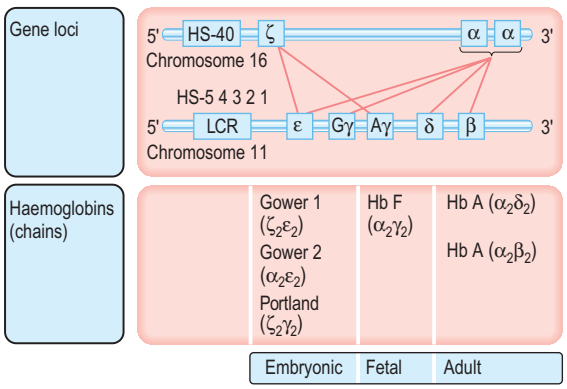


Fig. 8.20 **Loci of genes on chromosomes 16 and 11 and the combination of various chains to produce different haemoglobins.** HS-40 (for α genes) and locus control region (LCR) HS1–5 (for β genes) are regulatory control elements.

Table 8.10 Some types of haemoglobin

	Haemoglobin	Structure	Comment
Normal	A	$\alpha_2\beta_2$	Comprises 97% of adult haemoglobin
	A_2	$\alpha_2\delta_2$	Comprises 2% of adult haemoglobin Elevated in β-thalassaemia
	F	$\alpha_2\gamma_2$	Normal haemoglobin in fetus from 3rd to 9th month Increased in β-thalassaemia Comprises < 1% of haemoglobin in adult
Abnormal chain production	H	β_4	Found in α-thalassaemia Biologically useless
	Barts	γ_4	Comprises 100% of haemoglobin in homozygous α-thalassaemia Biologically useless
Abnormal chain structure	S	$\alpha_2\beta_2^S$	Substitution of valine for glutamine acid in position 6 of β chain
	C	$\alpha_2\beta_2^C$	Substitution of lysine for glutamic acid in position 6 of β chain

www.studentconsult.com
Ask the authors

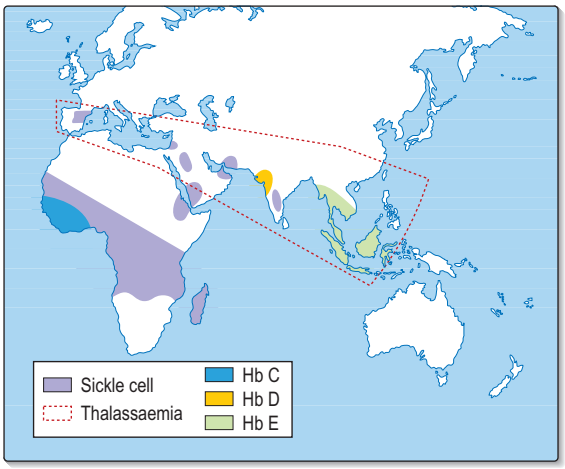


Fig. 8.21 **Major haemoglobin abnormalities: geographical distribution.**

Genetic defects in haemoglobin are the most common of all genetic disorders.

THE THALASSAEMIAS

The thalassaemias affect people throughout the world (Fig. 8.21). Normally there is balanced (1 : 1) production of α and β chains. The defective synthesis of globin chains in thalassaemia leads to 'imbalanced' globin chain production, leading to precipitation of globin chains within the red cell precursors and resulting in ineffective erythropoiesis. Precipitation of globin chains in mature red cells leads to haemolysis.

β-Thalassaemia

In *homozygous* β-thalassaemia, either no normal β chains are produced (β^0), or β-chain production is very reduced (β+). There is an excess of α chains, which precipitate in erythroblasts and red cells causing ineffective erythropoiesis and haemolysis. The excess α chains combine with whatever β, δ and γ chains are produced, resulting in increased quantities of Hb A_2 and Hb F and, at best, small amounts of Hb A. In heterozygous β-thalassaemia there is usually symptomless microcytosis with or without mild anaemia. Table 8.11 shows the findings in the homozygote and heterozygote for the common types of β-thalassaemia.

Molecular genetics

The molecular errors accounting for over 200 genetic defects leading to β-thalassaemia have been characterized. Unlike in α-thalassaemia, the defects are mainly point mutations rather than gene deletions. The mutations result in defects in transcription, RNA splicing and modification, translation via frame shifts and nonsense codons producing highly unstable β-globin which cannot be utilized.

Clinical syndromes

Clinically, β-thalassaemia can be divided into the following:

- thalassaemia minor (or trait), the symptomless heterozygous carrier state
- thalassaemia intermedia, with moderate anaemia, rarely requiring transfusions
- thalassaemia major, with severe anaemia requiring regular transfusions.

Table 8.11 β-Thalassaemia: common findings

Type of thalassaemia	Findings in homozygote	Findings in heterozygote
β^+	Thalassaemia major Hb A + F + A_2	Thalassaemia minor Hb A_2 raised
β^0	Thalassaemia major Hb F + A_2	Thalassaemia minor Hb A_2 raised
δβ	Thalassaemia intermedia	Thalassaemia minor Hb F 5–15%
	Hb F only	Hb A_2 normal
$\delta\beta^+$ (Lepore)	Thalassaemia major or intermedia Hb F and Lepore	Thalassaemia minor Hb Lepore 5–15% Hb A_2 normal

Adapted with permission from Weatherall DJ. Disorders of the synthesis of function of haemoglobin. In: Weatherall DJ, Warrell DA, Cox TM, Firth JD (eds) *Oxford Textbook of Medicine*, 4th edn. Oxford: Oxford University Press, 2003.
Hb Lepore is a cross fusion product of δ and β globin genes.

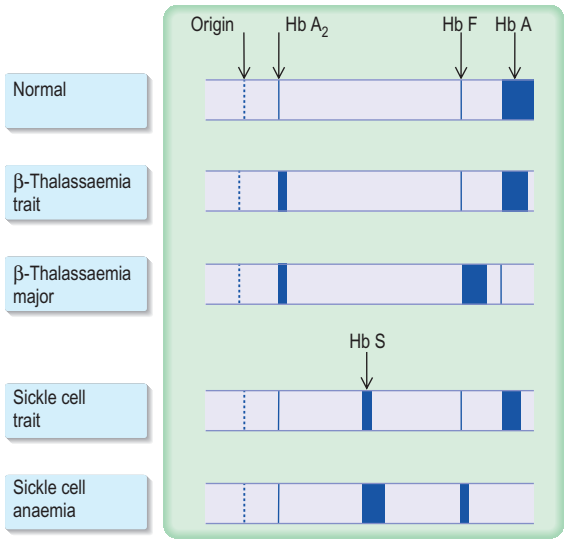


Fig. 8.22 **Patterns of haemoglobin electrophoresis.**

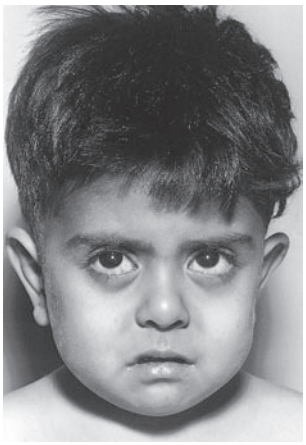
Thalassaemia minor (trait)

This common carrier state (heterozygous β-thalassaemia) is asymptomatic. Anaemia is mild or absent. The red cells are hypochromic and microcytic with a low MCV and MCH, and it may be confused with iron deficiency. However, the two are easily distinguished, as in thalassaemia trait the serum ferritin and the iron stores are normal (Table 8.3). The RDW is usually normal (see p. 389). Hb electrophoresis usually shows a raised Hb A_2 and often a raised Hb F (Fig. 8.22). Iron should not be given to these patients unless they have proven coincidental iron deficiency.

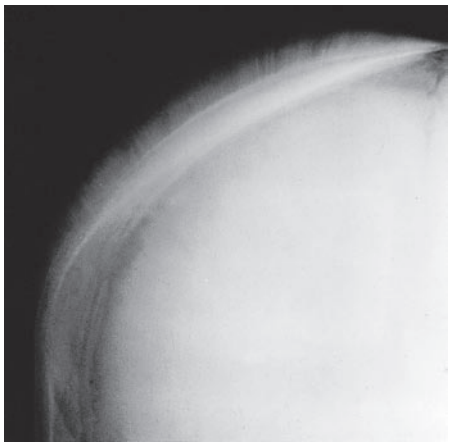
Thalassaemia intermedia

Thalassaemia intermedia includes patients who are symptomatic with moderate anaemia (Hb 7–10 g/dL) and who do not require regular transfusions.

Thalassaemia intermedia may be due to a combination of homozygous mild β+- and α-thalassaemia, where there is reduced α-chain precipitation and less ineffective erythropoi-



(a)



(b)

(C)

Fig. 8.23 **Thalassaemia. (a)** A child with thalassaemia, showing the typical facial features. **(b)** Skull X-ray of a child with β-thalassaemia, showing the 'hair on end' appearance. **(c)** X-ray of hand, showing expansion of the marrow and a thinned cortex.

esis and haemolysis. The inheritance of hereditary persistence of Hb F with homozygous β-thalassaemia also results in a milder clinical picture than unmodified β-thalassaemia major because the excess α chains are partially removed by the increased production of γ chains.

Patients may have splenomegaly and bone deformities. Recurrent leg ulcers, gallstones and infections are also seen. It should be noted that these patients may be iron overloaded despite a lack of regular blood transfusions. This is caused by excessive iron absorption which results from the underlying dyserythropoiesis (see iron absorption, p. 393).

Thalassaemia major (Cooley's anaemia)

Most children affected by homozygous β-thalassaemia present during the first year of life with:

- failure to thrive and recurrent bacterial infections
- severe anaemia from 3 to 6 months when the switch from γ- to β-chain production should normally occur
- extramedullary haemopoiesis that soon leads to hepatosplenomegaly and bone expansion, giving rise to the classical thalassaemic facies (Fig. 8.23a).

Skull X-rays in these children show the characteristic 'hair on end' appearance of bony trabeculation as a result of expansion of the bone marrow into cortical bone (Fig. 8.23b). The expansion of the bone marrow is also shown in an X-ray of the hand (Fig. 8.23c).

The classic features of untreated thalassaemia major are generally only observed in patients from countries without good blood transfusion support.

Management

The aims of treatment are to suppress ineffective erythropoiesis, prevent bony deformities and allow normal activity and development.

- *Long-term folic acid* supplements are required.
- *Regular transfusions* should be given to keep the Hb above 10 g/dL. Blood transfusions may be required every 4–6 weeks.
- If *transfusion requirements* increase, splenectomy may help, although this is usually delayed until after the age of 6 years because of the risk of infection. Prophylaxis against infection is required for patients undergoing splenectomy (see p. 423).
- *Iron overload* caused by repeated transfusions (transfusion haemosiderosis) may lead to damage to the endocrine glands, liver, pancreas and the myocardium by the time patients reach adolescence. Magnetic resonance imaging (myocardial T2- relaxation time) is useful for monitoring iron overload in thalassaemia, both the heart and the liver can be monitored. The standard iron-chelating agent remains desferrioxamine, although it has to be administered parenterally. Desferrioxamine is given as an overnight subcutaneous infusion on 5–7 nights each week. Ascorbic acid 200 mg daily is given, as it increases the urinary excretion of iron in response to desferrioxamine. Often young children have a very high standard of chelation as it is organized by their parents. However, when the children become adults and take on this role themselves they often rebel and chelation with desferrioxamine may become problematic. Preliminary results on a new once-daily oral iron chelator, deferasirox, indicate that it is safe, similar in effectiveness to desferrioxamine and well tolerated. Deferiprone, an oral iron chelator, is also now available.
- Intensive treatment with desferrioxamine has been reported to reverse damage to the heart in patients with severe iron overload, but excessive doses of desferrioxamine may cause cataracts, retinal damage and nerve deafness. Infection with *Yersinia enterocolitica* occurs in iron-loaded patients treated with desferrioxamine. Iron overload should be periodically assessed by measuring the serum ferritin and by assessment of hepatic iron stores by MRI.
- *Bone marrow transplantation* has been used in young patients with HLA-matched siblings. It has been successful in patients in good clinical condition with a 3-year mortality of less than 5%, but there is a high

mortality (>50%) in patients in poor condition with iron overload and liver dysfunction.

- *Prenatal diagnosis and gene therapy* are discussed on page 45.
- *Patients' partners should be tested.* If both partners have β-thalassaemia trait, there is a 1 in 4 chance of such pregnancy resulting in a child having β-thalassaemia major. Therefore, couples in this situation must be offered prenatal diagnosis.

α-Thalassaemia

Molecular genetics

In contrast to β-thalassaemia, α-thalassaemia is often caused by gene deletions, although mutations of the alpha globin genes may also occur. The gene for α globin chains is duplicated on both chromosomes 16, i.e. a normal person has a total of four alpha globin genes. Deletion of one α-chain gene (α+) or both α-chain genes (α0) on each chromosome 16 may occur (Table 8.12). The former is the most common of these abnormalities.

- **Four-gene deletion** (deletion of both genes on both chromosomes); there is no α-chain synthesis and only Hb Barts (γ4) is present. Hb Barts cannot carry oxygen and is incompatible with life (Tables 8.10 and 8.12). Infants are either stillborn at 28–40 weeks or die very shortly after birth. They are pale, oedematous and have enormous livers and spleens – a condition called hydrops fetalis.
- **Three-gene deletion**; there is moderate anaemia (Hb 7–10 g/dL) and splenomegaly (Hb H disease). The patients are not usually transfusion-dependent. Hb A, Hb Barts and Hb H (β4) are present. Hb A_2 is normal or reduced.
- **Two-gene deletion** (α-thalassaemia trait); there is microcytosis with or without mild anaemia. Hb H bodies may be seen on staining a blood film with brilliant cresyl blue.
- **One-gene deletion**; the blood picture is usually normal.

Globin chain synthesis studies for the detection of a reduced ratio of α to β chains may be necessary for the definitive diagnosis of α-thalassaemia trait.

Less commonly, α-thalassaemia may result from genetic defects other than deletions, for example mutations in the stop codon producing an α chain with many extra amino acids (Hb Constant Spring).

Table 8.12 The α-thalassaemias

Gene deletion		Haemoglobin type	Clinical picture
4 genes	α^0––/––	Hb Barts (γ4)	Hydrops fetalis
3 genes	α^0––/–α	Hb H (β4) Splenomegaly	Moderate anaemia
2 genes	α^0––/αα	Some Hb H bodies	Mild anaemia
	Or		
	–α/–α	Hb A	α-Thalassaemia trait
1 gene	α^+–α/αα	'Normal'	α-Thalassaemia trait

SICKLE SYNDROMES

Sickle cell haemoglobin (Hb S) results from a single-base mutation of adenine to thymine which produces a substitution of valine for glutamic acid at the sixth codon of the β-globin chain ($\alpha_2\beta_2^{6glu\rightarrow val}$). In the homozygous state (*sickle cell anaemia*) both genes are abnormal (Hb SS), whereas in the heterozygous state (sickle cell trait, Hb AS) only one chromosome carries the gene. As the synthesis of Hb F is normal, the disease usually does not manifest itself until the Hb F decreases to adult levels at about 6 months of age.

The sickle gene is commonest in Africans (up to 25% gene frequency in some populations) but is also found in India, the Middle East and Southern Europe.

Pathogenesis

Deoxygenated Hb S molecules are insoluble and polymerize. The flexibility of the cells is decreased and they become rigid and take up their characteristic sickle appearance (Fig. 8.24). This process is initially reversible but, with repeated sickling, the cells eventually lose their membrane flexibility and become irreversibly sickled. This is due to dehydration, partly caused by potassium leaving the red cells via calcium activated potassium channels called the Gados channel. These irreversibly sickled cells are dehydrated and dense and will not return to normal when oxygenated. Sickling can produce:

- a shortened red cell survival
- impaired passage of cells through the microcirculation, leading to obstruction of small vessels and tissue infarction.

Sickling is precipitated by infection, dehydration, cold, acidosis or hypoxia. In many cases the cause is unknown, but adhesion proteins on activated endothelial cells (VCAM-1) may play a causal role, particularly in vaso-occlusion when rigid cells are trapped, facilitating polymerization. Hb S releases its oxygen to the tissues more easily than does normal Hb, and patients therefore feel well despite being anaemic (except of course during crises or complications).

Depending on the type of haemoglobin chain combinations, three clinical syndromes occur:

- homozygous Hb SS have the most severe disease
- combined heterozygosity (Hb SC) for Hb S and C (see below) who suffer intermediate symptoms
- heterozygous Hb AS (sickle cell trait) have no symptoms (see p. 411).

Sickle cell anaemia

Clinical features

Vaso-occlusive crises

The earliest presentation in the first few years of life is acute pain in the hands and feet (dactylitis) owing to vaso-

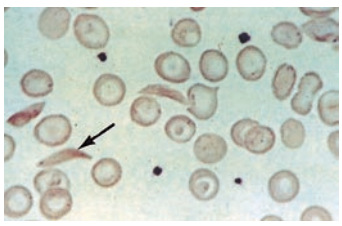


Fig. 8.24 **Sickle cells (arrowed) and target cells.**

occlusion of the small vessels. Severe pain in other bones, e.g. femur, humerus, vertebrae, ribs, pelvis, occurs in older children/adults. These attacks vary in frequency from daily to perhaps only once a year. Fever often accompanies the pain.

Anaemia

Chronic haemolysis produces a stable haemoglobin level, usually in the 6–8 g/dL range but an acute fall in the haemoglobin level can occur owing to:

- splenic sequestration
- bone marrow aplasia
- further haemolysis.

Splenic sequestration

Vaso-occlusion produces an acute painful enlargement of the spleen. There is splenic pooling of red cells and hypovolaemia, leading in some to circulatory collapse and death. The condition occurs in childhood before multiple infarctions have occurred. The latter eventually leads to a fibrotic non-functioning spleen. Liver sequestration can also occur.

Bone marrow aplasia

This most commonly occurs following infection with erythrovirus B19, which invades proliferating erythroid progenitors. There is a rapid fall in haemoglobin with no reticulocytes in the peripheral blood, because of the failure of erythropoiesis in the marrow.

Haemolysis due to drugs, acute infection or associated G6PD deficiency also occurs. Anaemia can also result from folate deficiency.

Long-term problems

In adults, nearly every organ is involved eventually, as patients survive longer with better treatment.

Growth and development. Young children are short but regain their height by adulthood. However, they remain below the normal weight. There is often delayed sexual maturation which may require hormone therapy.

Bones are a common site for vaso-occlusive episodes, leading to chronic infarcts. Avascular necrosis of hips, shoulders, compression of vertebrae and shortening of bones in the hands and feet occur. These episodes are the common cause for the painful crisis. Osteomyelitis is commoner in sickle cell disease and is caused by *Staphylococcus aureus*, *Staph. pneumoniae* and salmonella (see p. 541). Occasionally hip joint replacement may be required.

Infections are common in tissues susceptible to vaso-occlusion, e.g. bones, lungs, kidneys.

Respiratory. The acute chest syndrome occurs in up to 30%, and pulmonary hypertension and chronic lung disease are the commonest cause of death in adults with sickle cell disease. The acute chest syndrome is caused by infection, fat embolism from necrotic bone marrow or pulmonary infarction due to sequestration of sickle cells. It comprises shortness of breath, chest pain, hypoxia, and new chest X-ray changes due to consolidation. The presentation may be gradual or very rapid, leading to death in a few hours. Initial management is with pain relief, inspired oxygen, antibiotics and exchange transfusion to reduce the amount of Hb S to <20%; occasionally ventilation may be necessary. Infections can be due to chlamydia and mycoplasma, as well as *Streptococcus pneumoniae*.

Leg ulcers occur spontaneously (vaso-occlusive episodes) or following trauma and are usually over the medial or lateral malleoli. They often become infected and are quite resistant to treatment, sometimes blood transfusion may facilitate ulcer healing.

Cardiac problems occur, with cardiomegaly, arrhythmias and iron overload cardiomyopathy. Myocardial infarctions occur due to thrombotic episodes which are not secondary to atheroma.

Neurological complications occur in 25% of patients, with transient ischaemic attacks, fits, cerebral infarction, cerebral haemorrhage and coma. Strokes occur in about 11% of patients under 20 years of age. The most common finding is obstruction of a distal intracranial internal carotid artery or a proximal middle cerebral artery. 10% of children without neurological signs or symptoms have abnormal blood-flow velocity indicative of clinically significant arterial stenosis; such patients have very high risk of stroke. It has now been demonstrated that if children with stenotic cranial artery lesions, as demonstrated on transcranial Doppler ultrasonography, are maintained on a regular programme of transfusion that is designed to suppress erythropoiesis so that no more than 30% of the circulating red cells are their own, about 90% of strokes in such children could be prevented.

Cholelithiasis. Pigment stones occur as a result of chronic haemolysis.

Liver. Chronic hepatomegaly and liver dysfunction are caused by trapping of sickle cells.

Renal. Chronic tubulo-interstitial nephritis occurs (see p. 605).

Priapism. An unwanted painful erection occurs from vaso-occlusion and can be recurrent. This may result in impotence. Treatment is with an α-adrenergic blocking drug, analgesia and hydration.

Eye. Background retinopathy, proliferative retinopathy, vitreous haemorrhages and retinal detachments all occur. Regular yearly eye checks are required.

Pregnancy. Impaired placental blood flow causes spontaneous abortion, intrauterine growth retardation, pre-eclampsia and fetal death. Painful episodes, infections and severe anaemia occur in the mother. Prophylactic transfusion does not improve fetal outcome. Oral contraceptives with low-dose oestrogens are safe.

Investigations

- **Blood count**. The level of Hb is in the range 6–8 g/dL with a high reticulocyte count (10–20%).
- **Blood films** can show features of hyposplenism (see Fig. 8.29) and sickling (Fig. 8.24).
- **Sickle solubility test**. A mixture of Hb S in a reducing solution such as sodium dithionite gives a turbid appearance because of precipitation of Hb S, whereas normal Hb gives a clear solution. A number of commercial kits such as Sickledex are available for rapid screening for the presence of Hb S, for example before surgery in appropriate ethnic groups and in the A&E department.
- **Hb electrophoresis** (Fig. 8.22) is always needed to confirm the diagnosis. There is no Hb A, 80–95% Hb SS and 2–20% Hb F.
- **The parents** of the affected child will show features of sickle cell trait.

Table 8.13 Complications requiring inpatient management

Pain – uncontrolled by non-opiate analgesia
Swollen painful joints
Acute sickle chest syndrome or pneumonia
Mesenteric sickling and bowel ischaemia
Splenic or hepatic sequestration
Central nervous system deficit
Cholecystitis (pigment stones)
Cardiac arrhythmias
Renal papillary necrosis resulting in colic or severe haematuria
Hyphema and retinal detachment
Priapism

Management

Precipitating factors (see above) should be avoided or treated quickly. The complications requiring inpatient management are shown in Table 8.13.

Acute painful attacks require supportive therapy with intravenous fluids, and adequate analgesia. Oxygen and antibiotics are only given if specifically indicated. Crises can be extremely painful and require strong, usually narcotic, analgesia. Morphine is the drug of choice. Milder pain can sometimes be relieved by codeine, paracetamol and NSAIDs (Box 8.1).

Prophylaxis is with penicillin 500 mg daily and vaccination with polyvalent pneumococcal and *Haemophilus influenzae* type B vaccine (see p. 423). Folic acid is given to all patients with haemolysis.

Anaemia

Blood transfusions should only be given for clear indications. Patients with steady state anaemia, those having minor surgery or having painful episodes without complications should not be transfused. Transfusions should be given for heart failure, TIAs, strokes, acute chest syndrome, acute splenic sequestration and aplastic crises. Before elective operations and during pregnancy, repeated transfusions may be used to reduce the proportion of circulating Hb S to less than 20% to prevent sickling. Exchange transfusions may be necessary in patients with severe or recurrent crises, or before emergency surgery. Transfusion and splenectomy may be life-saving for young children with splenic sequestration. A full blood crossmatching compatibility screen should always be performed.

Hydroxycarbamide (hydroxyurea) is the first drug which has been widely used as therapy for sickle cell anaemia. It acts, at least in part, by increasing Hb F concentrations. Hydroxycarbamide has been shown in trials to reduce the episodes of pain, the acute chest syndrome, and the need for blood transfusions.

Inhaled nitric oxide inhibits platelet function, reduces vascular adhesion of red cells and is also a vasodilator. It has been shown to reduce opiate requirements in acute painful episodes.

Bone marrow transplantation has been used to treat sickle cell anaemia although in fewer numbers than for thalassaemia. Children and adolescents younger than 16 years of age who have severe complications (strokes, recurrent chest syndrome or refractory pain) and have an HLA-matched donor are the best candidates for transplantation.

Box 8.1 Management of acute painful crisis in opioid naïve adults with sickle cell disease. Higher doses may be required for patients who have previously received opioids.

Morphine/diamorphine

- 0.1 mg/kg i.v./s.c. every 20 min until pain controlled, then
- 0.05–0.1 mg/kg i.v./s.c. (or oral morphine) every 2–4 hours
- Patient controlled analgesia (PCA) when pain controlled

Patient controlled analgesia (PCA) (example for adults >50 kg)

Diamorphine

- Continuous infusion: 0–10 mg/h
- PCA bolus dose: 2–10 mg
- Dose duration: 1 min
- Lockout time: 20–30 min

Adjuvant oral analgesia

- Paracetamol 1 g 6 hourly
- +/– Ibuprofen* 400 mg 8 hourly
- Or diclofenac* 50 mg 8 hourly

Laxatives (all patients)

For example:

- Lactulose 10 ml × 2 daily
- Senna 2–4 tablets daily
- Sodium docusate 100 mg × 2 daily
- Macrogol 1 sachet daily
- Lubiprostone

Other adjuvants

Anti-pruritics

- Hydroxyzine 25 mg × 2 as required

Antiemetics

- Prochlorperazine 5–10 mg × 3 as required
- Cyclizine 50 mg × 3 as required

Anxiolytic

- Haloperidol 1–3 mg oral/i.m. × 2 as required

*Caution advised with NSAIDs in renal impairment.
Adapted from Rees DC, Olujohungbe AD, Parker NE et al. Guidelines for the management of the acute painful crisis in sickle cell disease. *British Journal of Haematology* 2003; **120**(5): 744–752.

Counselling

A multidisciplinary team should be involved, with regular clinic appointments to build up relationships. Adolescents require careful counselling over psychosocial issues, drug and birth control.

Prognosis

Some patients with Hb SS die in the first few years of life from either infection or episodes of sequestration. However, there is marked individual variation in the severity of the disease and some patients have a relatively normal lifespan with few complications.

Sickle cell trait

These individuals have no symptoms unless extreme circumstances cause anoxia, such as flying in non-pressurized aircraft. Sickle cell trait gives some protection against *Plasmodium falciparum* malaria (see p. 154), and consequently the sickle gene has been seen as an example of a

balanced polymorphism (where the advantage of the malaria protection in the heterozygote is balanced by the mortality of the homozygous condition). Typically there is 60% Hb A and 40% Hb S. It should be emphasized that unlike thalassaemia trait, the blood count and film of a person with sickle cell trait are normal. The diagnosis is made by a positive sickle test or by Hb electrophoresis (Fig. 8.22).

Other structural globin chain defects

There are many Hb variants (e.g. Hb C, D), many of which are not associated with clinical manifestations.

Hb C ($\alpha_2\beta_2^{6glu\rightarrow lys}$) disease may be associated with Hb S (Hb SC disease). The clinical course is similar to that with Hb SS, but there is an increased likelihood of thrombosis, which may lead to life-threatening episodes of thrombosis in pregnancy, and retinopathy.

Combined defects of globin chain production and structure

Abnormalities of Hb structure (e.g. Hb S, C) can occur in combination with thalassaemia. The combination of β-thalassaemia trait and sickle cell trait (sickle cell β-thalassaemia) resembles sickle cell anaemia (Hb SS) clinically.

Hb E ($\alpha_2\beta_2^{26glu\rightarrow lys}$) is the most common Hb variant in South East Asia, and the second most prevalent haemoglobin variant world-wide. Hb E heterozygotes are asymptomatic; the haemoglobin level is normal, but red cells are microcytic. Homozygous Hb E causes a mild microcytic anaemia, but the combination of heterozygosity for Hb E and β-thalassaemia produces a variable anaemia, which can be as severe as β-thalassaemia major.

Prenatal screening and diagnosis of severe haemoglobin abnormalities

Of the offspring of parents who both have either β-thalassaemia or sickle cell trait, 25% will have β-thalassaemia major or sickle cell anaemia, respectively. Recognition of these heterozygous states in parents and family counselling provide a basis for antenatal screening and diagnosis.

Prognosis

Pregnant women with either sickle cell trait or thalassaemia trait must be identified at antenatal booking either by selective screening of high-risk groups on the basis of ethnic origin or by universal screening of all pregnant women. β-Thalassaemia trait can always be detected by a low MCV and MCH and confirmed by haemoglobin electrophoresis. However, sickle cell trait is undetectable from a blood count and the laboratory need a specific request to screen for sickle cell trait.

If a pregnant woman is found to have a haemoglobin defect, her partner should be tested. Antenatal diagnosis is offered if both are affected as there is a risk of a severe fetal Hb defect, particularly β-thalassaemia major. Fetal DNA analysis can be carried out using amniotic fluid, chorionic villus or fetal blood samples. Abortion is offered if the fetus is found to be severely affected. Chorionic villus biopsy has the advantage that it can be carried out in the first trimester, thus avoiding the need for second trimester abortions.

Gene therapy would be the ultimate corrective therapy for severe Hb abnormalities. Normal Hb genes could be inserted into the patient's haemopoietic cells in vitro and these cells could be transplanted back into the patient after ablative bone marrow treatment.

METABOLIC DISORDERS OF THE RED CELL

Red cell metabolism

The mature red cell has no nucleus, mitochondria or ribosomes and is therefore unable to synthesize proteins. Red cells have only limited enzyme systems but they maintain the viability and function of the cells. In particular, energy is required in the form of ATP for the maintenance of the flexibility of the membrane and the biconcave shape of the cells to allow passage through small vessels, and for regulation of the sodium and potassium pumps to ensure osmotic equilibrium. In addition, it is essential that Hb be maintained in the reduced state.

The enzyme systems responsible for producing energy and reducing power are (Fig. 8.25):

- the glycolytic (Embden–Meyerhof) pathway, in which glucose is metabolized to pyruvate and lactic acid with production of ATP
- the hexose monophosphate (pentosephosphate) pathway, which provides reducing power for the red cell in the form of NADPH.

About 90% of glucose is metabolized by the former and 10% by the latter. The hexose monophosphate shunt maintains glutathione (GSH) in a reduced state. Glutathione is necessary to combat oxidative stress to the red cell, and failure of this mechanism may result in:

- rigidity due to cross-linking of spectrin, which decreases membrane flexibility (see Fig. 8.18) and causes 'leakiness' of the red cell membrane
- oxidation of the Hb molecule, producing methaemoglobin and precipitation of globin chains as Heinz bodies localized on the inside of the membrane; these bodies are removed from circulating red cells by the spleen.

2,3-BPG is formed from a side-arm of the glycolytic pathway (Fig. 8.25). It binds to the central part of the Hb tetramer, fixing it in the low-affinity state (Fig. 8.4). A decreased affinity with a shift in the oxygen dissociation curve to the right enables more oxygen to be delivered to the tissues.

In addition to the G6PD, pyruvate kinase and pyrimidine 5′ nucleotidase deficiencies described below, there are a number of rare enzyme deficiencies that need specialist investigation.

Glucose-6-phosphate dehydrogenase (G6PD) deficiency

The enzyme G6PD holds a vital position in the hexose monophosphate shunt (Fig. 8.25), oxidizing glucose-6-phosphate to 6-phosphoglycerate with the reduction of NADP to NADPH. The reaction is necessary in red cells where it is the only source of NADPH, which is used via glutathione to protect the red cell from oxidative damage. G6PD deficiency is a common condition that presents with a haemolytic anaemia and affects millions of people throughout the world, particu-

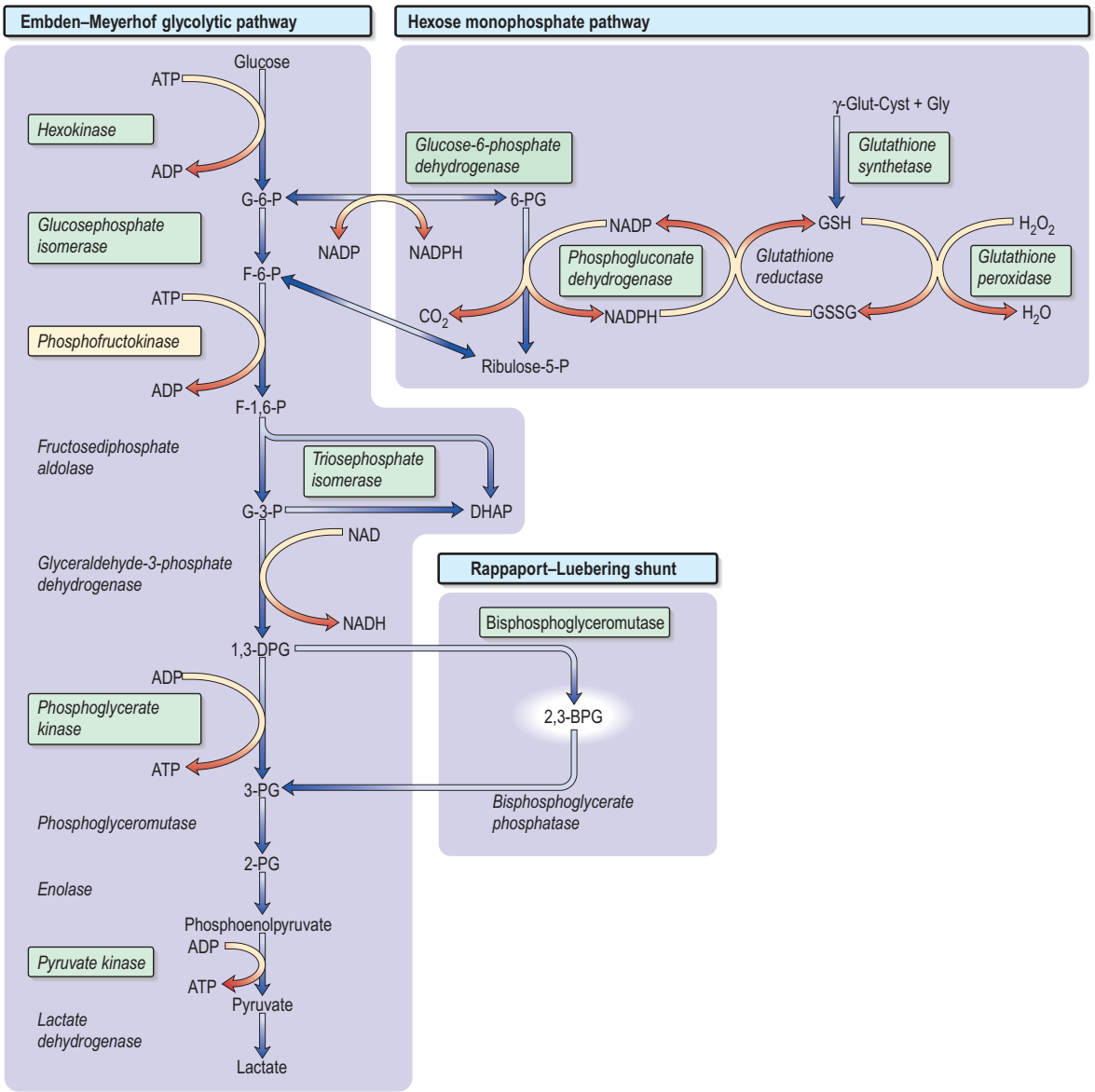


Fig. 8.25 **Glucose metabolism pathways in red cells.** The enzymes in green boxes indicate documented hereditary deficiency diseases. BPG, bisphosphoglycerate; DHAP, dihydroxyacetone phosphate; F, fructose; G, glucose; GSSG, oxidized glutathione; GSH, reduced glutathione; P, phosphate; PG, phosphoglycerate.

larly in Africa, around the Mediterranean, the Middle East (around 20%) and southeast Asia (up to 40% in some regions).

The gene for G6PD is localized to chromosome Xq28 near the factor VIII gene. The deficiency is more common in males than in females. However, female heterozygotes can also have clinical problems due to lyonization, whereby because of random X-chromosome inactivation female heterozygotes have two populations of red cells – a normal one and a G6PD-deficient one.

There are over 400 structural types of G6PD, and mutations are mostly single amino acid substitutions (missense point mutations). WHO has classified variants by the degree of enzyme deficiency and severity of haemolysis. The most common types with normal activity are called type B^+, which is present in almost all Caucasians and about 70% of black Africans, and type A^+, which is present in about 20% of black Africans. There are many variants with reduced activity but only two are common. In the African, or A^- type, the degree of deficiency is mild and more marked in older cells. Haemolysis is self-limiting as the young red cells newly produced by the bone marrow have nearly normal enzyme activity. However, in the Mediterranean type, both young and old red cells have very low enzyme activity. After an oxidant shock the Hb level may fall precipitously; death may follow unless the condition is recognized and the patient is transfused urgently.

Clinical syndromes

- Acute drug-induced haemolysis (Table 8.14) – usually dose related
- Favism (ingestion of fava beans)
- Chronic haemolytic anaemia
- Neonatal jaundice
- Infections and acute illnesses will also precipitate haemolysis in patients with G6PD deficiency.

Table 8.14	Drugs causing haemolysis in glucose-6-phosphate deficiency

Analgesics, such as:
Aspirin
Phenacetin (withdrawn in the UK)

Antimalarials, such as:
Primaquine
Pyrimethamine
Quinine
Chloroquine
Pamaquin

Antibacterials, such as:
Most sulphonamides
Dapsone
Nitrofurantoin
Chloramphenicol
Quinolones

Miscellaneous drugs, such as:
Vitamin K
Probenecid
Quinidine
Dimercaprol
Phenylhydrazine

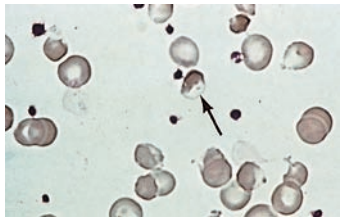


Fig. 8.26 **'Blister' cells (arrowed) in G6PD deficiency.**

Mothballs containing naphthalene can also cause haemolysis.

The clinical features are due to rapid intravascular haemolysis with symptoms of anaemia, jaundice and haemoglobinuria.

Investigations

- **Blood count** is normal between attacks.
- **During an attack** the blood film may show irregularly contracted cells, bite cells (cells with an indentation of the membrane), blister cells (cells in which the Hb appears to have become partially detached from the cell membrane; Fig. 8.26), Heinz bodies (best seen on films stained with methyl violet) and reticulocytosis.
- **Haemolysis** is evident (see p. 403).
- **G6PD deficiency** can be detected using several screening tests, such as demonstration of the decreased ability of G6PD-deficient cells to reduce dyes. The level of the enzyme may also be directly assayed. There are two diagnostic problems. Immediately after an attack the screening tests may be normal (because the oldest red cells with least 6GPD activity are destroyed selectively). The diagnosis of heterozygous females may be difficult because the enzyme level may range from very low to normal depending on lyonization. However, the risk of clinically significant haemolysis is minimal in patients with borderline G6PD activity.

Treatment

- Any offending drugs should be stopped.
- Underlying infection should be treated.
- Blood transfusion may be life-saving.
- Splenectomy is not usually helpful.

FURTHER READING

Cappellini MD, Fiorelli G. Glucose-6-phosphate dehydrogenase deficiency. *Lancet* 2008; **371:** 64–72.

Gladwin MT, Vichinsey E. Pulmonary complications of sickle cell disease. *New England Journal of Medicine* 2008; **357**: 2254–2265.

Rund D, Rachmilewitz E. Beta-thalassemia. *New England Journal of Medicine* 2005; **353**(ii): 1135–1146.

Stuart M, Nagel RL. Sickle cell disease. *Lancet* 2004; **364:** 1343–1360.

Pyruvate kinase deficiency

This is the most common defect of red cell metabolism after G6PD deficiency, affecting thousands rather than millions of people. The site of the defect is shown in Figure 8.25. There is reduced production of ATP, causing rigid red cells. Homozygotes have haemolytic anaemia and splenomegaly. It is inherited as an autosomal recessive.

Investigations

- **Anaemia** of variable severity is present (Hb 5–10 g/dL). The oxygen dissociation curve is shifted to the right as a result of the rise in intracellular 2,3-BPG, and this reduces the severity of symptoms due to anaemia.
- **Blood film** shows distorted ('prickle') cells and a reticulocytosis.
- **Pyruvate kinase activity** is low (affected homozygotes have levels of 5–20%).

Treatment

Blood transfusions may be necessary during infections and pregnancy. Splenectomy may improve the clinical condition and is usually advised for patients requiring frequent transfusions.

Pyrimidine 5′ nucleotidase deficiency

This autosomal disorder produces a haemolytic anaemia with basophilic stippling of the red cells. The enzyme degrades pyrimidine nucleotides to cytidine and uridine (pentose phosphate shunt), which in turn lead to the degradation of RNA in the reticulocytes. Lack of the enzyme results in accumulation of partially degraded RNA, which shows as basophilic stippling in mature red cells. The enzyme is also inhibited by lead (see p. 945) and thus basophilic stippling is seen in lead poisoning. The hereditary form can be diagnosed by measuring the enzyme in erythrocytes. A screening test using the ultraviolet absorption spectrum of red cells is available.

ACQUIRED HAEMOLYTIC ANAEMIA

These anaemias may be divided into those due to immune, non-immune, or other causes (Table 8.9)

Causes of immune destruction of red cells

- Autoantibodies
- Drug-induced antibodies
- Alloantibodies.

Causes of non-immune destruction of red cells

- Acquired membrane defects (e.g. paroxysmal nocturnal haemoglobinuria; see p. 418).
- Mechanical factors (e.g. prosthetic heart valves, or microangiopathic haemolytic anaemia; see p. 419)
- Secondary to systemic disease (e.g. renal and liver disease).

Table 8.15 Causes and major features of autoimmune haemolytic anaemias

	Warm	Cold
Temperature at which antibody attaches best to red cells	37°C	Lower than 37°C
Type of antibody	IgG	IgM
Direct Coombs' test	Strongly positive	Positive
Causes of primary conditions	Idiopathic	Idiopathic
Causes of secondary condition	Autoimmune disorders, e.g. systemic lupus erythematosus Chronic lymphocytic leukaemia Lymphomas Hodgkin's lymphoma Carcinomas Drugs, many including methyldopa, penicillins, cephalosporins, NSAIDs, quinine, interferon	Infections, e.g. infectious mononucleosis, *Mycoplasma pneumoniae*, other viral infections (rare) Lymphomas Paroxysmal cold haemoglobinuria (IgG)

Miscellaneous causes

- Various toxic substances can disrupt the red cell membrane and cause haemolysis (e.g. arsenic, and products of *Clostridium welchii*).
- Malaria frequently causes anaemia owing to a combination of a reduction in red cell survival and reduced production of red cells.
- Hypersplenism (see p. 423) results in a reduced red cell survival, which may also contribute to the anaemia seen in malaria.
- Extensive burns result in denaturation of red cell membrane proteins and reduced red cell survival.
- Some drugs (e.g. dapsone, sulfasalazine) cause oxidative haemolysis with Heinz bodies.
- Some ingested chemicals (e.g. weedkillers such as sodium chlorate) can cause severe oxidative haemolysis leading to acute renal failure.

AUTOIMMUNE HAEMOLYTIC ANAEMIAS

Autoimmune haemolytic anaemias (AIHA) are acquired disorders resulting from increased red cell destruction due to red cell autoantibodies. These anaemias are characterized by the presence of a positive direct antiglobulin (Coombs') test, which detects the autoantibody on the surface of the patient's red cells (Fig. 8.27).

AIHA is divided into 'warm' and 'cold' types, depending on whether the antibody attaches better to the red cells at body temperature (37°C) or at lower temperatures. The major features and the causes of these two forms of AIHA are shown in Table 8.15. In warm AIHA, IgG antibodies predominate and the direct antiglobulin test is positive with IgG alone, IgG and complement, or complement only. In cold AIHA, the antibodies are usually IgM. They easily elute off red cells, leaving complement which is detected as C3d.

Immune destruction of red cells

IgM or IgG red cell antibodies which fully activate the complement cascade cause lysis of red cells in the circulation (intravascular haemolysis).

IgG antibodies frequently do not activate complement and the coated red cells undergo extravascular haemolysis (Fig. 8.28). They are either completely phagocytosed in the spleen through an interaction with Fc receptors on

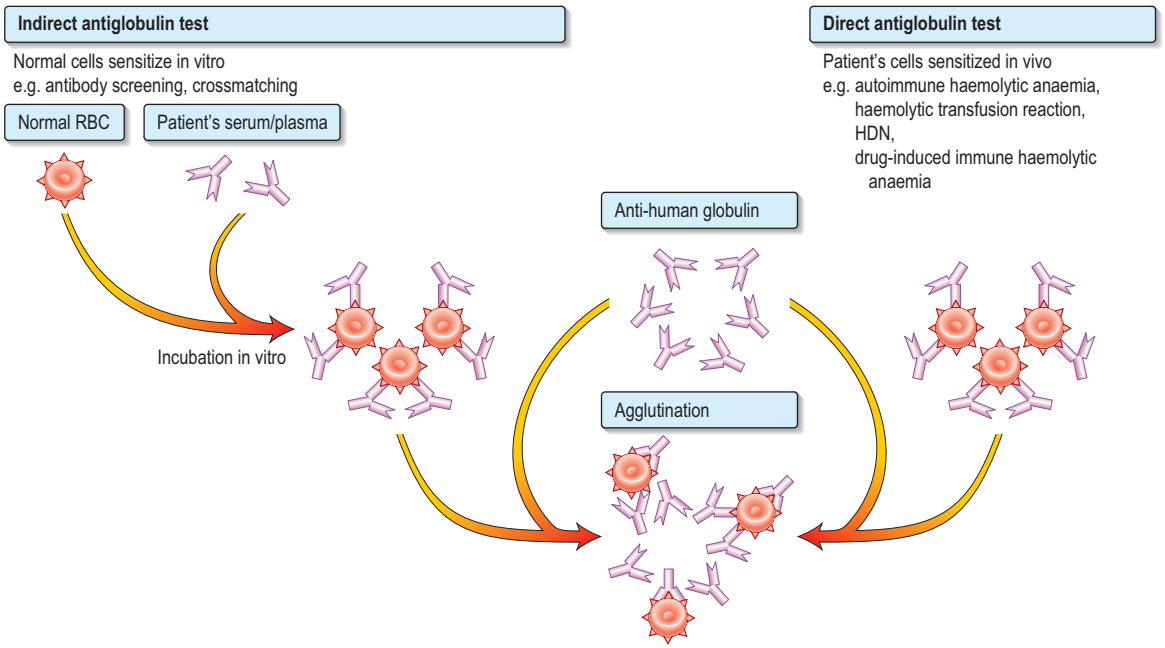


Fig. 8.27 **Antiglobulin (Coombs') tests.** The anti-human globulin forms bridges between the sensitized cells causing visible agglutination. The direct test detects patients' cells sensitized in vivo. The indirect test detects normal cells sensitized in vitro. HDN, haemolytic disease of newborn.

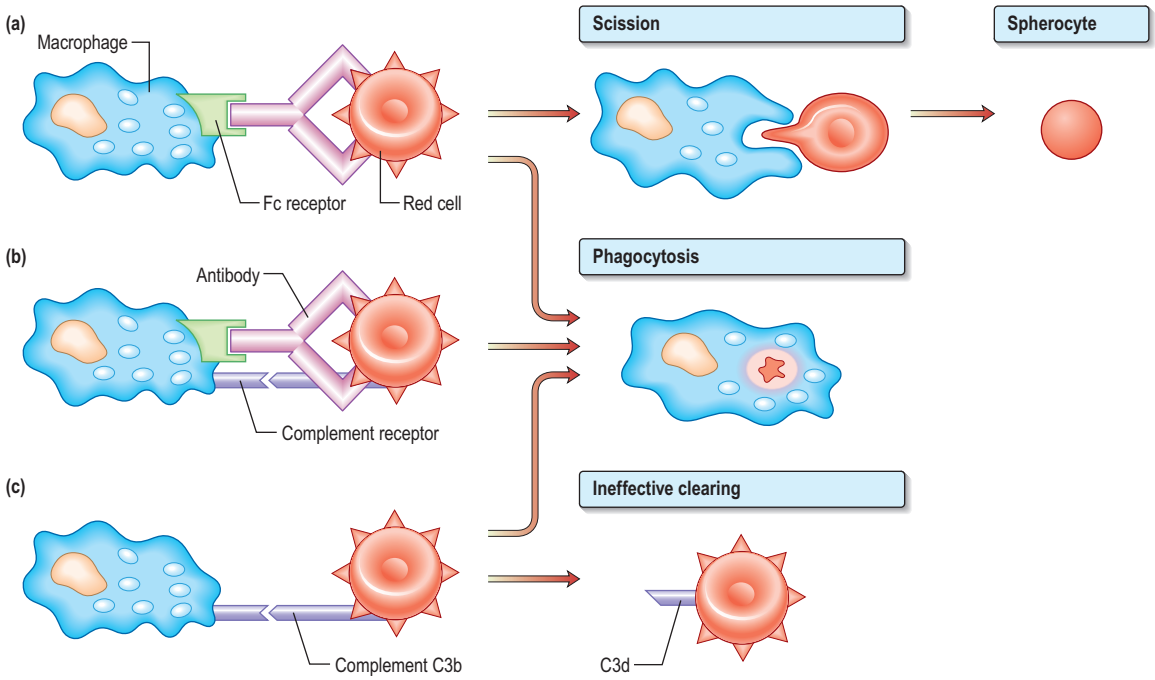


Fig. 8.28 **Extravascular haemolysis is due to interaction of antibody-coated cells with cells in the reticuloendothelial system, predominantly in the spleen. (a)** Spherocytosis results from partial phagocytosis. **(b)** Complete phagocytosis may occur and this is enhanced if there is complement as well as antibody on the cell surface. **(c)** Cells coated with complement only are ineffectively removed and circulate with C3d or C3b on their surface.

macrophages, or they lose part of the cell membrane through partial phagocytosis and circulate as spherocytes until they become sequestered in the spleen. Some IgG antibodies partially activate complement, leading to deposition of C3b on the red cell surface, and this may enhance phagocytosis as macrophages also have receptors for C3b.

Non-complement-binding IgM antibodies are rare and have little or no effect on red cell survival. IgM antibodies which partially rather than fully activate complement cause adherence of red cells to C3b receptors on macrophages, particularly in the liver, although this is an ineffective mechanism of haemolysis. Most of the red cells are released from the macrophages when C3b is cleaved to C3d and then circulate with C3d on their surface.

'Warm' autoimmune haemolytic anaemias

Clinical features

These anaemias may occur at all ages and in both sexes, although they are most frequent in middle-aged females. They can present as a short episode of anaemia and jaundice but they often remit and relapse and may progress to an intermittent chronic pattern. The spleen is often palpable. Infections or folate deficiency may provoke a profound fall in the haemoglobin level.

In more than 30% of cases, the cause remains unknown. These anaemias may be associated with lymphoid malignancies or diseases such as rheumatoid arthritis and SLE or drugs (Table 8.15).

Investigations

- **Haemolytic anaemia** is evident (see p. 403).
- **Spherocytosis** is present as a result of red cell damage.
- **Direct antiglobulin test** is positive, with either IgG alone (67%), IgG and complement (20%), or complement alone (13%) being found on the surface of the red cells.
- **Autoantibodies** may have specificity for the Rh blood group system (e.g. for the e antigen).
- **Autoimmune thrombocytopenia** and/or neutropenia may also be present (Evans' syndrome).

Treatment and prognosis

Corticosteroids (e.g. prednisolone in doses of 1 mg/kg daily) are effective in inducing a remission in about 80% of patients. Steroids reduce both production of the red cell autoantibody and destruction of antibody-coated cells. Splenectomy may be necessary if there is no response to steroids or if the remission is not maintained when the dose of prednisolone is reduced. Other immunosuppressive drugs, such as azathioprine and rituximab, may be effective in patients who fail to respond to steroids and splenectomy. Blood transfusion may be necessary if there is severe anaemia; compatibility testing is complicated by the presence of red cell autoantibodies.

'Cold' autoimmune haemolytic anaemias

Normally, low titres of IgM cold agglutinins reacting at 4°C are present in plasma and are harmless. At low temperatures these antibodies can attach to red cells and cause their agglutination in the cold peripheries of the body. In addition, activation of complement may cause intravascular haemolysis when the cells return to the higher temperatures in the core of the body.

After certain infections (such as *Mycoplasma*, cytomegalovirus, Epstein–Barr virus (EBV)) there is increased synthesis

of polyclonal cold agglutinins producing a mild to moderate transient haemolysis.

Chronic cold haemagglutinin disease (CHAD)

This usually occurs in the elderly with a gradual onset of haemolytic anaemia owing to the production of monoclonal IgM cold agglutinins. After exposure to cold, the patient develops an acrocyanosis similar to Raynaud's as a result of red cell autoagglutination.

Investigations

- **Red cells** agglutinate in the cold or at room temperature. Agglutination is sometimes seen in the sample tube after cooling but is more easily seen on the peripheral blood film made at room temperature. The agglutination is reversible after warming the sample. The agglutination may cause a spurious increase in the MCV (see p. 392).
- **Direct antiglobulin test** is positive with complement (C3d) alone.
- **Monoclonal IgM antibodies** with specificity for the Ii blood group system, usually for the I antigen but occasionally for the i antigen.

Treatment

The underlying cause should be treated, if possible. Patients should avoid exposure to cold. Steroids, alkylating agents and splenectomy are usually ineffective. Treatment with anti-CD20 (rituximab) has been successful in some cases. Blood transfusion may be necessary, and if so, the patient should be in a warm environment; compatibility testing may be difficult due to the cold agglutinin.

Paroxysmal cold haemoglobinuria (PCH)

This is a rare condition associated with common childhood infections, such as measles, mumps and chickenpox. Intravascular haemolysis is associated with polyclonal IgG complement-fixing antibodies. These antibodies are biphasic, reacting with red cells in the cold in the peripheral circulation, with lysis occurring due to complement activation when the cells return to the central circulation. The antibodies have specificity for the P red cell antigen. The lytic reaction is demonstrated in vitro by incubating the patient's red cells and serum at 4°C and then warming the mixture to 37°C (Donath–Landsteiner test). Haemolysis is self-limiting but red cell transfusions may be necessary.

DRUG-INDUCED IMMUNE HAEMOLYTIC ANAEMIA

Drug-induced haemolytic anaemias are rare, although over 100 drugs have been reported to cause immune haemolytic anaemia. The interaction between a drug and red cell membrane may produce three types of antibodies:

- *Antibodies to the drug* only, e.g. quinidine, rifampicin. Immune complexes attach to red cells, and may cause acute and severe intravascular haemolysis, sometimes associated with renal failure. The haemolysis usually resolves quickly once the drug is withdrawn.
- *Antibodies to the cell membrane* only, e.g. methyldopa, fludarabine. There is extravascular haemolysis and the clinical course tends to be more protracted.
- *Antibodies to part-drug, part-cell membrane*, e.g. penicillin. This develops only in patients receiving large doses of penicillin. The haemolysis typically develops over 7–10 days, and recovery is gradual after drug withdrawal.

Confirmation of the diagnosis requires a temporal association between administration of a drug and haemolytic anaemia, and recovery after withdrawal of the drug. In addition, the direct antiglobulin test should be positive, and drug-dependent red cell antibodies are detectable in the first and third mechanisms described above. In the second, the antibodies are not drug-dependent and are indistinguishable from autoantibodies.

ALLOIMMUNE HAEMOLYTIC ANAEMIA

Antibodies produced in one individual react with the red cells of another. This situation occurs in haemolytic disease of the newborn, haemolytic transfusion reactions (see p. 427) and after allogeneic bone marrow, renal, liver or cardiac transplantation when donor lymphocytes transferred in the allograft ('passenger lymphocytes') may produce red cell antibodies against the recipient and cause haemolytic anaemia.

Haemolytic disease of the newborn (HDN)

HDN is due to fetomaternal incompatibility for red cell antigens. Maternal alloantibodies against fetal red cell antigens pass from the maternal circulation via the placenta into the fetus, where they destroy the fetal red cells. Only IgG antibodies are capable of transplacental passage from mother to fetus.

The most common type of HDN is that due to ABO incompatibility, where the mother is usually group O and the fetus group A.

HDN due to ABO incompatibility is usually mild and exchange transfusion is rarely needed. HDN due to RhD incompatibility has become much less common in developed countries following the introduction of anti-D prophylaxis (see below). HDN may be caused by antibodies against antigens in many blood group systems (e.g. other Rh antigens such as c and E, and Kell, Duffy and Kidd; see p. 424).

Sensitization occurs as a result of passage of fetal red cells into the maternal circulation (which most readily occurs at the time of delivery), so that first pregnancies are rarely affected. However, sensitization may occur at other times, for example after a miscarriage, ectopic pregnancy or blood transfusion, or due to episodes during pregnancy which cause transplacental bleeding such as amniocentesis, chorionic villus sampling and threatened miscarriage.

Clinical features

These vary from a mild haemolytic anaemia of the newborn to intrauterine death from 18 weeks' gestation with the characteristic appearance of hydrops fetalis (hepatosplenomegaly, oedema and cardiac failure).

Kernicterus occurs owing to severe jaundice in the neonatal period, where the unconjugated (lipid-soluble) bilirubin

exceeds 250 µmol/L and bile pigment deposition occurs in the basal ganglia. This can result in permanent brain damage, choreoathetosis and spasticity. In mild cases it may present as deafness.

Investigations

Routine antenatal serology

All mothers should have their ABO and RhD groups determined and tested for atypical antibodies after attending the antenatal booking clinic. These tests should be repeated at 28 weeks' gestation.

If an antibody is detected, its blood group specificity should be determined and the mother should be retested at least monthly. A rising antibody titre of IgG antibodies or a history of HDN in a previous pregnancy is an indication for referral to a fetal medicine unit.

Antenatal assessment and treatment

If a clinically significant antibody capable of causing HDN, e.g. ant-D, anti-c or anti-K, is detected, the father's phenotype provides useful information to predict the likelihood of the fetus carrying the relevant red cell antigen. If the father is heterozygous, the genotype of the fetus can be determined from fetal DNA obtained by amniocentesis, chorionic villous sampling or fetal blood sampling. Soluble fetal DNA in maternal plasma can also be used for RhD typing, avoiding an invasive procedure.

The severity of anaemia is assessed by Doppler flow velocity of the fetal middle cerebral artery; measurement of bile pigments in the amniotic fluid is no longer routinely used. If the infant appears to have severe anaemia by non-invasive monitoring, ultrasound-guided fetal blood sampling is used to confirm this directly, and if necessary, an intravascular fetal transfusion of red cells is given.

At the birth of an affected infant

A sample of cord blood is obtained. This shows:

- anaemia with a high reticulocyte count
- a positive direct antiglobulin test
- a raised serum bilirubin.

Postnatal management

In mild cases, phototherapy may be used to convert bilirubin to water-soluble biliverdin. Biliverdin can be excreted by the kidneys and this therefore reduces the chance of kernicterus.

In more severely affected cases, exchange transfusion may be necessary to replace the infant's red cells and to remove bilirubin. Indications for exchange transfusion include:

- a cord Hb of <12 g/dL (normal cord Hb is 13.6–19.6 g/dL)
- a cord bilirubin of >60 µmol/L
- a later serum bilirubin of >300 µmol/L
- a rapidly rising serum bilirubin level.

Further exchange transfusions may be necessary to remove the unconjugated bilirubin.

The blood used for exchange transfusions should be ABO-compatible with the mother and infant, lack the antigen against which the maternal antibody is directed, be fresh (no more than 5 days from the day of collection) and be CMV-seronegative to prevent transmission of cytomegalovirus.

Prevention of RhD immunization in the mother

Anti-D should be given after delivery when all of the following are present:

- the mother is RhD negative
- the fetus is RhD positive
- there is no maternal anti-D detectable in the mother's serum; i.e. the mother is not already immunized.

The dose is 500 i.u. of IgG anti-D intramuscularly within 72 hours of delivery. The Kleihauer test is used to assess the number of fetal cells in the maternal circulation. A blood film prepared from maternal blood is treated with acid, which elutes Hb A. Hb F is resistant to this treatment and can be seen when the film is stained with eosin. If large numbers of fetal red cells are present in the maternal circulation, a higher or additional dose of anti-D will be necessary.

It may be necessary to give prophylaxis to RhD-negative women at other times when sensitization may occur, for example after an ectopic pregnancy, threatened miscarriage or termination of pregnancy. The dose of anti-D is 250 i.u. before 20 weeks' gestation and 500 i.u. after 20 weeks. A Kleihauer test should be carried out after 20 weeks to determine if more anti-D is required.

Of previously non-immunized RhD-negative women carrying RhD-positive fetuses, 1–2% became immunized by the time of delivery. Antenatal prophylaxis with administration of 500 i.u. anti-D to RhD-negative women at both 28 and 34 weeks' gestation has been shown to reduce the incidence of immunization, and its routine use is being implemented in the UK. Monoclonal anti-D could in principle replace polyclonal anti-D, which is collected from RhD-negative women immunized in pregnancy and deliberately immunized RhD-negative males, but it is likely to be some years before trials have been completed to confirm its safety and effectiveness.

FURTHER READING

Schwartz RS. Paroxysmal nocturnal hemoglobinuria. *New England Journal of Medicine* 2004; **350:** 537–538.

NON-IMMUNE HAEMOLYTIC ANAEMIA

Paroxysmal nocturnal haemoglobinuria (PNH)

Paroxysmal nocturnal haemogloblinuria is a rare form of haemolytic anaemia which results from the clonal expression of haematopoietic stems cells that have mutations in the X-linked gene *PIG-A*. These mutations result in an impaired synthesis of glycosylphosphatidylinositol (GPI), which anchors many proteins to the cell surface such as decay accelerating factor (DAF; CD55) and membrane inhibitor of reactive lysis (MIRL; CD59) to cell membranes. CD55 and CD59 and other proteins are involved in complement degradation (at the C3 and C5 levels), and in their absence the haemolytic action of complement continues.

Clinical features

The major clinical signs are intravascular haemolysis, venous thrombosis and haemoglobinuria. Haemolysis may be precipitated by infection, iron therapy or surgery. Characteristically only the urine voided at night and in the morning on waking is dark in colour, although the reason for this phenomenon is not clear. In severe cases all urine samples are dark. Urinary iron loss may be sufficient to cause iron deficiency.

Some patients present insidiously with signs of anaemia and recurrent abdominal pains.

Venous thrombotic episodes are very common in unusual places and severe thromboses may occur, for example in hepatic (Budd–Chiari syndrome), mesenteric or cerebral veins. The cause of the increased predisposition to thrombosis is not known, but may be due to complement-mediated activation of platelets deficient in CD55 and CD59. Another suggestion is that intravascular haemolysis, which releases haemoglobin in the plasma, lowers plasma nitric oxide, which causes the symptoms and venous thrombosis.

Investigations

- **Intravascular haemolysis** is evident (see p. 403).
- **Flow cytometric analysis** of red cells with anti-CD55 and anti-CD59 has replaced the Ham's test.
- **Bone marrow** is sometimes hypoplastic (or even aplastic) despite haemolysis.

Treatment and prognosis

PNH is a chronic disorder requiring supportive measures such as blood transfusions, which are necessary for patients with severe anaemia. Leucocyte-depleted blood should be used in order to prevent transfusion reactions resulting in complement activation and acceleration of the haemolysis. Treatment is with eculizumab, a recombinant humanized monoclonal antibody that prevents the cleavage of C5 (and therefore the formation of the membrane attack complex). It reduces intravascular haemolysis, haemoglobinuria, the need for transfusion and gives an improved quality of life.

Long-term anticoagulation may be necessary for patients with recurrent thrombotic episodes. In patients with bone marrow failure, treatment options include immunosuppression with antilymphocyte globulin, ciclosporin or bone marrow transplantation. Bone marrow transplantation has been successfully carried out using either HLA-matched sibling donors in patients under the age of 50 or matched unrelated donors in patients under the age of 25.

The course of PNH is variable. PNH may transform into aplastic anaemia or acute leukaemia, but it may remain stable for many years and the PNH clone may even disappear, which must be taken into account if considering potentially dangerous treatments such as bone marrow transplantation. The median survival is 10–15 years.

Gene therapy will perhaps be possible in the future.

MECHANICAL HAEMOLYTIC ANAEMIA

Red cells may be injured by physical trauma in the circulation. Direct injury may cause immediate cell lysis or be followed by resealing of the cell membrane with the formation of distorted red cells or 'fragments'. These cells may circulate for a short period before being destroyed prematurely in the reticuloendothelial system.

The causes of mechanical haemolytic anaemia include:

- damaged artificial heart valves
- march haemoglobinuria, where there is damage to red cells in the feet associated with prolonged marching or running
- microangiopathic haemolytic anaemia (MAHA), where fragmentation of red cells occurs in an abnormal microcirculation caused by malignant hypertension, eclampsia, haemolytic uraemic syndrome, thrombotic thrombocytopenic purpura, vasculitis or disseminated intravascular coagulation.

MYELOPROLIFERATIVE DISORDERS

In these disorders there is uncontrolled clonal proliferation of one or more of the cell lines in the bone marrow, namely erythroid, myeloid and megakaryocyte lines. Myeloproliferative disorders include polycythaemia vera (PV), essential thrombocythaemia (ET), myelofibrosis (all of which have a JAK-2 molecular lesion) and chronic myeloid leukaemia (CML) (a genetic BCR-ABL lesion). These disorders are grouped together as there can be transition from one disease to another; for example PV can lead to myelofibrosis. They may also transform to acute myeloblastic leukaemia. The non-leukaemic myeloproliferative disorders (PV, ET and myelofibrosis) will be discussed in this section. Chronic myeloid leukaemia is described on page 472.

POLYCYTHAEMIA

Polycythaemia (or erythrocytosis) is defined as an increase in haemoglobin, PCV and red cell count. PCV is a more reliable indicator of polycythaemia than is Hb, which may be disproportionately low in iron deficiency. Polycythaemia can be divided into absolute erythrocytosis where there is a true increase in red cell volume, or relative erythrocytosis where the red cell volume is normal but there is a decrease in the plasma volume (Fig. 8.6).

Absolute erythrocytosis is due to primary polycythaemia (PV) or secondary polycythaemia. Secondary polycythaemia is due to either an appropriate increase in red cells in response to anoxia, or an inappropriate increase associated with tumours, such as a renal carcinoma. The causes of polycythaemia are given in Table 8.16.

Primary polycythaemia: polycythaemia vera (PV)

PV is a clonal stem cell disorder in which there is an alteration in the pluripotent progenitor cell leading to excessive prolif-

Table 8.16 Causes of polycythaemia

Primary
Polycythaemia vera
Mutations in erythropoietin receptor
High-oxygen-affinity haemoglobins

Secondary

Due to an appropriate increase in erythropoietin:
High altitude
Lung disease
Cardiovascular disease (right-to-left shunt)
Heavy smoking
Increased affinity of haemoglobin, e.g. familial polycythaemia

Due to an inappropriate increase in erythropoietin:
Renal disease–renal cell carcinoma, Wilms' tumour
Hepatocellular carcinoma
Adrenal tumours
Cerebellar haemangioblastoma
Massive uterine fibroma

Relative:
Stress or spurious polycythaemia
Dehydration
Burns

eration of erythroid, myeloid and megakaryocytic progenitor cells. Over 95% of patients with PV have acquired mutations of the gene Janus Kinase 2 (JAK2). There is a V617F mutation which causes the substitution of phenylalanine for valine at position 617. JAK2 is a cytoplasmic tyrosine kinase that transduces signals, especially those triggered by haematopoietic growth factors such as erythropoietin, in normal and neoplastic cells. The significance of the discovery is twofold: first of immediate significance is the clinical utility of the detection of JAK2 mutations for the diagnosis of PV and second is the prospect of the development of new treatments for the myeloproliferative disorders based on targeting JAK2 activity.

Clinical features

The onset is insidious. It usually presents in patients aged over 60 years with tiredness, depression, vertigo, tinnitus and visual disturbance. It should be noted that these symptoms are also common in the normal population over the age of 60 and consequently PV is easily missed. These features, together with hypertension, angina, intermittent claudication and a tendency to bleed, are suggestive of PV.

Severe itching after a hot bath or when the patient is warm is common. Gout due to increased cell turnover may be a feature, and peptic ulceration occurs in a minority of patients. Thrombosis and haemorrhage are the major complications of PV.

The patient is usually plethoric and has a deep dusky cyanosis. Injection of the conjunctivae is commonly seen. The spleen is palpable in 70% and is useful in distinguishing PV from secondary polycythaemia. The liver is enlarged in 50% of patients.

Diagnosis

Box 8.2 shows the revised WHO criteria for diagnosis in adults. The measurement of red cell and plasma volume are not necessary. There may be a raised serum uric acid, leucocyte alkaline phosphatase and a raised serum vitamin B_{12} and vitamin B_{12} binding protein (transcobalamin 1).

Course and management

Treatment is designed to maintain a normal blood count and to prevent the complications of the disease, particularly thromboses and haemorrhage. Treatment is aimed at keeping the PCV below 0.45 L/L and the platelet count below 400×10^9/L. There are three types of specific treatment:

- **Venesection**. The removal of 400–500 mL weekly will successfully relieve many of the symptoms of PV. Iron deficiency limits erythropoiesis. Venesection is often used as the sole treatment and other therapy is reserved to control the thrombocytosis. The aim is to maintain a packed cell volume (PVC) of <0.45 L/L.
- **Chemotherapy**. Continuous or intermittent treatment with hydroxycarbamide (hydroxyurea) is used frequently because of the ease of controlling thrombocytosis and general safety in comparison to the alkylating agents such as busulfan, which carry an increased risk of acute leukaemia. Low-dose intermittent busulfan may be more convenient for elderly people, and this must be weighed against the potential risk of long-term complications.
- **Low dose aspirin** 100 mg daily with the above treatments is used for patients with recurrent thrombotic episodes.

Box 8.2 Modified from proposed revised WHO criteria for polycythaemia vera (PV)

Major criteria

Haemoglobin > 18.5 g/dL in men, 16.5 g/dL in women or other evidence of increased red cell volume

Presence of JAK2 tyrosine kinase V617F or other functionally similar mutation such as JAK2 exon 12 mutation

Minor criteria

Bone marrow biopsy, showing hypercellularity for age with trilineage growth (panmyelosis) with prominent erythroid, granulocytic and megakaryocytic proliferation

Serum erythropoietin level below the reference range for normal

Endogenous erythroid colony (EEC) formation in vitro*

Diagnosis requires the presence of both major criteria and one minor criterion or the presence of the first major criterion together with two minor criteria

*EEC. This is not routinely available but colony formation in the absence of exogenous erythropoietin in vitro is 100% specific and sensitive in patients without previous treatment.

This research was originally published in *Blood*. Tefferi A, Thiele J, Orazi A et al. Proposals and rationale for revision of the World Health Organization diagnostic criteria for polycythemia vera, essential thrombocythemia, and primary myelofibrosis: recommendations from an ad hoc international expert panel. *Blood* 2007; **110:** 1092–1097. © American Society of Hematology.

- **Anagrelide** inhibits megakaryocyte differentiation and is useful for thrombolysis.

General treatment

Radioactive ^{32}P is only given to patients over 70 years because of the increased risk of transformation to acute leukaemia. *Allopurinol* is given to block uric acid production. The pruritus is lessened by avoiding very hot baths. *H1-receptor antagonists* have largely proved unsuccessful in relieving distressing pruritus, but *H2-receptor antagonists* such as cimetidine are occasionally effective.

Surgery. Polycythaemia should be controlled before surgery. Patients with uncontrolled PV have a high operative risk; 75% of patients have severe haemorrhage following surgery and 30% of these patients die. In an emergency, reduction of the haematocrit by venesection and appropriate fluid replacement must be carried out.

Prognosis

PV develops into myelofibrosis in 30% of cases and into acute myeloblastic leukaemia in 5% as part of the natural history of the disease.

Secondary polycythaemias

Many high-oxygen affinity haemoglobin mutants (HOAHM) have been described which lead to increased oxygen affinity but decreased oxygen delivery to the tissues, resulting in compensatory polycythaemia. A congenital autosomal recessive disorder (Chuvasch polycythaemia) is due to a defect in the oxygen-sensing erythropoietin production pathway caused by a mutation of the von Hippel–Lindau (VHL) gene, resulting in an increased production of erythropoietin.

The causes of secondary polycythaemias are shown in Table 8.16.

Serum erythropoietin (EPO) levels are normal or raised in secondary polycythaemia. Rarely the discovery of a high EPO level may be the clue to the presence of an EPO secreting tumour.

The *treatment* is that of the precipitating factor; for example, renal or posterior fossa tumours need to be resected. The commonest cause is heavy smoking, which can produce as much as 10% carboxyhaemoglobin and this can produce polycythaemia because of a reduction in the oxygen-carrying capacity of the blood. Heavy smokers also often have respiratory disease.

Complications of secondary polycythaemia are similar to those seen in PV, including thrombosis, haemorrhage and cardiac failure, but the complications due to myeloproliferative disease such as progression to myelofibrosis or acute leukaemia do not develop. Venesection may be symptomatically helpful in the hypoxic patient, particularly if the PCV is above 0.55 L/L.

'Relative' or 'apparent' polycythaemia (Gaisböck's syndrome)

This condition was originally thought to be stress-induced. The red cell volume is normal, but as the result of a decreased plasma volume, there is a relative polycythaemia. 'Relative' polycythaemia is more common than PV and occurs in middle-aged men, particularly in smokers who are obese and hypertensive. The condition may present with cardiovascular problems such as myocardial or cerebral ischaemia. For this reason, it may be justifiable to venesect the patient. Smoking should be stopped.

Essential thrombocythaemia (ET)

Essential thrombocythaemia (ET) is a myeloproliferative disorder closely related to PV. Patients have normal Hb levels and WBC but elevated platelet counts. At diagnosis the platelet count will usually be $>600 \times 10^9$/L, and may be as high as 2000×10^9/L or rarely even higher. ET presents either symptomatically with thromboembolic or less commonly bleeding problems or incidentally (e.g. at a routine medical check).

The diagnosis of ET is not straightforward as there is no global gold standard test. The JAK2 mutation tests (see PV) are useful in that the gene is mutated in about half of all cases of ET, confirming a myeloproliferative disorder. For the remaining 50% of patient with a normal JAK2 gene, clinical assessment and observation over a period of time are required. As a generalization a person with a very high platelet count ($>1000 \times 10^9$/L) who is clinically normal with good health will most likely prove to have ET. In a patient with a lower platelet count, e.g. 600×10^9/L, and in poor health the diagnosis can be more difficult. Other disorders which may give rise to reactive high platelet counts include autoimmune rheumatic disorders and malignancy. Individuals who have been splenectomized (for any reason, including trauma) sometimes have high platelet counts.

Treatment

Treatment is with hydroxycarbamide (hydroxyurea), anagrelide or busulfan to control the platelet count to less than 400×10^9/L.

α-Interferon is also effective; it is administered by subcutaneous injection. ET may eventually transform into PV, myelofibrosis or acute leukaemia, but the disease may not progress for many years.

MYELOFIBROSIS (MYELOSCLEROSIS)

The terms myelosclerosis and myelofibrosis are interchangeable. There is clonal proliferation of stem cells and myeloid metaplasia in the liver, spleen and other organs. Increased fibrosis in the bone marrow is caused by hyperplasia of abnormal megakaryocytes which release fibroblast-stimulating factors such as platelet-derived growth factor. In about 25% of cases there is a preceding history of PV and 50% have the JAK2 mutation seen in PV.

Clinical features

The disease presents insidiously with lethargy, weakness and weight loss. Patients often complain of a 'fullness' in the upper abdomen due to splenomegaly. Severe pain related to respiration may indicate perisplenitis secondary to splenic infarction, and bone pain and attacks of gout can complicate the illness. Bruising and bleeding occur because of thrombocytopenia or abnormal platelet function. Other physical signs include anaemia, fever and massive splenomegaly (for other causes, see p. 423).

Investigations

- **Anaemia** with leucoerythroblastic features is present (see p. 430). Poikilocytes and red cells with characteristic tear-drop forms are seen. The WBC count may be over 100×10^9/L, and the differential WBC count may be very similar to that seen in chronic myeloid leukaemia (CML); later leucopenia may develop.
- **The platelet count** may be very high, but in later stages, thrombocytopenia occurs.
- **Bone marrow aspiration** is often unsuccessful and this gives a clue to the presence of the condition. A bone marrow trephine is necessary to show the markedly increased fibrosis. Increased numbers of megakaryocytes may be seen.
- **The Philadelphia chromosome** is absent; this helps to distinguish myelofibrosis from most cases of CML.
- **The leucocyte alkaline phosphatase (LAP)** score is normal or high.
- **A high serum urate** is present.
- **Low serum folate** levels may occur owing to the increased haemopoietic activity.

Differential diagnosis

The major diagnostic difficulty is the differentiation of myelofibrosis from CML as in both conditions there may be marked splenomegaly and a raised WBC count with many granulocyte precursors seen in the peripheral blood. The main distinguishing features are the appearance of the bone marrow and the absence of the Philadelphia chromosome in myelofibrosis.

Fibrosis of the marrow, often with a leucoerythroblastic anaemia, can also occur secondarily to leukaemia or lymphoma, tuberculosis or malignant infiltration with metastatic carcinoma, or to irradiation.

Treatment

This consists of general supportive measures such as blood transfusion, folic acid, analgesics and allopurinol. Drugs such as hydroxycarbamide (hydroxyurea) and busulfan are used to reduce metabolic activity and high WBC count and platelet levels. Chemotherapy and radiotherapy are used to reduce splenic size. If the spleen becomes very large and painful,

Ask the authors
www.studentconsult.com

and transfusion requirements are high, it may be advisable to perform splenectomy. Splenectomy may also result in relief of severe thrombocytopenia.

Prognosis

Patients may survive for 10 years or more; median survival is 3 years. Death may occur in 10–20% of cases from transformation to acute myeloblastic leukaemia. The most common causes of death are cardiovascular disease, infection and gastrointestinal bleeding.

MYELODYSPLASIA (MDS)

Myelodysplasia (MDS) describes a group of acquired bone marrow disorders that are due to a defect in stem cells. They are characterized by increasing bone marrow failure with quantitative and qualitative abnormalities of all three myeloid cell lines (red cells, granulocyte/monocytes and platelets). The natural history of MDS is variable, but there is a high morbidity and mortality owing to bone marrow failure, and transformation into acute myeloblastic leukaemia occurs in about 30% of cases. A classification of the myelodysplastic syndrome is shown in Table 8.17.

Clinical and laboratory features

MDS occurs mainly in the elderly, and presents with symptoms of anaemia, infection or bleeding due to pancytopenia. Serial blood counts show evidence of increasing bone marrow failure with anaemia, neutropenia, monocytosis and thrombocytopenia, either alone or in combination. By contrast, in chronic myelomonocytic leukaemias (CMML), monocytes are $>1 \times 10^9$/L and the WBC count may be $>100 \times 10^9$/L.

The bone marrow usually shows increased cellularity despite the pancytopenia. Dyserythropoiesis is present, and granulocyte precursors and megakaryocytes also have abnormal morphology. Ring sideroblasts are present in all types. In refractory anaemia with excess blasts (RAEB) and refractory anaemia with excess blasts in transformation (RAEB-t), the number of blasts in the bone marrow is increased, and the prognosis is worse than in those types with a low number of blast cells (<5%).

Table 8.17 Classification of myelodysplasia (based on WHO and FAB)

	Category	Peripheral blasts (%)	Bone marrow (%)	Median prognosis (months)
1a	Refractory anaemia (RA)	< 1	Blasts < 5 RS < 15	≈ 50
2a	RA and ring sideroblasts (RARS)	< 1	Blasts < 5 RS > 15	≈ 50
3a	RA with excess blasts (RAEB) I	1–4	Blasts 5–10	10
3b	RAEB-t II	5–19	Blasts 11–19	5

In all categories, monocytes in the peripheral blood < 1.0×10^9/L.
RS, ring sideroblasts; RAEB-t (in transformation) also includes the 5q syndrome (deletion of 5q–) – usually elderly females but with good prognosis.

Management

Patients with <5% blasts in the bone marrow are usually managed conservatively with red cell and platelet transfusions and antibiotics for infections, as they are needed. Haemopoietic growth factors (e.g. erythropoietin, G-CSF) may be useful in some patients.

Patients with >5% blasts have a less favourable prognosis, and a number of treatment options are available:

- **Supportive care** only is suitable for elderly patients with other medical problems.
- **'Gentle' chemotherapy** (low-dose or single-agent, e.g. azacytidine) may be useful in patients with high WBC counts.
- **Intensive chemotherapy** schedules used for acute myeloblastic leukaemia (see p. 469) may be tried in patients under the age of 60, but the remission rate is less, and prolonged pancytopenia may occur owing to poor haemopoietic regeneration because of the defect in stem cells.
- **Lenalidomide** (a thalidomide analogue) has been proven to be remarkably successful in the treatment of early stage myelodysplasia with a chromosome 5q deletion (the 5q– syndrome). Avoid use in women of child-bearing age.
- **Bone marrow transplantation** offers the hope of cure in the small proportion of MDS patients who are under the age of 50 and who have an HLA-identical sibling or an unrelated HLA-matched donor.

FURTHER READING

Cazzola M, Malcovati L. Myelodysplastic syndromes. *New England Journal of Medicine* 2005; **352:** 536–538.

Tefferi A. JAK2 mutations in polycythemia vera. Molecular mechanisms and clinical applications. *New England Journal of Medicine* 2007; **356:** 444–445.

THE SPLEEN

The spleen is the largest lymphoid organ in the body and is situated in the left hypochondrium. There are two anatomical components:

- the red pulp, consisting of sinuses lined by endothelial macrophages and cords (spaces)
- the white pulp, which has a structure similar to lymphoid follicles.

Blood enters via the splenic artery and is delivered to the red and white pulp. During the flow the blood is 'skimmed', with leucocytes and plasma preferentially passing to white pulp. Some red cells pass rapidly through into the venous system while others are held up in the red pulp.

Functions

Sequestration and phagocytosis. Normal red cells, which are flexible, pass through the red pulp into the venous system without difficulty. Old or abnormal cells are damaged by the hypoxia, low glucose and low pH found in the sinuses of the red pulp and are therefore removed by phagocytosis along with other circulating foreign matter. Howell–Jolly and Heinz bodies and sideroblastic granules have their particles removed by 'pitting' and are then returned to the circulation. IgG-coated red cells are removed through their Fc receptors by macrophages.

Extramedullary haemopoiesis. Pluripotential stem cells are present in the spleen and proliferate during severe haematological stress, such as in haemolytic anaemia or thalassaemia major.

Immunological function. About 25% of the body's T lymphocytes and 15% of B lymphocytes are present in the spleen. The spleen shares the function of production of antibodies with other lymphoid tissues.

Blood pooling. Up to one-third of the platelets are sequestrated in the spleen and can be rapidly mobilized. Enlarged spleens pool a significant percentage (up to 40%) of the red cell mass.

SPLENOMEGALY

Causes

A clinically palpable spleen can have many causes.

- Infection:
  (a) acute, e.g. septic shock, infective endocarditis, typhoid, infectious mononucleosis
  (b) chronic, e.g. tuberculosis and brucellosis
  (c) parasitic, e.g. malaria, kala-azar and schistosomiasis.
- Inflammation: rheumatoid arthritis, sarcoidosis, SLE.
- Haematological: haemolytic anaemia, haemoglobinopathies and the leukaemias, lymphomas and myeloproliferative disorders.
- Portal hypertension: liver disease.
- Miscellaneous: storage diseases, amyloid, primary and secondary neoplasias, tropical splenomegaly.

Massive splenomegaly. This is seen in myelofibrosis, chronic myeloid leukaemia, chronic malaria, kala-azar or, rarely, Gaucher's disease.

Hypersplenism

This can result from splenomegaly due to any cause. It is commonly seen with splenomegaly due to haematological disorders, portal hypertension, rheumatoid arthritis (Felty's syndrome) and lymphoma. Hypersplenism produces:

- pancytopenia
- haemolysis due to sequestration and destruction of red cells in the spleen
- increased plasma volume.

Treatment. This is often dependent on the underlying cause, but splenectomy is sometimes required for severe anaemia or thrombocytopenia.

Splenectomy

Splenectomy is performed mainly for:

- trauma
- immune thrombocytopenic purpura (see p. 436)
- haemolytic anaemias (see p. 408)
- hypersplenism.

Problems after splenectomy

An immediate problem is an increased platelet count (usually 600–1000 $\times$ 10^9/L) for 2–3 weeks. Thrombo-embolic phenomena may occur. In the longer term there is an increased risk of overwhelming infections, particularly pneumococcal infections.

Box 8.3 Prophylaxis against infection after splenectomy or splenic dysfunction

Vaccinate 2–3 weeks before elective splenectomy.

- A 23-valent unconjugated pneumococcal polysaccharide vaccine repeated every 5 years
- Meningococcal group C conjugate vaccine
- Annual influenza vaccine
- *Haemophilus influenzae* type b (Hib) vaccine
- Long-term penicillin V 500 mg 12-hourly (if sensitive, use erythromycin)
- Meningococcal polysaccharide vaccine (ACWY) for travellers to Africa/Saudi Arabia, e.g. during Hajj and Umrah pilgrimages

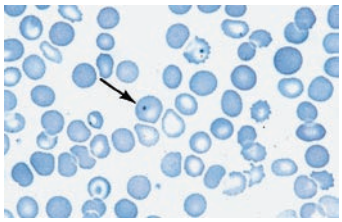


Fig. 8.29 **Postsplenectomy film with Howell–Jolly bodies (arrowed), target cells and irregularly contracted cells.**

Prophylaxis against infection after splenectomy or splenic dysfunction (Box 8.3)

All patients should be educated about the risk of infection and the importance of its early recognition and treatment. They should be given an information leaflet and should carry a card or bracelet to alert health professionals to their risk of overwhelming infection.

Postsplenectomy haematological features

- *Thrombocytosis* persists in about 30% of cases.
- *The WBC count* is usually normal but there may be a mild lymphocytosis and monocytosis.
- *Abnormalities in red cell morphology* are the most prominent changes and include Howell–Jolly bodies (contain basophilic nuclear remnants), Pappenheimer bodies (contain sideroblastic granules), target cells and irregular contracted red cells (Fig. 8.29). Pitted red cells can be counted.

Splenic atrophy

This is seen in sickle cell disease due to infarction. It is also seen in coeliac disease, in dermatitis herpetiformis, and occasionally in ulcerative colitis and essential thrombocythaemia. Postsplenectomy haematological features are seen.

FURTHER READING

Davies JM, Barnes R, Milligan D. Update of guidelines for the prevention and treatment of infection in patients with an absent or dysfunctional spleen. *Clinical Medicine* 2003; **2:** 440–443.

BLOOD TRANSFUSION

The cells and proteins in the blood express antigens which are controlled by polymorphic genes; that is, a specific antigen may be present in some individuals but not in others. A blood transfusion may immunize the recipient against donor antigens that the recipient lacks (alloimmunization), and repeated transfusions increase the risk of the occurrence

of alloimmunization. Similarly, the transplacental passage of fetal blood cells during pregnancy may alloimmunize the mother against fetal antigens inherited from the father. Antibodies stimulated by blood transfusion or pregnancy, such as Rh antibodies, are termed immune antibodies and are usually IgG, in contrast to naturally occurring antibodies, such as ABO antibodies, which are made in response to environmental antigens present in food and bacteria and which are usually IgM.

BLOOD GROUPS

The blood groups are determined by antigens on the surface of red cells; more than 280 blood groups are recognized. The ABO and Rh systems are the two major blood groups, but incompatibilities involving many other blood groups (e.g. Kell, Duffy, Kidd) may cause haemolytic transfusion reactions and/or haemolytic disease of the newborn (HDN).

ABO system

This blood group system involves naturally occurring IgM anti-A and anti-B antibodies which are capable of producing rapid and severe intravascular haemolysis of incompatible red cells.

The ABO system is under the control of a pair of allelic genes, *H* and *h*, and also three allelic genes, *A, B* and *O*, producing the genotypes and phenotypes shown in Table 8.18. The A, B and H antigens are very similar in structure; differences in the terminal sugars determine their specificity. The H gene codes for enzyme H, which attaches fucose to the basic glycoprotein backbone to form H substance, which is the precursor for A and B antigens (Fig. 8.30).

Table 8.18 The ABO system: antigens and antibodies

Phenotype	Genotype	Antigens	Antibodies	Frequency UK (%)
O	OO	None	Anti-A and anti-B	43
A	AA or AO	A	Anti-B	45
B	BB or BO	B	Anti-A	9
AB	AB	A and B	None	3

The *A* and *B* genes control specific enzymes responsible for the addition to H substance of *N*-acetylgalactosamine for Group A and d-galactose for Group B. The O gene is amorphic and does not transform H substance and therefore O is not antigenic. The A, B and H antigens are present on most body cells. These antigens are also found in soluble form in tissue fluids such as saliva and gastric juice in the 80% of the population who possess secretor genes.

Rh system

There is a high frequency of development of IgG RhD antibodies in RhD-negative individuals after exposure to RhD-positive red cells. The antibodies formed cause HDN and haemolytic transfusion reactions.

This system is coded by allelic genes, *C* and *c*, *E* and *e*, *D* and no *D*, which is signified as *d*; they are inherited as triplets on each chromosome 1, one from each pair of genes (i.e. *CDE/cde*). RhD-negative individuals have no D protein in the red cell membrane which explains why it is so immunogenic. In Caucasians, the RhD-negative phenotype almost always results from a complete deletion of the RhD gene; in black Africans, it can also result from an inactive gene containing stop codons in the reading frame.

PROCEDURE FOR BLOOD TRANSFUSION

The safety of blood transfusion depends on meticulous attention to detail at each stage leading to and during the transfusion. Avoidance of simple errors involving patient and blood sample identification at the time of collection of the sample for compatibility testing and at the time of transfusion would avoid most serious haemolytic transfusion reactions, almost all of which involve the ABO system.

Pretransfusion compatibility testing

Blood grouping

The ABO and RhD groups of the patient are determined.

Antibody screening

The patient's serum or plasma is screened for atypical antibodies that may cause a significant reduction in the survival of the transfused red cells. The patient's serum or plasma is

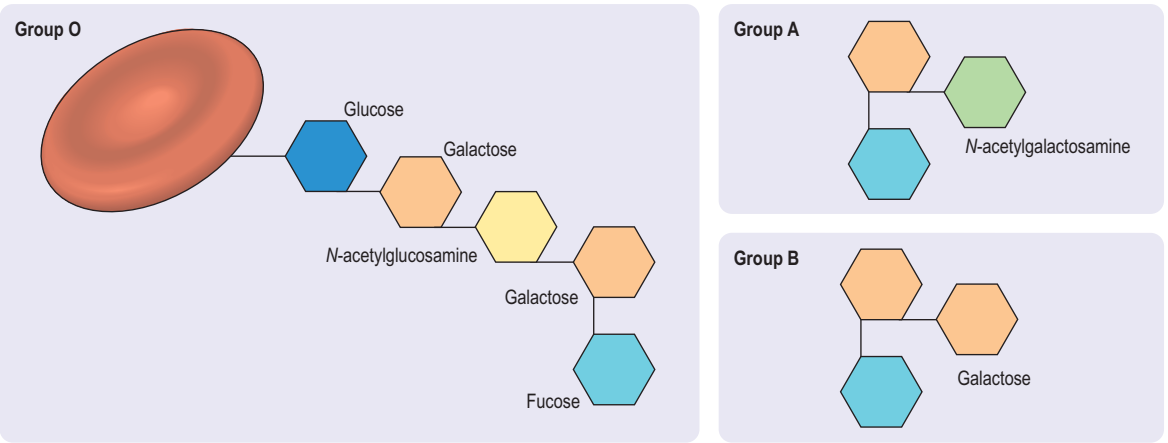


Fig. 8.30 **Sugar chains in the ABO blood group system.** Reproduced from Fricker J. Conversion of red blood cells to group O. *Lancet* 1996; **347**: 680, with permission from Elsevier.

tested against red cells from at least two group O donors, expressing a wide range of red cell antigens, for detection of IgM red cell alloantibodies (using a direct agglutination test of cells suspended in saline) and IgG antibodies (using an indirect antiglobulin test, see p. 415). About 10% of patients have a positive antibody screening result; in which case, further testing is carried out using a comprehensive panel of typed red cells to determine the blood group specificity of the antibody (clinically significant red cell antibodies are detected in about 20% of patients with positive antibody screens).

Selection of donor blood and crossmatching

Donor blood of the same ABO and RhD group as the patient is selected. Matching for additional blood groups is carried out for patients with clinically significant red cell antibodies (see below), for patients who are likely to be multitransfused and at high risk of developing antibodies, e.g. sickle cell disease, and many centres routinely provide c-negative and Kell-negative blood for women of child-bearing age to minimize the risk of alloimmunization and subsequent HDN.

Crossmatching procedures

Patients without atypical red cell antibodies. The full crossmatch involves testing the patient's serum or plasma against the donor red cells suspended in saline in a direct agglutination test, and also using an indirect antiglobulin test. In many hospitals this serological crossmatch has been omitted as the negative antibody screen makes it highly unlikely that there will be any incompatibility with the donor units. A greater risk is that of a transfusion error involving the collection of the patient sample or a mix-up of samples in the laboratory. Laboratories can use their information system to check the records of the patient and authorize the release of the donor units if a number of criteria are met (*computer or electronic crossmatching*), including:

- The system is automated for ABO and RhD grouping and antibody screening including positive sample identification and electronic transfer of results.
- The antibody screening procedure conforms to national recommendations.
- The patient's serum or plasma does not contain clinically significant red cell antibodies.
- The release of ABO incompatible blood must be prevented by conformation of laboratory computer software to the following requirements:
  (a) the issue of blood is not allowed if the patient has only been grouped once
  (b) the issue of blood is not allowed if the current group does not match the historical record
  (c) the system must not allow the reservation and release of units which are ABO incompatible with the patient.
- The laboratory must assure the validity of the ABO and RhD group of the donor blood either by written verification from the Blood Service supplying the donor units or confirmatory testing in the laboratory; the UK Blood Transfusion Services guarantee that the blood group information is correct.

Alternatively, the crossmatch can be shortened to an *'immediate spin crossmatch'* where the patient's serum or plasma is briefly incubated with the donor red cells, followed by centrifugation and examination for agglutination; this rapid crossmatch is an acceptable method of excluding ABO incompatibility in patients known to have a negative antibody screen.

Patients with atypical red cell antibodies. Donor blood should be selected that lacks the relevant red cell antigen(s), as well as being the same ABO and RhD group as the patient. A full crossmatch should always be carried out.

Several other systems for blood grouping, antibody screening and crossmatching are available to hospital transfusion laboratories. They do not depend on agglutination of red cells in suspension, but rather on the differential passage of agglutinated and unagglutinated red cells through a column of dextran gel matrix (e.g. DiaMed and Ortho Biovue systems), or on the capture of antibodies by red cells immobilized on the surface of a microplate well (e.g. Capture-R solid phase system).

Blood ordering

Elective surgery

Many hospitals have guidelines for the ordering of blood for elective surgery (maximum surgical blood ordering schedules). These are aimed at reducing unnecessary crossmatching and the amount of blood that eventually becomes outdated. Many operations in which blood is required only occasionally for unexpectedly high blood loss can be classified as 'group and save'; this means that, where the antibody screen is negative, blood is not reserved in advance but can be made available quickly if necessary, i.e. in a few minutes, using the immediate spin or electronic crossmatch procedures. If a patient has atypical antibodies, compatible blood should always be reserved in advance; this may take several days if the patient has multiple or unusual antibodies.

Emergencies

There may be insufficient time for full pretransfusion testing. The options include:

- Blood required immediately – use of 2 units of O RhD negative blood ('emergency stock'), to allow additional time for the laboratory to group the patient.
- Blood required in 10–15 minutes – use of blood of the same ABO and RhD groups as the patient (*'group compatible blood'*).
- Blood required in 45 minutes – most laboratories will be able to provide fully crossmatched blood within this time.

COMPLICATIONS OF BLOOD TRANSFUSION (Table 8.19)

In the USA, it has been mandatory to report transfusion-associated deaths to the Food and Drug Administration since 1975; such reports have provided useful data which have contributed to efforts to improve the safety of blood transfusion. Similar reporting schemes under the term '*haemovigilance'* have been set up in other countries, including the Serious Hazards of Transfusion (SHOT) scheme which produced its first report in the UK in 1997. Figure 8.31 shows the reports to SHOT up to 2005, indicating that '*incorrect blood component transfused'* is the most frequent type of

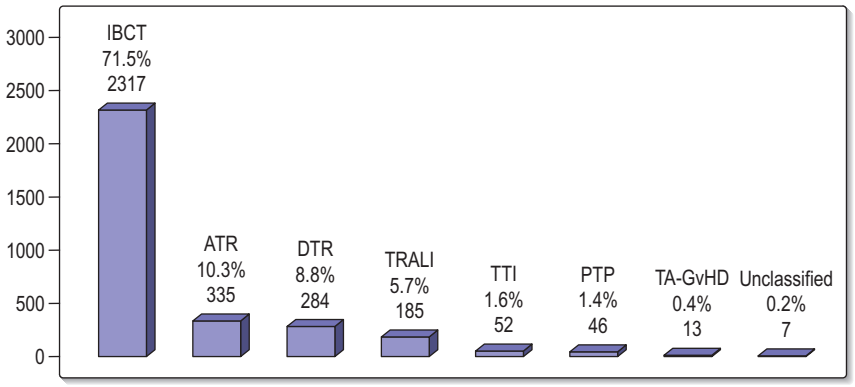


Fig. 8.31 **Overview of 3239 cases reported to the Serious Hazards of Transfusion (SHOT) scheme between 1996 and 2005.** ATR, acute transfusion reaction; DTR, delayed transfusion reaction; IBCT, incorrect blood component transfused; PTP, post-transfusion purpura; TA-GvHD, transfusion-associated graft-vs-host disease; TRALI, transfusion-related acute lung injury; TTI, transfusion-transmitted infection.

Table 8.19 Complications of blood transfusion

Immunological	Non-immunological
Alloimmunization and incompatibility	***Transmission of infection***
Red cells	Viruses
Immediate haemolytic transfusion reactions	– HAV, HBV, HCV
Delayed haemolytic transfusion reactions	– HIV, HHV8
Leucocytes and platelets	– CMV, EBV, HTLV-1, WNV
Non-haemolytic (febrile) transfusion reactions	Parasites
Post-transfusion purpura	– malaria, trypanosomiasis, toxoplasmosis
Poor survival of transfused platelets and granulocytes	Bacteria
Graft-versus-host disease	Prion – CJD
Lung injury (TRALI)	***Circulatory failure*** due to volume overload
Plasma proteins	***Iron overload*** due to multiple transfusions (see p. 408)
Urticarial and anaphylactic reactions	***Massive transfusion*** of stored blood may cause bleeding reactions and electrolyte changes
	Physical damage due to freezing or heating
	Thrombophlebitis
	Air embolism

serious incident. Errors at the time of collection of blood from the fridge and/or the administration of blood (53%) were the commonest source of error in 2005, followed by laboratory errors (43%) and mistakes in the collection of blood samples for compatibility testing (5%). Death or serious morbidity can also be attributed to other complications of blood transfusion including transfusion-associated lung injury (TRALI), transfusion-associated graft-versus-host disease (TA-GvHD) and bacterial infection of blood components.

Prevention

The serious consequences of such failures emphasize the need for meticulous checks at all stages in the procedure of blood transfusion. Written procedures and good training of staff is essential. Regular audit to ensure compliance with these procedures is also required. New approaches are also being used including the use of barcode patient identification and new technology at the bedside. Handheld devices can be used to prompt staff through the key steps and check that the barcode on the patient's wristband matches the barcode on the unit of blood (Fig. 8.32).

Fig. 8.32 The traditional method of pre-transfusion bedside checking requires two nurses and checks of multiple items of written documentation. With barcode technology, a handheld computer reads a barcode on the patient wristband containing full patient details. The handheld computer checks that the patient details on the wristband barcode match those on the barcode (in the red box) on the compatibility label attached to the unit after pre-transfusion testing. This barcode also contains the unique number of the unit, and is matched with the barcode number of the unit (top left of the bag) to ensure that the blood bank has attached the right compatibility label. This figure was originally published as the front cover of the September 2003 issue of *Transfusion* **43**. Reproduced with permission from Wiley–Blackwell.

Immunological complications

Alloimmunization

Blood transfusion carries a risk of alloimmunization to the many 'foreign' antigens present on red cells, leucocytes, platelets and plasma proteins. Alloimmunization also occurs during pregnancy – to fetal antigens inherited from the father and not shared by the mother (see p. 417).

Alloimmunization does not usually cause clinical problems with the first transfusion but these may occur with

subsequent transfusions. There may also be delayed consequences of alloimmunization, such as HDN and rejection of tissue transplants.

Incompatibility

This may result in poor survival of transfused cells, such as red cells and platelets, and also in the harmful effects of antigen–antibody reaction.

1. Red cells

Haemolytic transfusion reactions

Immediate reaction. This is the most serious complication of blood transfusion and is usually due to ABO incompatibility. There is complement activation by the antigen–antibody reaction, usually caused by IgM antibodies, leading to rigors, lumbar pain, dyspnoea, hypotension, haemoglobinuria and renal failure. The initial symptoms may occur a few minutes after starting the transfusion. Activation of coagulation also occurs and bleeding due to disseminated intravascular coagulation (DIC) is a bad prognostic sign. Emergency treatment for shock (p. 904) is needed to maintain the blood pressure and renal function.

Diagnosis

This is confirmed by finding evidence of haemolysis (e.g. haemoglobinuria), and incompatibility between donor and recipient. All documentation should be checked to detect errors such as:

- failure to check the identity of the patient when taking the sample for compatibility testing (i.e. sample from the wrong patient)
- mislabelling the blood sample with the wrong patient's name
- simple labelling or handling errors in the laboratory
- errors in the collection of blood, leading to delivery of the wrong blood to the ward/theatre
- failure to perform proper identity checks before the blood is transfused (i.e. blood transfused to the wrong patient).

Investigations

To confirm where the error occurred, blood grouping should be carried out on:

- the patient's original sample (used for the compatibility testing)
- a new sample taken from the patient after the reaction
- the donor units.

At the first suspicion of any serious transfusion reaction, the transfusion should always be stopped and the donor units returned to the blood transfusion laboratory with a new blood sample from the patient to exclude a haemolytic transfusion reaction.

Delayed reaction. This occurs in patients alloimmunized by previous transfusions or pregnancies. The antibody level is too low to be detected by pretransfusion compatibility testing, but a secondary immune response occurs after transfusion, resulting in destruction of the transfused cells, usually by IgG antibodies.

Haemolysis is usually extravascular as the antibodies are IgG, and the patient may develop anaemia and jaundice about a week after the transfusion, although most of these episodes are clinically silent. The blood film shows spherocytosis and reticulocytosis. The direct antiglobulin test is positive and detection of the antibody is usually straightforward.

2. Leucocytes and platelets

Non-haemolytic (febrile) transfusion reactions

Febrile reactions are a common complication of blood transfusion in patients who have previously been transfused or pregnant. The usual causes are the presence of leucocyte antibodies in an alloimmunized recipient acting against donor leucocytes in red cell concentrates leading to release of pyrogens, or the release of cytokines from donor leucocytes in platelet concentrates. Typical signs are flushing and tachycardia, fever (>38°C), chills and rigors. Aspirin may be used to reduce the fever, although it should not be used in patients with thrombocytopenia. The introduction of leucocyte-depleted blood in the UK, to minimize the risk of transmission of variant Creutzfeldt–Jakob disease (vCJD) by blood transfusion (see below), has reduced the incidence of febrile reactions.

Potent leucocyte antibodies in the plasma of donors, who are usually multiparous women, may cause severe pulmonary reactions (called transfusion-related acute lung injury or TRALI) characterized by dyspnoea, fever, cough, and shadowing in the perihilar and lower lung fields on the chest X-ray. Prompt respiratory support is essential; mechanical ventilation is frequently necessary. It usually resolves within 48–96 hours, but the mortality is approximately 20% in the 185 cases of TRALI reported to SHOT up to 2005.

Platelets

Post-transfusion purpura. See page 437.

3. Plasma proteins

Urticaria and anaphylaxis

Urticarial reactions are often attributed to plasma protein incompatibility, but in most cases, they are unexplained. They are common but rarely severe; stopping or slowing the transfusion and administration of chlorphenamine 10 mg i.v. are usually sufficient treatment.

Anaphylactic reactions (see p. 73) occasionally occur; severe reactions are seen in patients lacking IgA who produce anti-IgA that reacts with IgA in the transfused blood. The transfusion should be stopped and *epinephrine (adrenaline)* 0.5 mg i.m. and *chlorphenamine* 10 mg i.v. should be given immediately; endotracheal intubation may be required. Patients who have had severe urticarial or anaphylactic reactions should receive washed red cells, autologous blood or blood from IgA-deficient donors for patients with IgA deficiency.

Immunosuppression

Transfusions are known to have a favourable effect on the survival of subsequent renal allografts, due to transfusion-induced immunomodulation. Although the mechanism is unclear, the transfusion-induced immunosuppression is assumed to be due to allogeneic leucocytes Other clinical effects caused by transfusion-induced immunosuppression, such as postoperative infection and tumour recurrence, have been debated, but their clinical significance, if any, appears to be minor.

Non-immunological complications

Transmission of infection

Viral contamination: donor blood in the UK is currently tested for HBV, HCV, HIV-1 and HTLV-1. CMV-seronegative tested blood is given to immunosuppressed patients who are susceptible to acquiring CMV. Blood Services continue a vigilant search for new infectious agents such as HHV-8 which may be transmitted by blood transfusion, and for methods to prevent their transmission including donor screening, testing and pathogen inactivation. Donor questionnaires record recent travel to exclude possible risks of West Nile virus (WNV) and severe acute respiratory syndrome (SARS). WNV is the causal agent of meningoencephalitis and has been transmitted by transfusion and transplantation in the USA.

The incidence of transmission of HBV is about 1 in 400 000 units transfused, and it is less than 1 in 30 million units transfused for HCV since the introduction of testing minipools of donor plasma (48 donations per pool) for viral DNA.

In the UK the incidence of transmission of HIV by blood transfusion is extremely low – under 1 in 5 million units transfused. Prevention is based on self-exclusion of donors in 'high-risk' groups and testing each donation for anti-HIV.

There is still a potential risk of viral transmission from coagulation factor concentrates prepared from large pools of plasma. Measures for inactivating viruses such as treatment with solvents and detergents are undertaken. Viral transmission via blood transfusion is still a major issue in the developing world.

Bacterial contamination of blood components is rare but it is one of the most frequent causes of death associated with transfusion. Some organisms such as *Yersinia enterocolitica* can proliferate in red cell concentrates stored at 4°C, but platelet concentrates stored at 22°C are a more frequent cause of this problem. Systems to avoid bacterial contamination include automated culture systems and bacterial antigen detection systems, but these are not currently in routine use.

Transfusion-transmitted syphilis is very rare in the UK. Spirochaetes do not survive for more than 72 hours in blood stored at 4°C, and each donation is tested using the *Treponema pallidum haemagglutination assay* (TPHA).

There continues to be concern about the risk of transmitting the prion protein causing vCJD (see p. 121) by transfusion: four transmissions of vCJD have occurred following a blood transfusion in the UK. A number of measures have been taken in the UK, including universal leucocyte depletion of blood components (in 1999) because the prion protein is thought to be primarily associated with lymphocytes. Blood donors are excluded if they have had a blood transfusion since 1980. UK donor plasma is not used for the manufacture of blood products; imported plasma from the USA is used instead. For children under the age of 16 years, fresh frozen plasma (FFP) is sourced from plasma (from unremunerated donors) imported from the USA, on the basis that exposure to bovine spongiform encephalitis (BSE) from food was eliminated by 1 January 1996. FFP for this group is treated with methylthioninium chloride (methylene blue) to inactivate viruses.

While stringent measures are being taken to minimize the risk of transfusion-transmitted infection, it may never be possible to guarantee that donor blood is absolutely 'safe'. The current approach to the safety of blood components and plasma in the UK is extremely cautious, but it is not an absolute guarantee of safety. Clinicians should always carefully consider the patient's requirement for transfusion, and only transfuse if clinically appropriate (see below).

Circulatory failure due to volume overload – see management of acute heart failure, page 742.

Iron overload – see page 408.

Strategies for the avoidance of unnecessary transfusion

These include:

- strict criteria for the use of blood components and blood products
- stopping drug therapy (anticoagulants and antiplatelet drugs) that may potentiate bleeding in surgical patients
- the identification and treatment of anaemia prior to surgery
- the use of anti-fibrinolytic drugs, e.g. aprotinin and tranexamic acid in major surgery
- *recombinant factor VIIa* is being increasingly used for people with haemophilia with inhibitors and to treat patients with severe bleeding, e.g. postoperative, trauma, intracerebral haemorrhage. However, there is little evidence of its safety and effectiveness for this latter indication.

Artificial haemoglobin solutions and other blood substitutes continue to be evaluated. They have a short intravascular half-life and a recent meta-analysis found a significant risk of mortality as well as myocardial infarction.

Autologous transfusion

An alternative to using blood from volunteer donors is to use the patient's own blood. There are three types of autologous transfusion:

- *Predeposit*. The patient donates 2–5 units of blood at approximately weekly intervals before elective surgery.
- *Preoperative haemodilution*. 1 or 2 units of blood are removed from the patient immediately before surgery and retransfused to replace operative losses.
- *Blood salvage*. Blood lost during or after surgery may be collected and retransfused. Several techniques of varying levels of sophistication are available. The operative site must be free of bacteria, bowel contents and tumour cells.

Predeposit autologous transfusion has been abandoned in the UK except for patients where it is not possible to identify compatible blood because of multiple antibodies. There is little evidence that this approach reduces blood requirements, and because blood is perceived as being 'safe'. Blood salvage is increasingly being used as a way of avoiding the use of donor blood. In developing countries, autologous blood and blood from relatives are commonly used because of a lack of donor blood.

BLOOD, BLOOD COMPONENTS AND BLOOD PRODUCTS

Most blood collected from donors is processed as follows:

- **Blood components**, such as red cell and platelet concentrates, fresh frozen plasma (FFP) and

cryoprecipitate, are prepared from a single donation of blood by simple separation methods such as centrifugation and are transfused without further processing.
- **Blood products**, such as coagulation factor concentrates, albumin and immunoglobulin solutions, are prepared by complex processes using the plasma from many donors as the starting material (UK donor plasma is not used, see above).

In most circumstances it is preferable to transfuse only the blood component or product required by the patient ('*component therapy*') rather than use whole blood. This is the most effective way of using donor blood, which is a scarce resource, and reduces the risk of complications from transfusion of unnecessary components of the blood.

Whole blood
The average volume of blood withdrawn is 470 mL taken into 63 mL of anticoagulant. Blood stored at 4°C has a 'shelf-life' of 5 weeks when at least 70% of the transfused red cells should survive normally. Whole blood is now rarely used for transfusion; donated blood is processed into red cell concentrates and other blood components.

Red cell concentrates
Virtually all the plasma is removed and is replaced by about 100 mL of an optimal additive solution, such as SAG-M, which contains sodium chloride, adenine, glucose and mannitol. The mean volume is about 330 mL. The PCV is about 0.57 L/L, but the viscosity is low as there are no plasma proteins in the additive solution, and this allows fast administration if necessary. All blood components (red cell and platelet concentrates, and plasma) are leucocyte-depleted in the UK by filtration within 48 hours of collection of the donor blood.

Washed red cell concentrates
These are preparations of red cells suspended in saline, produced by cell separators to remove all but traces of plasma proteins. They are used in patients who have had severe recurrent urticarial or anaphylactic reactions.

Platelet concentrates
These are prepared either from whole blood by centrifugation or by plateletpheresis of single donors using cell separators. They may be stored for up to 5 days at 22°C with agitation. They are used to treat bleeding in patients with severe thrombocytopenia, and prophylactically to prevent bleeding in patients with bone marrow failure.

Granulocyte concentrates
These are prepared from single donors using cell separators and are used for patients with severe neutropenia with definite evidence of bacterial infection. The numbers of granulocytes are increased by treating donors with G-CSF and steroids.

Fresh frozen plasma
FFP is prepared by freezing the plasma from 1 unit of blood at 30°C within 6 hours of donation. The volume is approximately 200 mL. FFP contains all the coagulation factors present in fresh plasma and is used mostly for replacement of coagulation factors in acquired coagulation factor deficiencies. It may be further treated by a pathogen-inactivation process, e.g. methylene blue or solvent detergent, to minimize the risk of disease transmission. For children, see page 428.

Cryoprecipitate
This is obtained by allowing the frozen plasma from a single donation to thaw at 4–8°C and removing the supernatant. The volume is about 20 mL and it is stored at –30°C. It contains factor VIII:C, von Willebrand factor (VWF) and fibrinogen, and may be useful in DIC and other conditions where the fibrinogen level is very low. It is no longer used for the treatment of haemophilia A and von Willebrand disease because of the greater risk of virus transmission compared with virus-inactivated coagulation factor concentrates.

Factor VIII and IX concentrates
These are freeze-dried preparations of specific coagulation factors prepared from large pools of plasma. They are used for treating patients with haemophilia and von Willebrand disease, where recombinant coagulation factor concentrates are unavailable. Recombinant coagulation factor concentrates, where they are available, are the treatment of choice for patients with inherited coagulation factor deficiencies (see p. 438).

Albumin
There are two preparations:

- *Human albumin solution 4.5%,* previously called plasma protein fraction (PPF), contains 45 g/L albumin and 160 mmol/L sodium. It is available in 50, 100, 250 and 500 mL bottles.
- *Human albumin solution 20%,* previously called 'salt-poor' albumin, contains approximately 200 g/L albumin and 130 mmol/L sodium and is available in 50 and 100 mL bottles.

Human albumin solutions are generally considered to be inappropriate fluids for acute volume replacement or for the treatment of shock because they are no more effective in these situations than synthetic colloid solutions such as polygelatins (Gelofusine) or hydroxyethyl starch (Haemaccel). However, albumin solutions are indicated for treatment of acute severe hypoalbuminaemia and as the replacement fluid for plasma exchange. The 20% albumin solution is particularly useful for patients with nephrotic syndrome or liver disease who are fluid overloaded and resistant to diuretics. Albumin solutions should not be used to treat patients with malnutrition or chronic renal or liver disease with low serum albumin.

Normal immunoglobulin
This is prepared from normal plasma. It is used in patients with hypogammaglobulinaemia, to prevent infections, and in patients with, for example, immune thrombocytopenic purpura (see p. 436).

Specific immunoglobulins
These are obtained from donors with high titres of antibodies. Many preparations are available, such as anti-D, anti-hepatitis B and anti-varicella zoster.

FURTHER READING

Alter HJ, Stramer SL, Dodd RY. Emerging infectious diseases that threaten the blood supply. *Seminars in Hematology* 2007; **44:** 32–41.

Birchall J, Stanworth SJ, Duffy MR et al. Evidence for the use of recombinant factor VIIa in the prevention and treatment of bleeding in patients without hemophilia. *Transfusion Medicine Reviews* 2008; **22:** 177–187.

Goodnough LT, Shander A. Blood management. *Archives of Pathology and Laboratory Medicine* 2007; **131:** 695–701.

Murphy MF, Pamphilon D (eds). *Practical Transfusion Medicine*, 2nd edn. Oxford: Blackwell Publishing, 2005.

Natanson C, Kern SJ, Lurie P et al. Cell-free hemoglobin-based blood substitutes and risk of myocardial infarction and death: a meta-analysis. *Journal of the American Medical Association* 2008; **299:** 2304–2312.

Series on Transfusion Medicine. *Lancet* 2007; **370:** 415–448.

THE WHITE CELL (see also Ch. 3)

The five types of leucocytes found in peripheral blood are neutrophils, eosinophils and basophils (which are all called granulocytes) and lymphocytes and monocytes. The development of these cells is shown in Figure 8.1.

NEUTROPHILS

The earliest morphologically identifiable precursors of neutrophils in the bone marrow are myeloblasts, which are large cells constituting up to 3.5% of the nucleated cells in the marrow. The nucleus is large and contains 2–5 nucleoli. The cytoplasm is scanty and contains no granules. Promyelocytes are similar to myeloblasts but have some primary cytoplasmic granules, containing enzymes such as myeloperoxidase. Myelocytes are smaller cells without nucleoli but with more abundant cytoplasm and both primary and secondary granules. Indentation of the nucleus marks the change from myelocyte to metamyelocyte. The mature neutrophil is a smaller cell with a nucleus with 2–5 lobes, with predominantly secondary granules in the cytoplasm which contain lysozyme, collagenase and lactoferrin.

Peripheral blood neutrophils are equally distributed into a circulating pool and a marginating pool lying along the endothelium of blood vessels. In contrast to the prolonged maturation time of about 10 days for neutrophils in the bone marrow, their half-life in the peripheral blood is extremely short, only 6–8 hours. In response to stimuli (e.g. infection, corticosteroid therapy), neutrophils are released into the circulating pool from both the marginating pool and the marrow. Immature white cells are released from the marrow when a rapid response (within hours) occurs in acute infection (described as a 'shift to the left' on a blood film).

Function

The prime function of neutrophils is to ingest and kill bacteria, fungi and damaged cells. Neutrophils are attracted to sites of infection or inflammation by chemotaxins. Recognition of foreign or dead material is aided by coating of particles with immunoglobulin and complement (opsonization) as neutrophils have Fc and C3b receptors (see p. 56). The material is ingested into vacuoles where it is subjected to enzymic destruction, which is either oxygen-dependent with the generation of hydrogen peroxide (myeloperoxidase) or oxygen-independent (lysosomal enzymes and lactoferrin). Leucocyte alkaline phosphatase (LAP) is an enzyme found in leucocytes. It is raised when there is a neutrophilia due to an acute illness. It is also raised in polycythaemia and myelofibrosis and reduced in CML.

Neutrophil leucocytosis

A rise in the number of circulating neutrophils to $>10 \times 10^9$/L occurs in bacterial infections or as a result of tissue damage. This may also be seen in pregnancy, during exercise and after corticosteroid administration (Table 8.20). With any tissue necrosis there is a release of various soluble factors, causing a leucocytosis. Interleukin 1 is also released in tissue necrosis and causes a pyrexia. The pyrexia and leucocytosis accompanying a myocardial infarction are a good example of this and may be wrongly attributed to infection.

A leukaemoid reaction (an overproduction of white cells, with many immature cells) may occur in severe infections, tuberculosis, malignant infiltration of the bone marrow and occasionally after haemorrhage or haemolysis.

In leucoerythroblastic anaemia, nucleated red cells and white cell precursors are found in the peripheral blood. Causes include marrow infiltration with metastatic carcinoma, myelofibrosis, osteopetrosis, myeloma, lymphoma, and occasionally severe haemolytic or megaloblastic anaemia.

Neutropenia and agranulocytosis

Neutropenia is defined as a circulatory neutrophil count below 1.5×10^9/L. A virtual absence of neutrophils is called agranulocytosis. The causes are given in Table 8.21. It should

Table 8.20 Neutrophil leucocytosis

Bacterial infections
Tissue necrosis, e.g. myocardial infarction, trauma
Inflammation, e.g. gout, rheumatoid arthritis
Drugs, e.g. corticosteroids, lithium
Haematological:
Myeloproliferative disease
Leukaemoid reaction
Leucoerythroblastic anaemia
Physiological, e.g. pregnancy, exercise
Malignant disease, e.g. bronchial, breast, gastric
Metabolic, e.g. renal failure, acidosis
Congenital, e.g. leucocyte adhesion deficiency, hereditary neutrophilia

Table 8.21 Causes of neutropenia

Acquired
Viral infection
Severe bacterial infection, e.g. typhoid
Felty's syndrome
Immune neutropenia – autoimmune, autoimmune neonatal neutropenia
Pancytopenia from any cause, including drug-induced marrow aplasia (see p. 402)
Pure white cell aplasia
Inherited
Ethnic (neutropenia is common in black races)
Kostmann's syndrome (severe infantile agranulocytosis) due to mutation in elastane 2 (ELA 2) gene
Cyclical (genetic mutation in ELA2 gene with neutropenia every 2–3 weeks)
Others, e.g. Schwachman–Diamond syndrome, dyskeratosis congenita, Chédiak–Higashi syndrome

The white cell
Neutrophils
Eosinophils
Basophils
Monocytes
Lymphocytes
Haemostasis and thrombosis
Haemostasis

be noted that black patients may have somewhat lower neutrophil counts. Neutropenia caused by viruses is probably the most common type. Chemotherapy and radiotherapy predictably produce neutropenia; many other drugs have been known to produce an idiosyncratic cytopenia and a drug cause should always be considered.

Clinical features

Infections may be frequent, often serious, and are more likely as the neutrophil count falls. An absolute neutrophil count of less than 0.5×10^9/L is regarded as 'severe' neutropenia and may be associated with life-threatening infections such as pneumonia and septicaemia. A characteristic glazed mucositis occurs in the mouth, and ulceration is common.

Investigations

The blood film shows marked neutropenia. The appearance of the bone marrow will indicate whether the neutropenia is due to depressed production or increased destruction of neutrophils. Neutrophil antibody studies are performed if an immune mechanism is suspected.

Treatment

Antibiotics should be given as necessary to patients with acute severe neutropenia (see p. 461).

If the neutropenia seems likely to have been caused by a drug, all current drug therapy should be stopped. Recovery of the neutrophil count usually occurs after about 10 days. G-CSF (see p. 389) is used to decrease the period of neutropenia after chemotherapy and haemopoietic transplantation. It is also used successfully in the treatment of chronic neutropenia.

Steroids and high-dose intravenous immunoglobulin are used to treat patients with severe autoimmune neutropenia and recurrent infections, and G-CSF has produced responses in some cases.

EOSINOPHILS

Eosinophils are slightly larger than neutrophils and are characterized by a nucleus with usually two lobes and large cytoplasmic granules that stain deeply red. The eosinophil plays a part in allergic responses (see p. 56) and in the defence against infections with helminths and protozoa.

Eosinophilia is $>0.4 \times 10^9$/L eosinophils in the peripheral blood. The causes of eosinophilia are listed in Table 8.22.

Table 8.22 Causes of eosinophilia

Parasitic infestations, such as:
Ascaris
Hookworm
Strongyloides

Allergic disorders, such as:
Hayfever (allergic rhinitis)
Other hypersensitivity reactions, including drug reactions

Skin disorders, such as:
Urticaria
Pemphigus
Eczema

Pulmonary disorders, such as:
Bronchial asthma
Tropical pulmonary eosinophilia
Allergic bronchopulmonary aspergillosis
Churg–Strauss syndrome

Malignant disorders, such as:
Hodgkin's lymphoma
Carcinoma
Eosinophilic leukaemia

Miscellaneous, such as:
Hypereosinophilic syndrome
Sarcoidosis
Hypoadrenalism
Eosinophilic gastroenteritis

BASOPHILS

The nucleus of basophils is similar to that of neutrophils but the cytoplasm is filled with large black granules. The granules contain histamine, heparin and enzymes such as myeloperoxidase. The physiological role of the basophil is not known. Binding of IgE causes the cells to degranulate and release histamine and other contents involved in acute hypersensitivity reactions (see p. 437).

Basophils are usually few in number ($<1 \times 10^9$/L) but are significantly increased in myeloproliferative disorders.

MONOCYTES

Monocytes are slightly larger than neutrophils. The nucleus has a variable shape and may be round, indented or lobulated. The cytoplasm contains fewer granules than neutrophils. Monocytes are precursors of tissue macrophages and dendritic cells and spend only a few hours in the blood but can continue to proliferate in the tissues for many years.

A monocytosis ($>0.8 \times 10^9$/L) may be seen in chronic bacterial infections such as tuberculosis or infective endocarditis, chronic neutropenia and patients with myelodysplasia, particularly chronic myelomonocytic leukaemia.

LYMPHOCYTES

Lymphocytes form nearly half the circulating white cells. They descend from pluripotential stem cells (Fig. 8.1). Circulating lymphocytes are small cells, a little larger than red cells, with a dark-staining central nucleus. There are two main types: T and B lymphocytes (see p. 63).

Lymphocytosis (lymphocyte count $>5 \times 10^9$/L) occurs in response to viral infections, particularly EBV, CMV and HIV, and chronic infections such as tuberculosis and toxoplasmosis. It also occurs in chronic lymphocytic leukaemia and in some lymphomas.

HAEMOSTASIS AND THROMBOSIS

The integrity of the circulation is maintained by blood flowing through intact vessels lined by endothelial cells. Injury to the vessel wall exposes collagen and together with tissue injury sets in motion a series of events leading to haemostasis.

HAEMOSTASIS

Haemostasis is a complex process depending on interactions between the vessel wall, platelets and coagulation and fibrinolytic mechanisms. The formation of the haemostatic plug is shown in Figure 8.33.

Vessel wall

The vessel wall is lined by endothelium which, in normal conditions, prevents platelet adhesion and thrombus formation. This property is partly due to its negative charge but also to:

- thrombomodulin and heparan sulphate expression
- synthesis of prostacyclin (PGI_2) and nitric oxide (NO), which cause vasodilatation and inhibit platelet aggregation
- production of plasminogen activator.

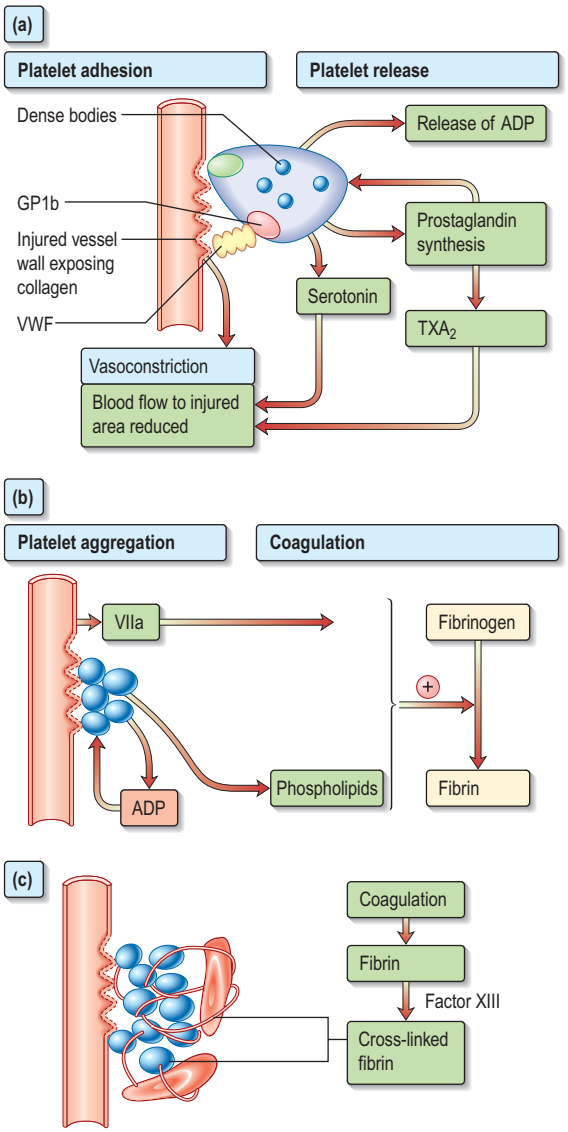


Fig. 8.33 **Formation of the haemostatic plug: sequential interactions between the vessel wall, platelets and coagulation factors.**
(a) Contact of platelets with collagen via the platelet receptor GP1b and factor VWF in plasma activates platelet prostaglandin synthesis which stimulates release of ADP from the dense bodies. Vasoconstriction of the vessel occurs as a reflex and by release of serotonin and thromboxane A_2 (TXA_2) from platelets.
(b) Release of ADP from platelets induces platelet aggregation and formation of the platelet plug. The coagulation pathway is stimulated leading to formation of fibrin.
(c) Fibrin strands are cross-linked by factor XIII and stabilize the haemostatic plug by binding platelets and red cells.

Injury to vessels causes reflex vasoconstriction, while endothelial damage results in loss of antithrombotic properties, activation of platelets and coagulation and inhibition of fibrinolysis (Fig. 8.33).

Platelets

Platelet adhesion. When the vessel wall is damaged, the platelets escaping come into contact with and adhere to collagen and von Willebrand factor that is bound below the endothelium. This is mediated through glycoprotein Ib (GPIb).

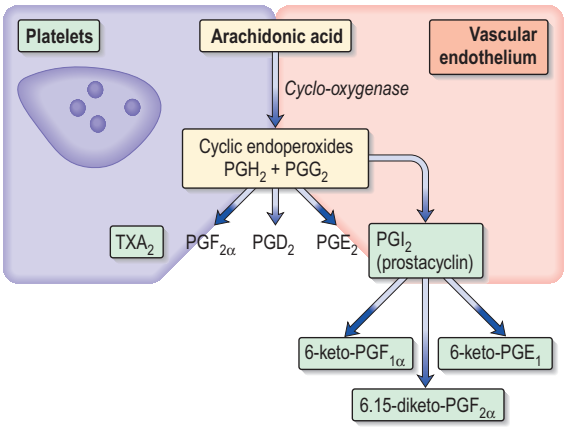


Fig. 8.34 **Prostaglandin synthesis.**

Glycoprotein IIb–IIIa is then exposed, forming a second binding site for VWF. Within seconds of adhesion to the vessel wall platelets begin to undergo a shape change, from a disc to a sphere, spread along the subendothelium and release the contents of their cytoplasmic granules, i.e. the dense bodies (containing ADP and serotonin) and the α-granules (containing platelet-derived growth factor, platelet factor 4, β-thromboglobulin, fibrinogen, VWF, fibronectin, thrombospondin and other factors).

Platelet release. The release of ADP leads to a conformational change in the fibrinogen receptor, the glycoprotein IIb–IIIa complex (GPIIb–IIIa), on the surfaces of adherent platelets allowing it to bind to fibrinogen (see also Fig. 8.41).

Platelet aggregation (Fig. 8.33b). As fibrinogen is a dimer it can form a direct bridge between platelets and so binds platelets into activated aggregates (platelet aggregation) and further platelet release of ADP occurs. A self-perpetuating cycle of events is set up leading to formation of a platelet plug at the site of the injury.

Coagulation. After platelet aggregation and release of ADP, the exposed platelet membrane phospholipids are available for the assembly of coagulation factor enzyme complexes (tenase and prothrombinase); this platelet phospholipid activity has been called platelet factor 3 (PF-3). The presence of thrombin encourages fusion of platelets, and fibrin formation reinforces the stability of the platelet plug. Central to normal platelet function is platelet *prostaglandin synthesis*, which is induced by platelet activation and leads to the formation of TXA_2 in platelets (Fig. 8.34). Thromboxane (TXA_2) is a powerful vasoconstrictor and also lowers cyclic AMP levels and initiates the platelet release reaction. Prostacyclin (PGI_2) is synthesized in vascular endothelial cells and opposes the actions of TXA_2. It produces vasodilatation and increases the level of cyclic AMP, preventing platelet aggregation on the normal vessel wall as well as limiting the extent of the initial platelet plug after injury.

Coagulation and fibrinolysis

Coagulation involves a series of enzymatic reactions leading to the conversion of soluble plasma fibrinogen to fibrin clot (Fig. 8.35). Roman numerals are used for most of the factors, but I and II are referred to as fibrinogen and prothrombin respectively; III, IV and VI are redundant. The active forms are denoted by 'a'. The coagulation factors are primarily synthesized in the liver and are either serine protease enzyme

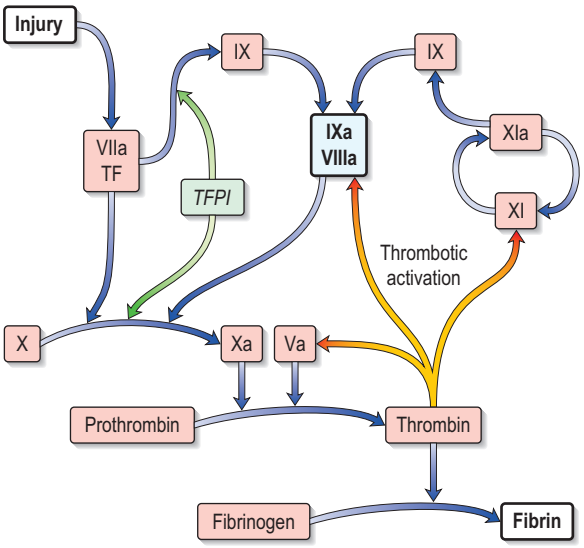


Fig. 8.35 **Coagulation mechanism.** The in vivo pathway begins with tissue factor–factor VIIa complex activating factor X and also factor IX. Factor XI is activated by thrombin. TF, tissue factor; TFPI, tissue factor pathway inhibitor.

precursors (factors XII, XI, X, IX and thrombin) or cofactors (V and VIII), except for fibrinogen, which is degraded to form fibrin.

Coagulation pathway

This enzymatic amplification system was traditionally divided into 'extrinsic' and 'intrinsic' pathways. This concept is useful for the interpretation of clinical laboratory tests such as the prothrombin time (PT) and activated partial thromboplastin time (APTT) (see p. 435) but unrepresentative and an oversimplification of in vivo coagulation. Coagulation is initiated by tissue damage. This exposes tissue factor (TF) which binds to factor VII. The TF–factor VII complex directly converts factor X to active factor Xa and some factor IX to factor IXa. In the presence of factor Xa, tissue factor pathway inhibitor (TFPI) inhibits further generation of factor Xa and factor IXa. Following inhibition by TFPI the amount of factor Xa produced is insufficient to maintain coagulation. Further factor Xa, to allow haemostasis to progress to completion, can only be generated by the alternative factor IX/factor VIII pathway. However, enough thrombin exists at this point to activate factor VIII (and factor V) and together with factor IXa (generated by TF-factor VIIa) further activation of factor X can proceed. The presence of activated factor V dramatically enhances the conversion of prothrombin to thrombin by factor Xa. Without the amplification and consolidating action of factor VIII/factor IX, bleeding will ensue as generation of factor Xa is insufficient to sustain haemostasis.

Thrombin hydrolyses the peptide bonds of fibrinogen, releasing fibrinopeptides A and B, and allowing polymerization between fibrinogen molecules to form fibrin. At the same time, thrombin, in the presence of calcium ions, activates factor XIII, which stabilizes the fibrin clot by cross-linking adjacent fibrin molecules.

Factor VIII consists of a molecule with coagulant activity (VIII:C) associated with von Willebrand factor. Factor VIII increases the activity of factor IXa by ~200000 fold. VWF functions to prevent premature factor VIII:C breakdown and

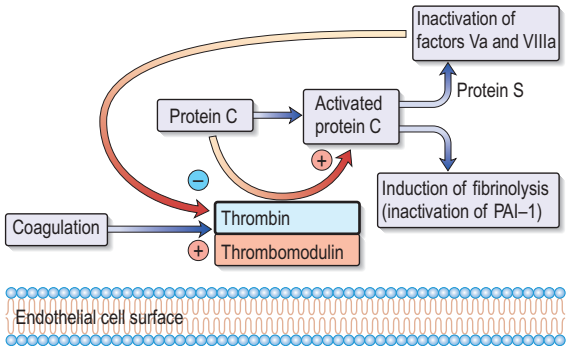


Fig. 8.36 **Activation of protein C.** PAI-1, plasminogen activator inhibitor 1.

locate it to areas of vascular injury. VIII:C has a molecular weight of about 350000.

Von Willebrand Factor (VWF) is a glycoprotein with a molecular weight of about 200000 which readily forms multimers in the circulation with molecular weights of up to 20×10^6. It is synthesized by endothelial cells and megakaryocytes and stored in platelet granules as well as the endothelial cells. The high-molecular-weight multimeric forms of VWF are the most biologically active (see p. 432 and Fig. 8.39).

Physiological limitation of coagulation

Without a physiological system to limit blood coagulation dangerous thrombosis could ensue. The natural anticoagulant mechanism regulates and localizes thrombosis to the site of injury.

Antithrombin. Antithrombin (AT), a member of the serine protease inhibitor (serpin) superfamily, is a potent inhibitor of coagulation. It inactivates the serine proteases by forming stable complexes with them, and its action is greatly potentiated by heparin.

Activated protein C. This is generated from its vitamin K-dependent precursor, Protein C, by thrombin; thrombin activation of protein C is greatly enhanced when thrombin is bound to thrombomodulin on endothelial cells (Fig. 8.36). Activated protein C inactivates factor V and factor VIII, reducing further thrombin generation.

Protein S. This is a cofactor for protein C, which acts by enhancing binding of activated protein C to the phospholipid surface. It circulates bound to C4b binding protein but some 30–40% remains unbound and active (free protein S).

Other inhibitors. Other natural inhibitors of coagulation include α_2-macroglobulin, α_1-antitrypsin and α_2-antiplasmin.

Fibrinolysis

Fibrinolysis is a normal haemostatic response that helps to restore vessel patency after vascular damage. The principal component is the enzyme plasmin, which is generated from its inactive precursor plasminogen (Fig. 8.37). This is achieved principally via tissue plasminogen activator (t-PA) released from endothelial cells. Some plasminogen activation may also be promoted by urokinase, produced in the kidneys. Other plasminogen activators (factor XII and prekallikrein) are of minor physiological importance.

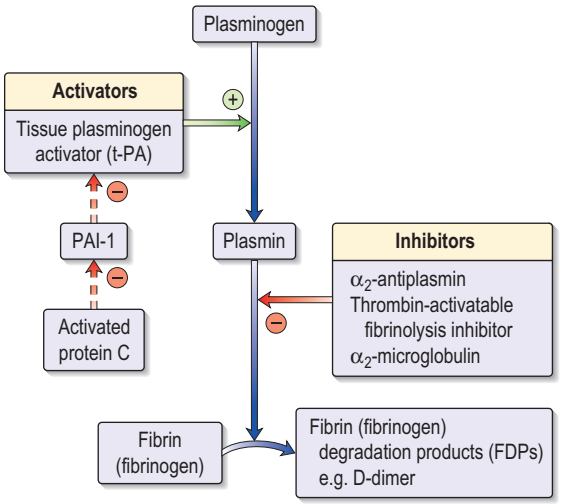


Fig. 8.37 **Fibrinolytic system.** PA-1, plasminogen activator inhibitor-1.

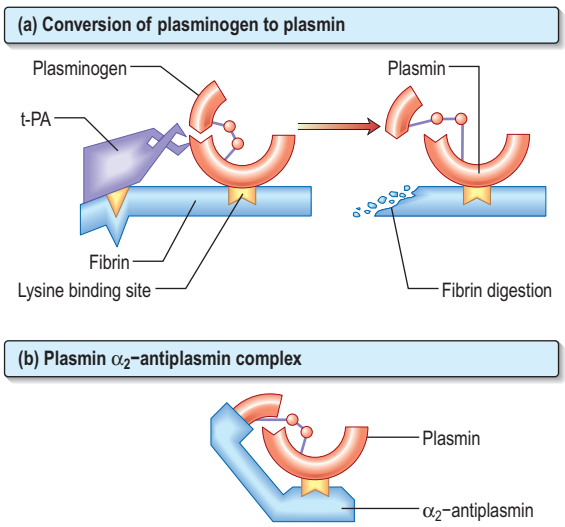


Fig. 8.38 **Fibrinolysis. (a)** The conversion of plasminogen to plasmin by plasminogen activator (t-PA) occurs most efficiently on the surface of fibrin, which has binding sites for both plasminogen and t-PA. **(b)** Free plasmin in the blood is rapidly inactivated by α_2-antiplasmin. Plasmin generated on the fibrin surface is partially protected from inactivation. The lysine-binding sites on plasminogen are necessary for the interaction between plasmin(ogen) and fibrin and between plasmin and α_2-antiplasmin.

Plasmin is a serine protease, which breaks down fibrinogen and fibrin into fragments X, Y, D and E, collectively known as fibrin (and fibrinogen) degradation products (FDPs). D-dimer is produced when cross-linked fibrin is degraded. Its presence in the plasma indicates that the coagulation mechanism has been activated.

The fibrinolytic system is activated by the presence of fibrin. Plasminogen is specifically adsorbed to fibrin and fibrinogen by lysine-binding sites. However, little plasminogen activation occurs in the absence of polymerized fibrin, as fibrin also has a specific binding site for plasminogen activators, whereas fibrinogen does not (Fig. 8.38).

t-PA is inactivated by plasminogen activator inhibitor-1 (PAI-1). Activated protein C inactivates PAI-1 and therefore induces fibrinolysis (Fig. 8.36). Inactivators of plasmin, such as α_2-antiplasmin (Fig. 8.38) and thrombin-activatable fibrinolysis inhibitor (TAFI), also contribute to the regulation of fibrinolysis.

Investigation of bleeding disorders

Although the precise diagnosis of a bleeding disorder will depend on laboratory tests, much information is obtained from the history and physical examination:

- **Is there a generalized haemostatic defect?** Supportive evidence for this includes bleeding from multiple sites, spontaneous bleeding, and excessive bleeding after injury.
- **Is the defect inherited or acquired?** A family history of a bleeding disorder should be sought. Severe inherited defects usually become apparent in infancy, while mild inherited defects may only come to attention later in life, for example with excessive bleeding after surgery, childbirth, dental extractions or trauma. Some defects are revealed by routine coagulation screens which are performed before surgical procedures.
- **Is the bleeding suggestive of a vascular/platelet defect or a coagulation defect?**

Vascular/platelet bleeding is characterized by easy bruising and spontaneous bleeding from small vessels. There is often bleeding into the skin. The term purpura includes both petechiae, which are small skin haemorrhages varying from pinpoint size to a few millimetres in diameter and which do not blanch on pressure, and ecchymoses, which are larger areas of bleeding into the skin. Bleeding also occurs from mucous membranes especially the nose and mouth. *Coagulation disorders* are typically associated with bleeding after injury or surgery, and in more severe forms, haemarthroses and muscle haematomas. There is often a short delay between the precipitating event and overt haemorrhage or haematoma formation.

Laboratory investigations

- **Blood count and film** show the number and morphology of platelets and any blood disorder such as leukaemia or lymphoma. The normal range for the platelet count is 150–400 $\times$ 10^9/L.
- **Bleeding time** measures platelet plug formation in vivo. It is determined by applying a sphygmomanometer cuff to the arm and inflating it to 40 mmHg. Two 1 mm deep, 1 cm long incisions are made in the forearm with a template. Each wound is blotted every 30 seconds and the time taken for bleeding to stop is recorded, normally between 3 and 10 minutes. Prolonged bleeding times are found in patients with platelet function defects, and there is a progressive prolongation with platelet counts less than 100 $\times$ 10^9/L. The bleeding time should not be performed at low platelet counts.
- **Coagulation tests** are performed using blood collected into citrate, which neutralizes calcium ions and prevents clotting.

The *prothrombin time* (PT) (also see p. 445) is measured by adding tissue factor (thromboplastin) and calcium to the patient's plasma. The normal PT is 12–16 seconds and may be expressed as the international normalized ratio, INR (see p. 324) The PT measures VII, X, V, prothrombin and fibrinogen (classic 'extrinsic' pathway) and is prolonged with abnor-

malities of these factors. It may also be abnormal in liver disease, or if the patient is on warfarin.

The *activated partial thromboplastin time* (APTT) is also sometimes known as the PTT with kaolin (PTTK). It is performed by adding a surface activator (such as kaolin), phospholipid (to mimic platelet membrane) and calcium to the patient's plasma. The normal APTT is 30–50 seconds and depends on the exact methodology. The APTT measures XII, XI, IX, VIII, X, V, prothrombin and fibrinogen (classic 'intrinsic' pathway) and is prolonged with deficiencies of one or more of these factors. It is not dependent on factor VII.

The *thrombin time* (TT) is performed by adding thrombin to the patient's plasma. The normal TT is 12–14 seconds, and it is prolonged with fibrinogen deficiency, qualitative defects of fibrinogen (dysfibrinogenaemia) or inhibitors such as heparin or FDPs.

Correction tests can be used to differentiate prolonged times in the PT, APTT and TT due to various coagulation factor deficiencies and inhibitors of coagulation. Prolonged PT, APTT or TT due to coagulation factor deficiencies can be corrected by addition of normal plasma to the patient's plasma. Failure to correct after addition of normal plasma is suggestive of the presence of an inhibitor of coagulation.

Factor assays are used to confirm coagulation defects, especially where a single inherited disorder is suspected.

Special tests of coagulation will often be required to confirm the precise haemostatic defect. Such tests include estimation of fibrinogen and FDPs, platelet function tests such as platelet aggregation and platelet granule contents.

VASCULAR DISORDERS

The vascular disorders (Table 8.23) are characterized by easy bruising and bleeding into the skin. Bleeding from mucous membranes sometimes occurs but the bleeding is rarely severe. Laboratory investigations including the bleeding time are normal. The vascular disorders include the following.

Hereditary haemorrhagic telangiectasia is a rare disorder with autosomal dominant inheritance. Dilatation of capillaries and small arterioles produces characteristic small red spots that blanch on pressure in the skin and mucous membranes, particularly the nose and gastrointestinal tract. Recurrent epistaxis and chronic gastrointestinal bleeding are the major problems which causes chronic iron deficiency anaemia.

Easy bruising syndrome is a common benign disorder occurring in otherwise healthy women. It is characterized by bruises on the arms, legs and trunk with minor trauma, possibly because of skin vessel fragility. It may give rise to the suspicion of a serious bleeding disorder.

Senile purpura and purpura due to steroids are both due to atrophy of the vascular supporting tissue.

Purpura due to infections is mainly caused by damage to the vascular endothelium. The rash of meningococcal septicaemia is particularly characteristic (see p. 136).

Henoch–Schönlein purpura (see p. 592) occurs mainly in children. It is a type III hypersensitivity (immune complex) reaction that is often preceded by an acute upper respiratory tract infection. Purpura is mainly seen on the legs and buttocks. Abdominal pain, arthritis, haematuria and glomerulonephritis also occur. Recovery is usually spontaneous, but some patients develop renal failure.

Episodes of inexplicable bleeding or bruising may represent abuse, either self-inflicted or caused by others. These various forms of artificial or factitious purpura are expressions of severe emotional or psychiatric disturbances.

Table 8.23 Vascular disorders

Congenital
Hereditary haemorrhagic telangiectasia (Osler–Weber–Rendu disease)
Connective tissue disorders (Ehlers–Danlos syndrome, osteogenesis imperfecta, pseudoxanthoma elasticum, Marfan's syndrome)

Acquired
Severe infections:
- Septicaemia
- Meningococcal infections
- Measles
- Typhoid

Allergic
Henoch–Schönlein purpura
Autoimmune disorders (SLE, rheumatoid arthritis)

Drugs
Steroids
Sulphonamides

Others
Senile purpura
Easy bruising syndrome
Scurvy
Factitious purpura

PLATELET DISORDERS

Bleeding due to thrombocytopenia or abnormal platelet function is characterized by purpura and bleeding from mucous membranes. Bleeding is uncommon with platelet counts above 50×10^9/L, and severe spontaneous bleeding is unusual with platelet counts above 20×10^9/L (Table 8.24).

Thrombocytopenia

This is caused by reduced platelet production in the bone marrow, excessive peripheral destruction of platelets or sequestration in an enlarged spleen (Table 8.25). The underlying cause may be revealed by history and examination but a bone marrow examination will show whether the numbers of megakaryocytes are reduced, normal or increased, and will provide essential information on morphology. Specific laboratory tests may be useful to confirm the presence of such conditions as paroxysmal nocturnal haemoglobinuria (PNH) or systemic lupus erythematosus (SLE).

In patients with thrombocytopenia due to failure of production, no specific treatment may be necessary but the underlying condition should be treated if possible. Where the platelet count is very low or the risk of bleeding is very high, then platelet transfusion is indicated.

Table 8.24 Clinical effects caused by different levels of platelet count

Platelet count ($\times 10^9$/L)	Clinical defect
>500	Haemorrhage or thrombosis
500–100	No clinical effect
100–50	Moderate haemorrhage after injury
50–20	Purpura may occur Haemorrhage after injury
<20	Purpura common Spontaneous haemorrhage from mucous membranes Intracranial haemorrhage (rare)

From Colvin BT. *Medicine* 2004; **32**(5): 27–33 with permission from Elsevier.

Ask the authors
www.studentconsult.com

Table 8.25 Causes of thrombocytopenia

Impaired production
Selective megakaryocyte depression:
- Rare congenital defects
- Drugs, chemicals and viruses

As part of a general bone marrow failure:
- Cytotoxic drugs and chemicals
- Radiation

Megaloblastic anaemia
Leukaemia
Myelodysplastic syndromes
Myeloma
Myelofibrosis
Solid tumour infiltration
Aplastic anaemia
HIV infection

Excessive destruction or increased consumption
Immune
- Autoimmune – ITP
- Drug induced, e.g. GP IIb/IIIa inhibitors, pencillins, thiazides
- Secondary immune (SLE, CLL, viruses, drugs, e.g. heparin, bivalirudin)
- Alloimmune neonatal thrombocytopenia
- Post-transfusion purpura

Disseminated intravascular coagulation
Thrombotic thrombocytopenic purpura

Sequestration
Splenomegaly
Hypersplenism

Dilutional
Massive transfusion

Immune thrombocytopenic purpura (ITP)

Thrombocytopenia is due to immune destruction of platelets. The antibody-coated platelets are removed following binding to Fc receptors on macrophages.

ITP in children

This occurs most commonly in children (age 2–6 years). It has an acute onset with mucocutaneous bleeding and there may be a history of a recent viral infection, including varicella zoster or measles. Although bleeding may be severe, life threatening hemorrhage is rare (~ 1%). Bone marrow examination is not usually performed unless treatment becomes necessary on clinical grounds.

ITP in adults

The presentation is usually less acute than in children. Adult ITP is characteristically seen in women and may be associated with other autoimmune disorders such as SLE, thyroid disease and autoimmune haemolytic anaemia (Evans' syndrome), in patients with chronic lymphocytic leukaemia and solid tumours, and after infections with viruses such as HIV. Platelet autoantibodies are detected in about 60–70% of patients, and are presumed to be present, although not detectable, in the remaining patients; the antibodies often have specificity for platelet membrane glycoproteins IIb/IIIa and/or Ib.

Clinical features

Major haemorrhage is rare and is seen only in patients with severe thrombocytopenia. Easy bruising, purpura, epistaxis and menorrhagia are common. Physical examination is normal except for evidence of bleeding. Splenomegaly is rare.

Investigation

The only blood count abnormality is thrombocytopenia. Normal or increased numbers of megakaryocytes are found in the bone marrow (if examination is performed), which is otherwise normal. The detection of platelet autoantibodies is not essential for confirmation of the diagnosis, which often depends on exclusion of other causes of excessive destruction of platelets.

Treatment

Children

Children do not usually require treatment. Where this is necessary on clinical grounds, high-dose prednisolone is effective, given for a very short course. Intravenous immunoglobulin (i.v. IgG) should be reserved for very serious bleeding or urgent surgery. Chronic ITP is rare and requires specialist management.

Adults

Patients with platelet counts greater than 30×10^9/L require no urgent treatment unless they are about to undergo a surgical procedure.

First-line therapy consists of oral corticosteroids 1 mg/kg body weight. Approximately 66% will respond to prednisolone but relapse is common when the dose is reduced. Only 33% of patients can expect a long-term response and long-term remission is seen in only 10–20% of patients following stopping prednisolone. Patients who fail to respond to corticosteroids or require high doses to maintain a safe platelet count should be considered for splenectomy.

Intravenous immunoglobulin (i.v. IgG) is effective. It raises platelet count in 75% and in 50% the platelet count will normalize. Responses are only transient (3–4 weeks) with little evidence of any lasting effect. However, it is very useful where a rapid rise in platelet count is desired, especially before surgery. There are also advocates for high dose corticosteroids for additional therapy.

Second-line therapy involves splenectomy, to which the majority of patients respond – two-thirds will achieve a normal platelet count. Patients who do not have a complete response can still expect some improvement.

Third-line therapy. For those that fail splenectomy, a wide range of other therapies are available. These include high-dose corticosteroids, intravenous immunoglobulin, Rh0(D) immune globulin (anti-D), vinca alkaloids, danazol, immunosuppressive agents such as azathioprine, ciclosporin and dapsone, combination chemotherapy, mycophenolate mofetil. Major difficulties with many third-line therapies are modest response rates and slow onset of action. Consequently there is also interest in the use of specific monoclonal antibodies such as rituximab, as well as recombinant thrombopoietin. However, clinical trials of thrombopoietin were stopped because of thrombocytopenia but eltrombopag, a thrombopoietin receptor agonist (which binds to another point of the thrombopoietin receptor), has been shown to increase platelets in ITP. Romiplostim, a novel thrombopoiesis protein given weekly subcutaneously, has also been shown to significantly increase platelet count in ITP on a long-term basis; there were no major adverse effects in this trial. Platelet transfusions are reserved for intracranial or other extreme haemorrhage, where emergency splenectomy may be justified.

Other immune thrombocytopenias

Drugs cause immune thrombocytopenia by the same mechanisms as described for drug-induced immune haemolytic anaemia (see p. 417). The same drugs can be responsible

for immune haemolytic anaemia, thrombocytopenia or neutropenia in different patients.

Heparin-induced thrombocytopenia. See page 444.

Neonatal alloimmune thrombocytopenia is due to fetomaternal incompatibility for platelet-specific antigens, usually for HPA-1a (human platelet alloantigen), and is the platelet equivalent of haemolytic disease of the newborn (HDN). The mother is HPA-1a-negative and produces antibodies, which destroy the HPA-1a-positive fetal platelets. Thrombocytopenia is self-limiting after delivery, but platelet transfusions may be required initially to prevent or treat bleeding associated with severe thrombocytopenia; platelets are prepared from HPA-1a-negative volunteers or the mother herself. Severe bleeding such as intracranial haemorrhage may also occur in utero.

Antenatal treatment of the mother – usually with high dose IgG and/or steroids – has been effective in preventing haemorrhage in severely affected cases.

Post-transfusion purpura (PTP) is rare, occurring 2–12 days after a blood transfusion. PTP is associated with a platelet-specific alloantibody, usually anti-HPA-1a in an HPA-1a-negative individual. PTP almost invariably occurs in females who have been previously immunized by pregnancy or blood transfusion. The cause of the destruction of the patient's own platelets is not well understood, but they may be destroyed as 'bystanders' during the acute immune response to HPA-1a. PTP is self-limiting, but high-dose intravenous IgG may limit the period of thrombocytopenia.

Thrombotic thrombocytopenic purpura (TTP) *(see p. 599)*

TTP is a rare, but very serious condition, in which platelet consumption leads to profound thrombocytopenia. There is a characteristic symptom complex of florid purpura, fever, fluctuating cerebral dysfunction and haemolytic anaemia with red cell fragmentation, often accompanied by renal failure. The coagulation screen is usually normal but lactic dehydrogenase (LDH) levels are markedly raised as a result of haemolysis. TTP arises due to endothelial damage and microvascular thrombosis. This occurs due to a reduction in ADAMTS 13 (A Disintegrin-like and Metalloproteinase domain with Thrombospondin-type motifs), a protease which is normally responsible for VWF degradation. ADAMTS 13 is needed to break down ultra large von Willebrand factor multimers (UL VWFMs) into smaller haemostatically active fragments that interact with platelets. Reduction in ADAMTS 13 results in the adhesion and aggregation of platelets to UL VWFMs and multiorgan microthrombi. In most sporadic cases there is a true deficiency of the ADAMTS 13, associated with antibodies to ADAMTS 13. In some congenital cases the deficiency is due to mutations in the ADAMTS 13 gene. Secondary causes of acute TTP include pregnancy, oral contraceptives, SLE, infection and drug treatment, including the use of ticlopidine and clopidogrel. Such cases may have a variable ADAMTS 13 activity at presentation and may or may not have associated antibodies to ADAMTS 13.

Treatment consists of plasma exchange as the mainstay of treatment. It provides a source of ADAMTS 13 and removes associated autoantibody in acute TTP. Cryoprecipitate and Solvent-Detergent FFP (fresh frozen plasma) both contain ADAMTS 13. Pulsed intravenous methylprednisolone is given acutely, as is increasingly rituximab as a primary treatment of choice. Disease activity is monitored by measuring the platelet count and serum LDH. Platelet concentrates are contraindicated. The untreated condition has a mortality of up to 90% but modern management has reduced this figure to about 10%. Recurrent and relapsing TTP occurs, often associated with a persistent lack of ADAMTS 13. In secondary TTP cases, identifiable precipitating drugs should be stopped.

More online
www.studentconsult.com

Platelet function disorders

These are usually associated with excessive bruising and bleeding and, in some of the acquired forms, with thrombosis. The platelet count is normal or increased and the bleeding time is prolonged. The rare inherited defects of platelet function require more detailed investigations such as platelet aggregation studies and factor VIII:C and VWF assays, if von Willebrand disease is suspected.

Inherited types of platelet dysfunction

- *Glanzmann's thrombasthenia* – lack of the platelet membrane glycoprotein IIb–IIIa complex resulting in defective fibrinogen binding and failure of platelet aggregation.
- *Bernard–Soulier syndrome* – lack of platelet membrane glycoprotein Ib, the binding site for VWF. This causes a failure of platelet adhesion and moderate thrombocytopenia.
- *Storage pool disease* – lack of the storage pool of platelet dense bodies, causing poor platelet function.

Acquired types of platelet dysfunction

- Myeloproliferative disorders
- Renal and liver disease
- Paraproteinaemias
- Drug-induced, such as NSAIDs (aspirin) or other platelet inhibitory drugs.

If there is serious bleeding or if the patient is about to undergo surgery, drugs with antiplatelet activity should be withdrawn and any underlying condition should be corrected if possible.

Bleeding in renal disease is multifactorial, although platelet dysfunction is a major component. The degree of the defect of haemostasis is broadly proportional to the plasma urea concentration – platelet function is impaired by urea, guanidinosuccinic acid and other phenolic metabolites that accumulate in renal failure. Dialysis partially corrects platelet function. The haematocrit should be increased to greater than 0.30 and the use of desmopressin may be helpful. Platelet transfusions may be required if these measures are unsuccessful or if the risk of bleeding is high.

Thrombocytosis

The platelet count may rise above 400 $\times$ 10^9/L as a result of:

- splenectomy
- malignant disease
- inflammatory disorders such as rheumatoid arthritis and inflammatory bowel disease
- major surgery and post haemorrhage
- myeloproliferative disorders
- iron deficiency.

Thus thrombocytosis is part of the acute-phase reaction, although following splenectomy platelet numbers are also elevated because of the loss of a major site of platelet destruction.

Essential thrombocythaemia, a myeloproliferative disorder which is described on page 421, and other myeloproliferative conditions such as polycythaemia vera (PV), myelofibrosis and chronic myeloid leukaemia (CML) may also be associated with a high platelet count.

A persistently elevated platelet count can lead to arterial or venous thrombosis. It is usual to treat the underlying cause of the thrombocytosis but a small dose of aspirin (75 mg) is also sometimes given. In myeloproliferative diseases the primary risk is thrombosis and specific action to reduce the platelet count, usually with hydroxycarbamide (hydroxyurea), is often taken. Paradoxically there is also a risk of abnormal bleeding if the platelet count is very high.

INHERITED COAGULATION DISORDERS

Coagulation disorders may be inherited or acquired. The inherited disorders are uncommon and usually involve deficiency of one factor only. The acquired disorders occur more frequently and almost always involve several coagulation factors; they are considered in the next subsection.

In inherited coagulation disorders, deficiencies of all factors have been described. Those leading to abnormal bleeding are rare, apart from haemophilia A (factor VIII deficiency), haemophilia B (factor IX deficiency) and von Willebrand disease.

Haemophilia A

This is due to a lack of factor VIII. VWF is normal in haemophilia (Fig. 8.39). The prevalence of haemophilia A is about 1 in 5000 of the male population. It is inherited as an X-linked disorder. If a female carrier has a son, he has a 50% chance of having haemophilia, and a daughter has a 50% chance of being a carrier. All daughters of men with haemophilia are carriers and the sons are normal.

The human factor VIII gene is large, constituting about 0.1% of the X chromosome, encompassing 186 kilobases of DNA. Various genetic defects have been found, including deletions, duplications, frameshift mutations and insertions. In approximately 50% of families with severe disease, the defect is an inversion. There is a high mutation rate, with one-third of cases being apparently sporadic with no family history of haemophilia.

Clinical and laboratory features

The clinical features depend on the level of factor VIII. The normal level of factor VIII is 50–150 iu/dL.

- *Levels of less than 1 iu/dL (severe haemophilia)* are associated with frequent spontaneous bleeding from early life. Haemarthroses are common and recurrent bleeding into joints leads to joint deformity and crippling if adequate treatment is not given. Bleeding into muscles is also common, and intramuscular injections should be avoided.
- *Levels of 1–5 iu/dL (moderate haemophilia)* are associated with severe bleeding following injury and occasional spontaneous bleeds.
- *Levels above 5 iu/dL (mild haemophilia)* are associated usually with bleeding only after injury or surgery. It should be noted that patients with mild haemophilia can still bleed badly once haemostasis has failed. Diagnosis in this group is often delayed until quite late in life.

The most common causes of death in people with haemophilia are cancer and heart disease, as for the general population. Cerebral haemorrhage is much more frequent. In recent years, HIV infection and liver disease (due to hepatitis C) have become a more common cause of death. These infections were acquired from blood transfusion by many patients that were treated with factor concentrates prior to 1986. Since 1986 such plasma derived products are all virally inactivated with heat or chemicals.

The main laboratory features of haemophilia A are shown in Table 8.26. The abnormal findings are a prolonged APTT and a reduced level of factor VIII. The PT, bleeding time and VWF level are normal.

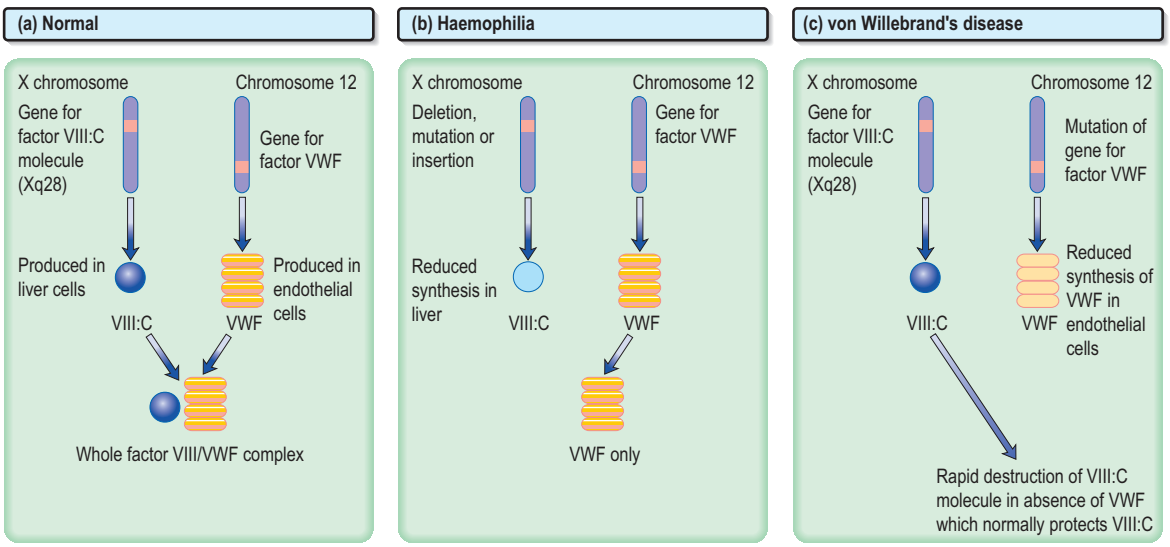


Fig. 8.39 **(a) Normal factor VIII synthesis. (b) Haemophilia A showing defective synthesis of factor VIIIc. (c) von Willebrand disease showing reduced synthesis of VWF.**

Table 8.26 Blood changes in haemophilia A, von Willebrand disease and vitamin K deficiency

	Haemophilia A	von Willebrand disease	Vitamin K deficiency
Bleeding time	Normal	↑	Normal
PT	Normal	Normal	↑
APTT	↑+	↑±	↑
VIII:C	↓++	↓	Normal
VWF	Normal	↓	Normal

Treatment

Bleeding is treated by administration of factor VIII concentrate by intravenous infusion.

- *Minor bleeding*: the factor VIII level should be raised to 20–30 iu/dL.
- *Severe bleeding*: the factor VIII should be raised to at least 50 iu/dL.
- *Major surgery*: the factor VIII should be raised to 100 iu/dL preoperatively and maintained above 50 iu/dL until healing has occurred.

Factor VIII has a half-life of 12 hours and therefore must be administered at least twice daily to maintain the required therapeutic level. Continuous infusion is sometimes used to cover surgery. Factor VIII concentrate is freeze-dried and can be stored in domestic refrigerators at 4°C. This allows it to be administered by the patient immediately after bleeding has started, reducing the likelihood of chronic damage to joints and the need for inpatient care.

Recombinant factor VIII concentrate is well established as the treatment of choice for people with haemophilia, but economic constraints and limited production capacity for recombinant factors have resulted in many patients still being offered treatment with plasma derived concentrates, particularly in developing countries.

To prevent recurrent bleeding into joints and subsequent joint damage, patients with severe haemophilia are given factor VIII infusions regularly three times per week. Such 'prophylaxis' treatment is usually started in early childhood (around 2 years of age).

Synthetic vasopressin (desmopressin – an analogue of vasopressin) – intravenous, subcutaneous or intranasal – produces a 3–5 fold rise in factor VIII and is very useful in patients with a baseline level of factor VIII > 10 iu/dL. It avoids the complications associated with blood products and is useful for treating bleeding episodes in mild haemophilia and as prophylaxis before minor surgery. It is ineffective in severe haemophilia.

People with haemophilia should be registered at comprehensive care centres (CCC), which take responsibility for their full medical care, including social and psychological support. Each person with haemophilia carries a special medical card giving details of the disorder and its treatment.

Complications

Up to 30% of people with severe haemophilia will, during their lifetime, develop antibodies to factor VIII that inhibit its action. Such inhibitors usually develop after the first few treatment doses of factor VIII. The prevalence of inhibitors is, however, only 5–10% because inhibitory antibodies develop only rarely in moderate and mild haemophilia and often disappear spontaneously or with continued treatment.

Management of such patients may be very difficult, as even extremely high doses of factor VIII may not produce a rise in the plasma factor VIII level. Recombinant factor VIIa at very high 'pharmacological' levels can bypass factor IX/VIII activity in the coagulation pathway and is an effective treatment in more than 80% of bleeding episodes in patients with high factor VIII inhibitor levels. Some factor IX concentrates are also deliberately activated to produce factors, which also may 'bypass' the inhibitor and stop the bleeding.

In the longer term the aim of management is to eradicate the inhibitory antibody, particularly in those that have recently developed inhibitors. This is done using immune tolerance induction strategies, sometimes using immunosuppression and immunoabsorption.

The risk of viral transmission has been virtually eliminated in developed countries by excluding high-risk blood donors, testing all donations for HBsAg, HCV and HIV antibodies, and by including steps to inactivate viruses during the preparation of plasma-derived concentrate. However, there is now a concern that the infectious agent that causes variant CJD may be transmissible by blood and blood products. Pooled plasma products are now manufactured solely from plasma drawn from countries with a low incidence of BSE.

Hepatitis A and B *vaccination* is offered routinely to all patients with haemophilia and von Willebrand disease. The clinical consequences of haemophilia patients infected with HIV are similar to other HIV-infected patients (see p. 187), except that Kaposi's sarcoma does not occur. A number of patients with hepatitis C will progress to develop chronic liver disease and cirrhosis (see p. 340).

The use of recombinant factor VIII effectively eliminates any residual risk of transfusion-transmitted infection, and it is safe and effective; but there is a similar incidence of inhibitor development as with plasma-derived factor VIII.

Carrier detection and antenatal diagnosis

Determination of carrier status in females begins with a family history and coagulation factor assays. Female carriers may have a low factor VIII level but the exact value is very variable, partly because of lyonization. Owing to this process early in embryonic life (that is, random inactivation of one chromosome; see p. 37), some carriers have very low levels of factor VIII while others will have normal levels. Carrier detection is best carried out using molecular genetic testing, in the main by direct detection of mutations within the factor VIII gene. Antenatal diagnosis may be carried out by molecular analysis of fetal tissue obtained by chorionic villus biopsy at 11–12 weeks' gestation.

Haemophilia B (Christmas disease)

Haemophilia B is caused by a deficiency of factor IX. The inheritance and clinical features are identical to haemophilia A, but the incidence is only about 1 in 30 000 males. The gene is smaller at 34 kilobases and the half-life of the factor is longer at 18 hours. Haemophilia B is treated with factor IX concentrates, recombinant factor IX being generally available, and prophylactic doses are given twice a week. Desmopressin is ineffective.

Von Willebrand disease (VWD)

In VWD, there is defective platelet function as well as factor VIII deficiency, and both are due to a deficiency or abnormality of VWF (Fig. 8.39). VWF plays a role in platelet adhesion to damaged subendothelium as well as stabilizing factor VIII in plasma (see p. 432).

The VWF gene is located on chromosome 12 and numerous mutations of the gene have been identified. VWD has been classified into three types:

- *Type 1* is partial quantitative deficiency of VWF and significant type 1 VWD is usually inherited as an autosomal dominant.
- *Type 2* is due to a qualitative abnormality of VWF, and it too is usually inherited as an autosomal dominant.
- *Type 3* is recessively inherited and patients have virtually complete deficiency of VWF. Their parents are often phenotypically normal.

Many subtypes of VWD have also been described, particularly type 2, which reflect the specific qualitative changes in the VWF protein.

Clinical features. These are very variable. Type 1 and type 2 patients usually have relatively mild clinical features. Bleeding follows minor trauma or surgery, and epistaxis and menorrhagia often occur. Haemarthroses are rare. Type 3 patients have more severe bleeding but rarely experience the joint and muscle bleeds seen in haemophilia A.

Characteristic laboratory findings are shown in Table 8.26. These also include defective platelet aggregation with ristocetin.

Treatment depends on the severity of the condition and may be similar to that of mild haemophilia, including the use of desmopressin where possible. Some plasma derived factor VIII concentrates contain von Willebrand factor. These specific products are used to treat bleeding or to cover surgery in patients who require replacement therapy, such as those with type 3 (severe) VWD and those who do not respond adequately to desmopressin. Cryoprecipitate can be used as a source of VWF, but should be avoided if possible since it is not a virally inactivated product.

ACQUIRED COAGULATION DISORDERS

Vitamin K deficiency (see also p. 219)

Vitamin K is necessary for the γ-carboxylation of glutamic acid residues on coagulation factors II, VII, IX and X and on proteins C and S. Without it, these factors cannot bind calcium.

Deficiency of vitamin K may be due to:

- *inadequate stores*, as in haemorrhagic disease of the newborn and severe malnutrition (especially when combined with antibiotic treatment) (see p. 324)
- *malabsorption of vitamin K*, a fat-soluble vitamin, which occurs in cholestatic jaundice owing to the lack of intraluminal bile salts
- *oral anticoagulant drugs*, which are vitamin K antagonists.

The PT and APTT are prolonged (Table 8.26) and there may be bruising, haematuria and gastrointestinal or cerebral bleeding. Minor bleeding is treated with phytomenadione (vitamin K_1) 10 mg intravenously. Some correction of the PT is usual within 6 hours but it may not return to normal for 2 days.

Newborn babies have low levels of vitamin K, and this may cause minor bleeding in the first week of life (*classical haemorrhagic disease of the newborn*). Vitamin K deficiency also causes *late haemorrhagic disease of the newborn*, which occurs 2–26 weeks after birth and results in severe bleeding such as intracranial haemorrhage. Most infants with these syndromes have been exclusively breast-fed, and both conditions are prevented by administering 1 mg i.m. vitamin K to all neonates (see p. 219). Concerns about the safety of this are unfounded.

Liver disease

Liver disease may result in a number of defects in haemostasis:

- *Vitamin K deficiency.* This occurs owing to intrahepatic or extrahepatic cholestasis.
- *Reduced synthesis.* Reduced synthesis of coagulation factors may be the result of severe hepatocellular damage. The use of vitamin K does not improve the results of abnormal coagulation tests, but it is generally given to ensure that a treatable cause of failure of haemostasis has not been missed.
- *Thrombocytopenia.* This results from hypersplenism due to splenomegaly associated with portal hypertension or from folic acid deficiency.
- *Functional abnormalities.* Functional abnormalities of platelets and fibrinogen are found in many patients with liver failure.
- *Disseminated intravascular coagulation.* DIC (see below) occurs in acute liver failure.

Disseminated intravascular coagulation (DIC)

There is widespread generation of fibrin within blood vessels, owing to activation of coagulation by release of procoagulant material, and by diffuse endothelial damage or generalized platelet aggregation. Activation of leucocytes, particularly monocytes causing expression of tissue factor and the release of cytokines, may play a role in the development of DIC. There is consumption of platelets and coagulation factors and secondary activation of fibrinolysis leading to production of fibrin degradation products (FDPs), which contributes to the coagulation defect by inhibiting fibrin polymerization (Fig. 8.40). The consequences of these changes are a mixture of initial thrombosis followed by a bleeding tendency due to consumption of coagulation factors and fibrinolytic activation.

Causes of DIC

These include:

- malignant disease
- septicaemia (e.g. Gram-negative and meningococcal)
- haemolytic transfusion reactions
- obstetric causes (e.g. abruptio placentae, amniotic fluid embolism)
- trauma, burns, surgery
- other infections (e.g. falciparum malaria)

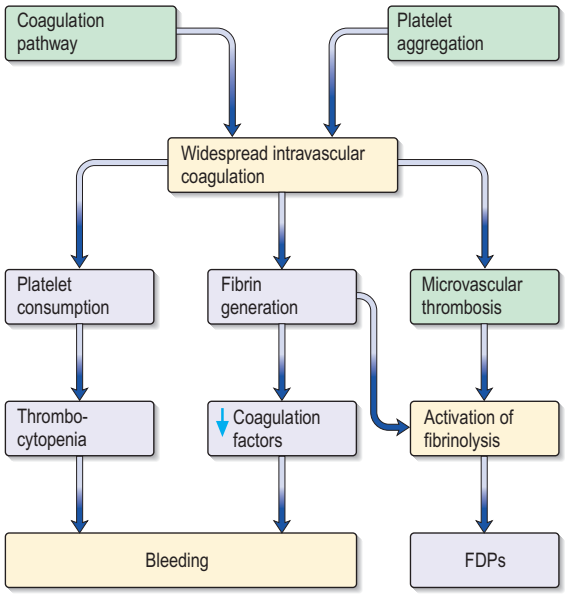


Fig. 8.40 **Disseminated intravascular coagulation.** FDPs, fibrin degradation products.

- liver disease
- snake bite.

Clinical features

The underlying disorder is usually obvious. The patient is often acutely ill and shocked. The clinical presentation of DIC varies from no bleeding at all to profound haemostatic failure with widespread haemorrhage. Bleeding may occur from the mouth, nose and venepuncture sites and there may be widespread ecchymoses.

Thrombotic events occur as a result of vessel occlusion by fibrin and platelets. Any organ may be involved, but the skin, brain and kidneys are most often affected.

Investigations

The diagnosis is often suggested by the underlying condition of the patient.

Severe cases with haemorrhage

- The PT, APTT and TT are usually very prolonged and the fibrinogen level markedly reduced.
- High levels of FDPs, including D-dimer, are found owing to the intense fibrinolytic activity stimulated by the presence of fibrin in the circulation.
- There is severe thrombocytopenia.
- The blood film may show fragmented red blood cells.

Mild cases without bleeding

- Increased synthesis of coagulation factors and platelets
- Normal PT, APTT, TT and platelet counts
- FDPs are raised.

Treatment

The underlying condition is treated and this is often all that is necessary in patients who are not bleeding. Maintenance of blood volume and tissue perfusion is essential. Transfusions of platelet concentrates, FFP, cryoprecipitate and red cell concentrates is indicated in patients who are bleeding. Inhibitors of fibrinolysis such as tranexamic acid should not be used in DIC as dangerous fibrin deposition may result. Activated protein C concentrates have been used in selected cases. In those cases with a dominant thrombotic component the use of heparin seems logical but there is little evidence to suggest any benefit.

Excessive fibrinolysis

Excessive fibrinolysis occurs during surgery involving tumours of the prostate, breast, pancreas and uterus owing to release of tissue plasminogen activators.

Primary hyperfibrinolysis is very rare but activation of fibrinolysis occurs in DIC as a secondary event in response to intravascular deposition of fibrin.

The clinical picture is similar to DIC with widespread bleeding. Laboratory investigations are also similar with a prolonged PT, APTT and TT, a low fibrinogen level, and increased FDPs, although fragmented red cells and thrombocytopenia are not seen, since disseminated coagulation is not present.

If the diagnosis is certain, fibrinolytic inhibitors such as epsilon-aminocaproic acid (EACA) or tranexamic acid can be given but evidence for efficacy is lack.

Massive transfusion

Few platelets and reduced levels of clotting factors are found in stored blood, although there are adequate amounts of the other coagulation factors. During massive transfusion (defined as transfusion of a volume of blood equal to the patient's own blood volume within 24 hours, e.g. approximately 10 units in an adult), the platelet count and PT and APTT should be checked at intervals.

Transfusion of platelet concentrates and FFP should be given if thrombocytopenia or defective coagulation are thought to be contributing to continued blood loss. Other problems of massive transfusion are described on page 905.

Inhibitors of coagulation

Factor VIII autoantibodies arise occasionally in patients without haemophilia but with autoimmune disorders such as SLE, in elderly patients, with malignant disease and sometimes after childbirth. There can be severe bleeding. Immediate bleeding problems are managed as with bypassing factor concentrates (see above). Longer-term therapy is to eliminate the autoantibody using immunosuppression, such as steroids, cyclophosphamide and more recently, in severe cases, rituximab.

Lupus anticoagulant antibodies (see p. 544) are autoantibodies directed against phospholipids (antiphospholipid antibodies) and lead to prolongation of phospholipid dependent coagulation tests, particularly the APTT, but do not inhibit coagulation factor activity.

FURTHER READING

Bolton-Maggs PHB, Pasi KJ. Haemophilias A and B. *Lancet* 2003; **361:** 1801–1809.

George JN. Thrombocytopenic purpura. *New England Journal of Medicine* 2006; **354:** 1927–1935.

Schafer AI. Thrombocytosis. *New England Journal of Medicine* 2004; **350:** 1211–1219.

Wilde JT. Von Willebrand Disease. *Clinical Medicine* 2007; **7:** 629–631.

THROMBOSIS

A thrombus is defined as a solid mass formed in the circulation from the constituents of the blood during life. Fragments of thrombi (emboli) may break off and block vessels downstream. Thromboembolic disease is much more common than abnormal bleeding; nearly half of adult deaths in England and Wales are due to coronary artery thrombosis, cerebral artery thrombosis or pulmonary embolism.

A thrombus results from a complex series of events involving coagulation factors, platelets, red blood cells and the vessel wall.

Arterial thrombosis

This usually occurs in association with atheroma, which tends to form at areas of turbulent blood flow such as the bifurcation of arteries. Platelets adhere to the damaged vascular endothelium and aggregate in response to ADP and TXA_2 to form a 'white thrombus'. The growth of the platelet thrombus is limited at its margins by PGI_2 and NO. Plaque rupture leads to the exposure of blood containing factor VIIa to tissue factor within the plaque, which may trigger blood coagulation and lead to thrombus formation. This results in complete occlusion of the vessel or embolization that produces distal obstruction. The risk factors for arterial thrombosis are related to the development of atherosclerosis (see p. 744).

Arterial thrombi may also form in the heart, as mural thrombi in the left ventricle after myocardial infarction, in the left atrium in mitral valve disease, or on the surfaces of prosthetic valves.

Venous thrombosis

Unlike arterial thrombosis, venous thrombosis often occurs in normal vessels. Major causes are stasis and hypercoagulability. The majority of venous thrombi occur in the deep veins of the leg, originating around the valves as 'red thrombi' consisting mainly of red cells and fibrin. The propagating thrombus is formed of fibrin and platelets and is particularly liable to embolize. Chronic venous obstruction following thrombosis in the deep veins of the leg frequently results in a permanently swollen limb and may lead to ulceration (postphlebitic syndrome).

Risk factors for venous thrombosis are shown in Table 8.27. Venous thrombosis may occur with changes in blood cells such as polycythaemia and thrombocythaemia, and with coagulation abnormalities (thrombophilia; see below).

The clinical features and diagnosis of venous thrombosis are discussed on page 809.

Thrombophilia

Thrombophilia is a term describing inherited or acquired defects of haemostasis leading to a predisposition to venous or arterial thrombosis. It should be considered in patients with:

- recurrent venous thrombosis
- venous thrombosis for the first time under age 40 years
- an unusual venous thrombosis such as mesenteric or cerebral vein thrombosis
- unexplained neonatal thrombosis

Table 8.27 Risk factors of venous thromboembolism

Patient factors	Disease or surgical procedure
Age	Trauma or surgery, especially of pelvis, hip or lower limb
BMI >30 kg/m^2	Malignancy
Varicose veins	Cardiac or respiratory failure
Continuous travel more than 3 h in preceding 4 weeks	Recent myocardial infarction or stroke
Immobility (bed rest > 4 days)	Acute medical illness/severe infection
Pregnancy and puerperium	Inflammatory bowel disease
Previous deep vein thrombosis or pulmonary embolism	Behçet's disease
Thrombophilia	Nephrotic syndrome
Antithrombin deficiency	Myeloproliferative disorders
Protein C or S deficiency	Paroxysmal nocturnal haemoglobinuria
Factor V Leiden	Paraproteinaemia
Resistance to activated protein C (caused by factor V Leiden variant)	Sickle cell anaemia
Prothrombin gene variant	Central venous catheter in situ
Hyperhomocysteinaemia	
Antiphospholipid antibody/lupus anticoagulant	
Oestrogen therapy including HRT	
Dysfibrinogenaemia	
Plasminogen deficiency	

- recurrent miscarriages
- arterial thrombosis in the absence of arterial disease.

Coagulation abnormalities

Factor V Leiden

Factor V Leiden differs from normal factor V by a single nucleotide substitution (Arg506Gln). This variation makes factor V less likely to be cleaved by activated protein C. As factor V is a cofactor for thrombin generation (Fig. 8.33) impaired inactivation by activated protein C (Fig. 8.34) results in a tendency to thrombosis. Factor V Leiden is found in 3–5% of healthy individuals in the western world and in about 20-30% of patients with venous thrombosis.

Factor V Leiden acts synergistically with other acquired thrombosis risk factors, for example in those taking oral contraceptive pills or when pregnant. Thrombosis risk rises 35-fold in those on a combined oral contraceptive pill although the absolute risk for thrombosis remains at significantly less than 0.5% per year for any single individual.

Prothrombin variant

A mutation in the 3′ untranslated region of the prothrombin gene has been described (G20210A). This variant is associated with elevated levels of prothrombin and a two- to three-fold increase in the risk of venous thrombosis. There is an interaction with factor V Leiden and contraceptive pill use or pregnancy. The prevalence is 2% in Caucasian populations, 6% in unselected patients with thrombosis.

Antithrombin (AT) deficiency

This deficiency can be inherited as an autosomal dominant. Many variations have been described that lead to a conformational change in the protein. It can also be acquired following trauma, with major surgery and with the contraceptive pill. Low levels are also seen in severe proteinuria (e.g. the nephrotic syndrome). Recurrent thrombotic episodes occur starting at a young age in the inherited variety. Patients are

relatively resistant to heparin as antithrombin is required for its action. Antithrombin concentrates are available.

Protein C and S deficiency

These autosomal dominant conditions result in an increased risk of venous thrombosis, often before the age of 40 years. Homozygous protein C or S deficiency causes neonatal purpura fulminans, which is fatal without immediate replacement therapy. Protein C concentrate and a recombinant activated protein C are available.

Antiphospholipid antibody

See pages 504 and 544.

Homocysteine

When elevated this amino acid is associated with both arterial thrombosis and venous thromboembolism. The mechanism of vascular damage is unclear. Folate, B_{12} and B_6 supplementation are often helpful in reducing levels.

Investigations

Haemostatic screening tests

- **Full blood count** including platelet count
- **Coagulation screen** including a fibrinogen level. These tests will detect erythrocytosis, thrombocytosis, and dysfibrinogenaemia and the possible presence of a lupus anticoagulant.

Testing for specific causes of thrombophilia

- **Assays** for naturally occurring anticoagulants such as AT, protein C and protein S
- **Assay** for activated protein C resistance and molecular testing for factor V Leiden and the prothrombin variant
- **Screen for a coagulation factor** inhibitor including a lupus anticoagulant (and anticardiolipin antibodies) (see p. 544).

Prevention and treatment of arterial thrombosis

Attempts to prevent or reduce arterial thrombosis are directed mainly at minimizing factors predisposing to atherosclerosis. Treatment of established arterial thrombosis includes the use of antiplatelet drugs and thrombolytic therapy.

Antiplatelet drugs

Platelet activation at the site of vascular damage is crucial to the development of arterial thrombosis, and this can be altered by the following drugs (Table 8.28):

- *Aspirin* irreversibly inhibits the enzyme cyclo-oxygenase (COX), resulting in reduced platelet production of TXA_2 (Fig. 8.34). At the low doses used in cardiovascular disease prevention or treatment, there is selective inhibition of the isoform COX-1 found within platelets. This inhibition cannot be repaired and is effective for the life of the circulating platelet, which is about 1 week. In recent years, it has been suggested there may be significant individual variability in the response to aspirin, although there is no clear reason for this. The term 'aspirin resistance' has been loosely applied when the clinical effects of aspirin are less than expected. No large body of clinical trial data is specifically available to correlate clinical events and laboratory findings with respect to aspirin response and so it is difficult to determine if the breakthrough events experienced by patients treated with aspirin represent aspirin resistance or are related to more mundane issues such as aspirin dose, drug interactions or drug non-compliance.
- *Dipyridamole* – which inhibits platelet phosphodiesterase, causing an increase in cyclic AMP with potentiation of the action of PGI_2 – has been used widely as an antithrombotic agent, but there is little evidence that it is effective.
- *Clopidogrel* – irreversibly blockades the ADP ($P2Y_{12}$) receptor on platelet cell membranes, so affecting the ADP-dependent activation of the glycoprotein IIb/IIIa complex. It is similar to ticlopidine but has fewer side-effects. Trials support its use in acute coronary syndromes (see p. 753), particularly if aspirin is contraindicated.
- *Prasugrel*, a novel thienopyridine, is like clopidogrel. It is being trialled in acute coronary syndromes (see p. 753).
- *Glycoprotein IIb/IIIa receptor antagonists* block the receptor on the platelet for fibrinogen and von Willebrand factor (Fig. 8.41). Three classes have been described:
  (a) murine–human chimeric antibodies (e.g. abciximab)
  (b) synthetic peptides (e.g. eptifibatide)
  (c) synthetic non-peptides (e.g. tirofiban).
  They have been used as an adjunct in invasive coronary artery intervention and as primary medical therapy in coronary heart disease. Excessive bleeding has been a problem.
- *Epoprostenol* is a prostacyclin, which is used to inhibit platelet aggregation during renal dialysis (with or without heparin) and is also used in primary pulmonary hypertension.

The indications for and results of antiplatelet therapy are discussed in the appropriate sections (pp. 753 and 1132).

Table 8.28 Drugs used in the treatment of thrombotic disorders

Antiplatelet	Anticoagulant
Aspirin	Heparin:
Dipyridamole	Unfractionated (or standard)
Clopidogrel	Low molecular weight
GP IIb/IIIa inhibitors, e.g. abciximab, eptifibatide, tirofiban	Hirudin-like, e.g. lepirudin, bivalirudin
Epoprostenol	Fondaparinux
Prasugrel	Warfarin
Thrombolytic	Bivalirudin
Streptokinase	Dabigatran
Tissue-type plasminogen activator (t-PA or alteplase)	Apixaban
Reteplase (r-PA)	Rivaroxaban
Tenecteplase (TNK-tPA)	

Thrombolytic therapy

Streptokinase

Streptokinase is a purified fraction of the filtrate obtained from cultures of haemolytic streptococci. It forms a complex with plasminogen, resulting in a conformational change, which activates other plasminogen molecules to form plasmin. Streptokinase is antigenic and the development of streptococcal antibodies precludes repeated use. Activation

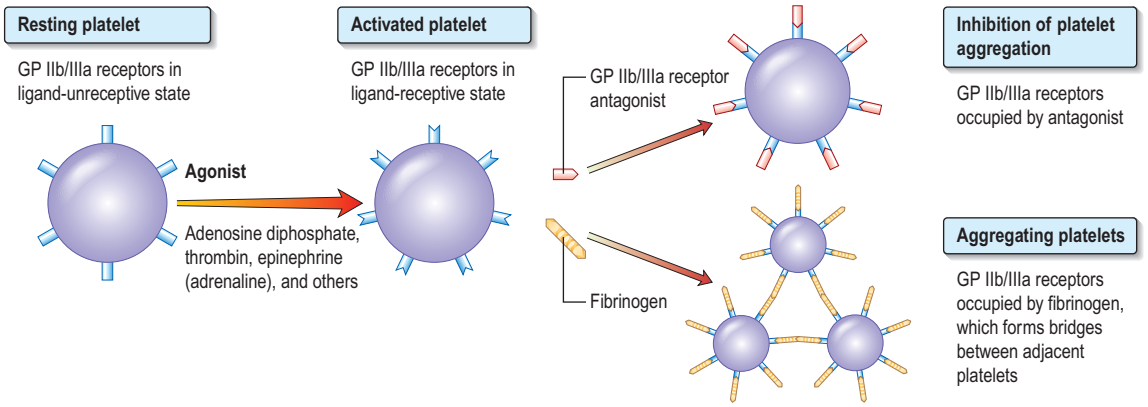


Fig. 8.41 **The role of glycoprotein IIb/IIIa in platelet aggregation and the inhibition of platelet aggregation by inhibitors of glycoprotein IIb/IIIa receptors.** Modified from Lefkovits J, Plow EF, Topol EJ. Mechanisms of disease: platelet glycoprotein IIb/IIIa receptors in cardiovascular medicine. *New England Journal of Medicine* 1995; **332:** 1554, with permission. Copyright © 1995 Massachusetts Medical Society. All rights reserved.

of plasminogen is indiscriminate so that both fibrin in clots and free fibrinogen are lysed, leading to low fibrinogen levels and the risk of haemorrhage.

Plasminogen activators (PA)

Tissue-type plasminogen activators (alteplase (t-PA), tenecteplase (TNK-tPA)) are produced by recombinant technology. Reteplase (r-PA) is also a recombinant plasminogen activator. They are not antigenic and do not give allergic reactions. They are relatively fibrin-specific, have relatively little systemic activity, and short half-lives (~5 minutes). The bleeding complications observed are similar in severity and frequency to those observed with streptokinase, suggesting that fibrin specificity does not confer protection against hemorrhage.

Indications

The use of thrombolytic therapy in myocardial infarction is discussed on page 753. The combination of aspirin with thrombolytic therapy produces better results than thrombolytic therapy alone. The extent of the benefit depends on how quickly treatment is given. They are also used in cerebral infarction (see p. 1131) and occasionally in massive pulmonary embolism. The main risk of thrombolytic therapy is bleeding. Treatment should not be given to patients who have had recent bleeding, uncontrolled hypertension or a haemorrhagic stroke, or surgery or other invasive procedures within the previous 10 days.

Prevention and treatment of venous thromboembolism

Venous thromboembolism is a common problem after surgery, particularly in high-risk patients such as the elderly, those with malignant disease and those with a history of previous thrombosis (Table 8.29). The incidence is also high in patients confined to bed following trauma, myocardial infarction or other illnesses. The prevention and treatment of venous thrombosis includes the use of anticoagulants.

Anticoagulants

Heparin (standard or unfractionated)

Heparin is not a single substance but a mixture of polysaccharides. Commercially available unfractionated heparin consists of components with molecular weights varying from 5000 to 35 000 with an average of about 13 000. It was initially extracted from liver (hence its name) but it is now prepared from porcine gastric mucosa. Heparin acts immediately, binding to antithrombin. This induces a conformational change which increases the inhibitory activity of antithrombin (at least 5000-fold) towards activated serine protease coagulation factors (thrombin, XIIa, XIa, Xa, IXa and VIIa).

Low-molecular-weight heparins (LMW heparins)

These are produced by enzymatic or chemical degradation of standard heparin, producing fractions with molecular weights in the range of 2000–8000. Potentiation of thrombin inhibition (anti-IIa activity) requires a minimum length of the heparin molecule with an approximate molecular weight of 5400, whereas the inhibition of factor Xa requires only a smaller heparin molecule with a molecular weight of about 1700. LMW heparins have the following properties:

- Bioavailability is better than that of unfractionated heparin.
- They have greater activity against factor Xa than against factor IIa, suggesting that they may produce an equivalent anticoagulant effect to standard heparin but have a lower risk of bleeding, although this has not generally been confirmed. In addition, LMW heparins cause less inhibition of platelet function.
- They have a longer half-life than standard heparin and so can be given as a once-daily subcutaneous injection instead of every 8–12 hours.
- They produce little effect on tests of overall coagulation, such as the APTT at doses recommended for prophylaxis. They are not fully neutralized by protamine.

LMW heparins are excreted renally and therefore dose reductions are required in those with renal impairment.

LMW heparins are widely used for antithrombotic prophylaxis, e.g. high-risk surgical patients and for the treatment of established thrombosis (see p. 810).

The main complication of all heparin treatment is bleeding. This is managed by stopping heparin. Very occasionally it is necessary to neutralize unfractionated heparin with protamine. Other complications include osteoporosis with prolonged therapy and thrombocytopenia.

Heparin-induced thrombocytopenia (HIT). HIT is an uncommon complication of heparin therapy and usually occurs 5–

Table 8.29 Classification of risk of deep vein thrombosis and pulmonary embolism for hospital patients (International Consensus Statement Guidelines)

Risk category	General surgery	Gynaecology	Obstetrics*	Medical patients
High	Major general surgery, age > 60 years Major general surgery, age 40–60 years and cancer or history of DVT/PE Thrombophilia Hip/knee arthroplasty	Major gynaecological surgery, age > 60 years Major gynaecological surgery, age 40–60 years and cancer or history of DVT/PE Thrombophilia	History of DVT/PE Thrombophilia	Stroke Age > 70 years Congestive heart failure (CHF) Shock History of DVT/PE Thrombophilia
Moderate	Major general surgery, age 40–60 years without other risk factors† Minor surgery, age > 60 years Minor surgery, age 40–60 years with history of DVT/PE or on oestrogen therapy	Major gynaecological surgery, age 40–60 years Major gynaecological surgery, age < 40 years on oestrogen therapy Minor surgery, age > 60 years	Age > 40 years	Immobilized patient with active disease Cardiac failure
Low	Major general surgery, age < 40 years No other risk factors† Minor surgery, age 40–60 years No other risk factors†	Minor gynaecological surgery, age < 40 without any other risk factors† Minor gynaecological surgery, age 40–60 years without any other risk factors†	Age < 40 years without any other risk factors	Minor medical illnesses

Adapted from Prevention of venous thromboembolism. International Consensus Statement. Guidelines compiled in accordance with the scientific evidence 2001. *International Angiology* **20**: 1–37.
Minor surgery, operations other than abdominal lasting less than 45 minutes. Major surgery, any intra-abdominal operation and all other operations lasting more than 45 minutes.
*The risk of DVT in obstetric patients with pre-eclampsia and the other factors is unknown, but prophylaxis should be considered.
†The risk is increased by infectious disease, presence of varicose veins, general immobility.

14 days after first heparin exposure. It is due to an immune response directed against heparin/platelet factor 4 complexes. All forms of heparin have been implicated but the problem occurs less often with LMW heparins.

HIT is paradoxically associated with severe thrombosis and when diagnosed all forms of heparin must be discontinued, including heparin flush. Unfortunately the diagnosis can be difficult to make because patients on heparin are often very sick and may be thrombocytopenic for many other reasons. Laboratory tests based on bioassay or immunoassay are available but are neither sensitive nor specific and management decisions often have to be made before results are available.

It is necessary to continue some form of anticoagulation in patients with HIT and the choice lies between the heparinoid danaparoid and the direct thrombin inhibitor hirudin. The introduction of warfarin should be covered by one of these agents, as warfarin alone may exacerbate thrombosis as protein C levels fall.

Fondaparinux

This is a synthetic pentasaccharide, which inhibits activated factor X, similar to the LMW heparins. It will now be necessary to establish comparative efficacy (prevention or treatment of thrombosis) against risk of side-effects (bleeding) before the position of fondaparinux becomes clear.

Direct thrombin inhibitors

A recombinant form of hirudin, lepirudin, is available. Hirudins bind directly to thrombin and are effectively irreversible inhibitors. They can be monitored by the use of the APTT and are excreted by the kidney, so must be used with caution in renal failure. *Lepirudin* is used for anticoagulation in patients with HIT.

Bivalirudin is a 20 amino acid synthetic analogue of hirudin. Compared to hirudin it appears to cause less bleeding, is a reversible thrombin inhibitor (as it is broken down by thrombin) and has a shorter half-life. It is used in percutaneous coronary interventions.

Oral anticoagulants

These act by interfering with vitamin K metabolism. There are two types of oral anticoagulants, the coumarins and indanediones. The coumarin warfarin is most commonly used because it has a low incidence of side-effects other than bleeding.

The dosage is controlled by prothrombin tests (PT). Thromboplastin reagents for PT testing are derived from a variety of sources and give different PT results for the same plasma.

It is standard practice to compare each thromboplastin with an international reference preparation so that it can be assigned an international sensitivity index (ISI). The international normalized ratio (INR) is the ratio of the patient's PT to a normal control when using the international reference preparation. Therapeutic ranges using the INR for oral anticoagulation in various conditions are shown in Box 8.4.

Each laboratory can use a chart adapted to the ISI of their thromboplastin to convert the patient's PT to the INR. Suitably selected control plasmas can also be used to achieve the same objective. The use of this system means that PT tests on a given plasma sample using different thromboplastins result in the same INR and that anticoagulant control is comparable in different hospitals across the world.

Contraindications to the use of oral anticoagulants are seldom absolute and include:

Box 8.4 Indications for oral anticoagulation and target INR (British Society for Haematology 2000)

Target INR	
2.5	Pulmonary embolism, proximal and calf deep vein thrombosis, recurrence of venous thromboembolism when no longer on warfarin therapy, symptomatic inherited thrombophilia, atrial fibrillation, cardioversion, mural thrombus, cardiomyopathy
3.5	Recurrence of venous thromboembolism while on warfarin therapy, antiphospholipid syndrome, mechanical prosthetic heart valve, coronary artery graft thrombosis

- severe uncontrolled hypertension
- non-thromboembolic strokes
- peptic ulceration (unless cured by *Helicobacter pylori* eradication)
- severe liver and renal disease
- pre-existing haemostatic defects
- non-compliance.

Oral anticoagulants should be avoided in pregnancy because they are teratogenic in the first trimester and may be associated with fetal haemorrhage later in pregnancy. When anticoagulation is considered essential in pregnancy, specialist advice should be sought. Self-administered subcutaneous heparin should be used as an alternative, although this may not be as effective for women with prosthetic cardiac valves.

Many drugs interact with warfarin (see Ch. 16). More frequent PT testing should accompany changes in medication, which should occur with the full knowledge of the anticoagulant clinic.

An increased anticoagulant effect due to warfarin is usually produced by one of the following mechanisms:

- drugs causing a reduction in the metabolism of warfarin, including tricyclic antidepressants, cimetidine, sulphonamides, phenothiazines and amiodarone
- drugs such as clofibrate and quinidine which increase the sensitivity of hepatic receptors to warfarin
- drugs interfering with vitamin K absorption (such as broad-spectrum antibiotics and colestyramine) which also potentiate the action of warfarin
- displacement of warfarin from its binding site on serum albumin by drugs such as sulphonamides (this is not usually responsible for clinically important interactions)
- drugs that inhibit platelet function (such as aspirin) which increase the risk of bleeding
- alcohol excess, cardiac failure, liver or renal disease, hyperthyroidism and febrile illnesses which result in potentiation of the effect of warfarin.

A decreased anticoagulant effect due to warfarin. This is usually produced by drugs that increase the clearance of warfarin by induction of hepatic enzymes that metabolize warfarin, such as rifampicin and barbiturates.

Anticoagulant related bleeding. Bleeding is the most important serious side-effect of warfarin. Bleeding occurs in up to 4% of patients on oral anticoagulants per year, and is serious in half of these with a 0.5% morbidity associated with it. The benefit of anticoagulants must therefore be notably more than the risk of bleeding. Management of warfarin related bleeding (Emergency Box 8.1) is given dependent upon the INR and the degree of bleeding. Minor bleeding may be treated with cessation of warfarin alone, whilst serious bleeding will require additional use of vitamin K and factor concentrates.

Emergency Box 8.1

Management of bleeding and excessive oral anticoagulation

INR > 3.0 < 6.0 (target INR 2.5)	(1) reduce warfarin dose or stop
INR > 4.0 < 6.0 (target INR 3.5)	(2) restart warfarin when INR < 5.0
INR > 6.0 < 8.0 no bleeding or minor bleeding	(1) stop warfarin
	(2) restart when INR < 5.0
INR > 8.0, no bleeding or minor bleeding	(1) stop warfarin
	(2) restart warfarin when INR < 5.0
	(3) if other risk factors for bleeding give 0.5–2.5 mg of vitamin K (oral)
Major bleeding	(1) stop warfarin
	(2) give prothrombin complex concentrate 50 units/kg or FFP 15 mL/kg if concentrate not available
	(3) give 5 mg of vitamin K (oral or i.v.)

If unexpected bleeding occurs investigate the possibility of a local anatomical cause.

Modified from British Society for Haematology 1998.

New orally active anticoagulant drugs. A number of orally active direct thrombin and Xa inhibitor drugs (e.g. dabigatran, rivaroxaban, apixaban) have been introduced for the prevention of thrombosis. Such drugs have a much broader therapeutic window than warfarin and offer the prospect of fixed drug dosing without the need to monitor coagulation. They do not however have specific antidotes.

Prophylaxis to prevent venous thromboembolism (VTE)

Risk factors for VTE are well defined. Most hospitalized patients have one or more of these risk factors and VTE is common in hospitalized patients. The risk of developing deep vein thrombosis (DVT) after hip replacement surgery has been estimated to be as high as 50% when thromboprophylaxis is not used. Approximately 10% of hospital deaths may be due to pulmonary embolism (PE) and more people die from hospital acquired venous thrombosis than the combined deaths from road traffic accidents, AIDS and breast cancer. PE is the most common preventable cause of hospital death.

Thromboprophylaxis is highly effective and cost effective. Such prophylactic measures include early mobilization, elevation of the legs, compression stockings, intermittent compression devices and use of anticoagulant drugs, such as LMW heparins. All patients, medical and surgical, admitted to hospital should be assessed for thrombotic risk and given

appropriate thromboprophylaxis. Care must be exercised in those patients undergoing neurosurgical procedures.

Low-risk patients (Table 8.29) require no specific measures other than early mobilization.

Moderate and high-risk patients are most effectively managed using graduated compression stockings and LMW heparin subcutaneously daily.

The antithrombin agents dabigatran and rivaroxaban are now being used after lower limb joint replacement surgery. They are as effective as LMW heparins and, as they are given orally, can be used for longer periods.

Treatment of established venous thromboembolism

- The *aim* of anticoagulant treatment is to prevent further thrombosis and pulmonary embolization while resolution of venous thrombi occurs by natural fibrinolytic activity. Anticoagulation is started with heparin as it produces an immediate anticoagulant effect. There is no evidence that it is necessary to use heparin for any longer than it takes for simultaneously administered warfarin to produce an anticoagulant effect (INR 2.5), usually about 5 days.
- *LMW heparin* (e.g. tinzaparin 175 units/kg daily) is equally effective and as safe as unfractionated heparin in the immediate treatment of deep vein thrombosis and pulmonary embolism. This creates the opportunity for treatment of venous thromboembolism without admission to hospital, in compliant patients without coexisting risk factors for haemorrhage.
- *Length of anticoagulation*. This is recommended for at least 6 weeks after isolated calf vein thrombosis and at least 3 months after proximal DVT or PE in patients who have temporary risk factors. For patients with idiopathic VTE or permanent risk factors at least 6 months' anticoagulation is recommended.
- Use of *longer-term anticoagulation* in patients with previous thrombosis. It has been suggested that a lower INR might be safer and equally effective but the current view is that the target INR should be 2.0–3.0 where oral anticoagulation is used. Indefinite anticoagulation is considered appropriate for those with two or more episodes of VTE.
- *Outpatient anticoagulation* is best supervised in anticoagulant clinics. Patients are issued with national booklets for recording INR results and anticoagulant doses. Home monitoring is possible in well-motivated patients.

The role of thrombolytic therapy in the treatment of venous thrombosis is not established. It is used in patients with massive pulmonary embolism who are haemodynamically unstable and in patients with extensive deep venous thrombi.

Thrombolytic therapy should be followed by anticoagulation with heparin for a few days and then by oral anticoagulants to prevent rethrombosis.

CHAPTER BIBLIOGRAPHY

Baglin T, Barrowcliffe A, Greaves M. The British Committee for Standards in Haematology. Guidelines on the use and monitoring of heparin 2006. *British Journal of Haematology* 2006; **133:** 19–34.

Baglin TP, Cousins D, Keeling DM et al. Recommendations from the British Committee for Standards in Haematology and National Patient Safety Agency. *British Journal of Haematology* 2006; **136:** 26–29.

Baglin TP, Keeling DM, Watson HG, British Committee for Standards in Haematology. Guidelines on oral anticoagulation (warfarin): 3rd edn – 2005 update. *British Journal of Haematology* 2005; **132:** 277–285.

British Society for Haematology. Guidelines on oral anticoagulation. *British Journal of Haematology* 2000; **132:** 277–285.

British Society for Haematology. Investigation and management of heritable thrombophilia. *British Journal of Haematology* 2001; **114:** 512–528.

British Society for Haematology. Guideline on the diagnosis and management of the thrombotic microangiopathic haemolytic anaemias. *British Journal of Haematology* 2003; **120:** 556–573.

British Society for Haematology. Guidelines for the investigation and management of idiopathic thrombocytopenic purpura in adults, children and in pregnancy. *British Journal of Haematology* 2003; **120:** 574–596.

British Society for Haematology. The diagnosis for deep vein thrombosis in symptomatic outpatients and the potential for clinical assessment and D-dimer assays to reduce the need for diagnostic imaging. *British Journal of Haematology* 2004; **124:** 15–25.

Buller, HR, Prins MH. Secondary prophylaxis with warfarin for venous thromboembolism. *New England Journal of Medicine* 2003; **349:** 702–704.

Keeling D, Davidson S, Watson H, The Haemostasis and Thrombosis Task Force of the British Committee for Standards in Haematology. Management of heparin-induced thrombocytopenia. *British Journal of Haematology* 2006; **133:** 259–269.

Shapiro SS. Treating thrombosis in the 21st century. *New England Journal of Medicine* 2003; **349:** 1762–1764.

SIGNIFICANT WEBSITES

http://www.bcshguidelines.com
British Society for Haematology guidelines

http://www.blood.co.uk
UK National Blood Service

http://www.bloodline.net
General website on haematology

http://www.doh.gov.uk/bbt2
UK CMO's Better Blood Transfusion Conference

http://www.hemophilia.org
US National Hemophilia Foundation

http://www.med.unc.edu/isth/
International Society on Thrombosis and Haemostasis (ISTH)

http://www.shotuk.org
Serious Hazards of Transfusion (SHOT) scheme, covering UK and Ireland NHS and private hospitals, affiliated to the Royal College of Pathologists (based Manchester Blood Transfusion Centre)

http://www.transfusion.org
Journal of the American Association of Blood Banks

http://www.transfusionguidelines.org.uk
UK Blood Transfusion and Tissue Transplantation Services professional guidelines

http://www.wfh.org
World Federation of Hemophilia

This page intentionally left blank

9 Malignant disease

Aetiology and epidemiology 449
The biology of cancer 451
The diagnosis of malignancy 454
Cancer treatment 457
Principles of chemotherapy 459
Principles of endocrine therapy 463
Principles of biological therapy 463
Principles of radiation therapy 465
Haematological malignancies 466
The leukaemias 467
Acute leukaemias 468
Chronic leukaemias 472
The lymphomas 475
Hodgkin's lymphoma (HL) 476
Non-Hodgkin's lymphoma (NHL) 479
Myeloma 484
Common solid tumour treatment 486
Palliative medicine and symptom control 494

The term 'malignant disease' encompasses a wide range of illnesses, including common ones such as lung, breast and colorectal cancer (Table 9.1), as well as rare ones, like the acute leukaemias. Malignant disease is widely prevalent and, in the West, almost a third of the population will develop cancer at some time during their life. It is second only to cardiovascular disease as the cause of death. Although the mortality of cancer is high, many advances have been made, both in terms of treatment, and in understanding the biology of the disease at the molecular level.

Treatment is given with curative or palliative intent, depending upon the evidence from continuing clinical trials. For many people, the word 'cancer' implies certain death, although this is clearly not always the case. Physicians have an obligation to be honest with their patients, combining realism about the prognosis with compassion and understanding so that patients can take an informed part in treatment decisions.

Table 9.1 Epidemiology of cancer by site of origin in England and Wales

Type	% of all cancers (1998)	% of cancer deaths	M:F
Oral cavity/pharynx	1	1.2	2.3:1
Oesophagus	2.3	4.4	2:1
Stomach	3.7	4.7	2.4:1
Colorectal	11.1	11.0	1.4:1
Pancreas	2.3	4.3	1.5:1
Lung	13.6	22.2	3:1
Melanoma	1.8	1.0	0.6:1
Other skin	14.1	0.3	1.5:1
Breast	12.2	8.6	0.01:1
Cervix	1.2	1.1	
Uterus	1.5	0.7	
Ovary	2.0	2.9	
Prostate	7.4	6.3	
Bladder	4.5	3.5	3.6:1
Kidney	1.8	2.0	2.1:1
Brain	1.3	2.1	1.4:1
Non-Hodgkin's lymphoma	2.8	2.9	1.5:1
Myeloma	1.1	1.6	1.5:1
Leukaemias	2.1	2.6	1.5:1

Cancers < 1% have been excluded.
Derived with permission from Doll R, Peto R. In: *Oxford Textbook of Medicine*, 4th edn. Oxford: Oxford University Press, 2003.

AETIOLOGY AND EPIDEMIOLOGY

In most patients the cause of their cancer remains unknown and is probably multifactorial. Several environmental factors have, however, been identified as being associated with the development of malignancy (Table 9.2) and may be amenable to preventative action such as smoking cessation, dietary modification and antiviral immunization.

Tobacco

The incidence of lung cancer in both men and women had increased dramatically in the last 25 years, but is now falling in many developed countries. The association of smoking with lung cancer is indisputable and causative mechanisms have been identified: cigarette tobacco is responsible for one-third of all deaths from cancer in the UK. Smoking not only causes lung cancer, it is also associated with cancer of the mouth, larynx, oesophagus and bladder. Smoking is discussed on page 829.

Alcohol

Alcohol is associated with cancers of the upper respiratory and gastrointestinal tracts (Table 9.2), and it also interacts with tobacco in the aetiology of these tumours. It may be associated with an increased risk of breast cancer.

Table 9.2 Some causative factors associated with the development of cancer

Smoking	Mouth, pharynx, oesophagus, larynx, lung, bladder, lip
Alcohol	Mouth, pharynx, larynx, oesophagus, colorectal
Iatrogenic:	
Alkylating agents	Bladder, bone marrow
Oestrogens	Endometrium, vagina, breast, cervix
Androgens	Prostate
Radiotherapy e.g. mantle radiotherapy	Carcinoma of breast and bronchus
Diet:	
High-fat diet	Colorectal cancer
Environmental/occupation	
Vinyl chloride	Liver (angiosarcoma)
Polycyclic hydrocarbons	Skin, lung, bladder, myeloid leukaemia
Aromatic amines	Bladder
Asbestos	Lung, mesothelium
Ultraviolet light	Skin, lip
Radiation	e.g. leukaemia, thyroid cancer
Aflatoxin	Liver
Biological agents:	
Hepatitis B virus	Liver (hepatocellular carcinoma)
Hepatitis C virus	Liver (hepatocellular carcinoma)
Human T cell leukaemia virus	Leukaemia/lymphoma
Epstein–Barr virus	Burkitt's lymphoma Hodgkin's lymphoma
Human papilloma virus types 16, 18	Cervix Oral cancer (type 16)
Schistosoma japonicum	Bladder
Helicobacter pylori	Stomach

Diet

Dietary factors have been attributed to account for a third of cancer deaths, although it is often difficult to differentiate these from other epidemiological factors. For example, the incidence of stomach cancer is particularly high in the Far East, while breast and colon cancers are more common in the western, economically more developed countries. Many associations have been observed without a causative mechanism being identified between the incidence of cancer and the consumption of dietary fibre, red meat, saturated fats, salted fish, vitamin E, vitamin A and many others. Food and its role in the causation of gastrointestinal cancer is discussed in Chapter 5. Increasing levels of obesity in the developed world have been associated with increases in women of cancers associated with oestrogenic stimulation of the breast and endometrium.

Environmental/occupational

Ultraviolet light is known to increase the risk of skin cancer (basal cell, squamous cell and melanoma). The incidence of melanoma is therefore particularly high in the white Anglo-Celtic population of Australia, New Zealand and South Africa, where exposure to UV light is combined with a genetically predisposed population.

Occupational factors. In 1775, Percival Pott described the association between carcinogenic hydrocarbons in soot and the development of scrotal epitheliomas in chimney sweeps.

The principal causes now are asbestos (lung and mesothelial cancer) and combustion of fossil fuels (see p. 880) releasing polycyclic hydrocarbons (skin, lung, bladder cancers). Organic chemicals such as benzene may cause molecular abnormalities associated with the development of myeloid leukaemia.

Infectious agents

The geographical distribution of a rare malignancy suggests that it might be caused by, or associated with, an infective agent. Chronic persistent infection provides growth stimulation while many viruses contain transforming viral oncogenes. For example, a specific type of T cell leukaemia, seen almost exclusively in residents of the southern island of Japan and in the West Indies, is caused by infection with the retrovirus HTLV-1 (human T cell leukaemia virus) which is endemic in these areas.

Hepatocellular carcinoma occurs in patients with hepatitis B and C virus infections, and Burkitt's lymphoma and nasopharyngeal carcinoma are associated with the Epstein–Barr virus. EBV is also linked with Hodgkin's lymphoma (see p. 476). Patients with HIV infection or immunosuppression from organ transplantation have an increased incidence of EBV-related lymphoma and herpesvirus-8-associated Kaposi's sarcoma. The incidence of cervical cancer had increased among younger women in association with sexually transmitted human papillomavirus infection types 16 and 18 but a vaccine is now available.

Bacterial infection with *Helicobacter pylori* predisposes to the development of gastric cancer and gastric lymphoma, while *Schistosoma japonicum* infection predisposes to the development of squamous carcinomas in the bladder.

Iatrogenic

Drugs. Oestrogens have been implicated in the development of vaginal, endometrial and breast carcinoma. Alkylating agents given, for example, for Hodgkin's lymphoma (see later) are themselves associated with an increased incidence of secondary acute myelogenous leukaemia (AML), bladder and lung cancer. The anthracyclines and the epipodophyllotoxin drug, etoposide, have also been shown to be associated with the development of secondary AML amongst long-term survivors.

Radiation. The nuclear disasters of Hiroshima, Nagasaki and Chernobyl led to an increased incidence of leukaemia after 5–10 years in the exposed population. Increased incidences of thyroid and breast cancer have also been reported. Long-term survivors following radiotherapy used, for example, in ankylosing spondylitis, Hodgkin's lymphoma, testicular germ cell tumours and childhood cancers has led to increased incidences of cancer at the radiation field margins.

Geographical distribution

The incidence and mortality from cancer varies by tumour type and geographical region across the world. The

Ask the authors
www.studentconsult.com

incidence varies according to a number of known factors such as age, since most cancers occur in those over the age of 65 who comprise 3.3% of the population in Africa compared with 15.2% in Europe. Smoking prevalence in the previous 20 years is reflected in the rates of lung, oesophageal and oral cancers, reproductive patterns influence breast cancer, diet is associated with breast, bowel and stomach cancers, and infection with stomach, liver and cervix cancers. Incidence and mortality are closely linked for cancers for which treatment has yet to make significant improvements such as lung, stomach and liver, while in countries with effective screening programmes there is an increasing incidence and decreasing mortality for breast, cervix, bowel, and prostate cancers.

Lung cancer incidence and mortality is highest in eastern Europe (66M + 35F per 100 000 population) with 50–60% smoking prevalence. *Breast cancer* incidence is highest in North America at 97/100 000 but mortality is lower at 17/100 000 than north and western Europe. *Bowel cancer* incidence is highest in Australia and New Zealand (49M + 38F/100 000) but mortality at 20/100 000 is the same as in eastern and western Europe. *East Asia* has the highest incidence (46M + 22F/100 000) and mortality (35M + 15F/100 000) from stomach cancer. The diagnosis of *prostate cancer* has been increased by PSA testing in North America to 120/100 000 but mortality at 15/100 000 is lower than the Caribbean, 28/100 000. *Liver cancer* is most frequent in East Asia (37M + 12F/100 000) related to infection with hepatitis B and C, while in East Africa *cervical cancer* incidence (43/100 000) and mortality (35/100 000) are highest due to HPV infection. *Oesophageal cancer* incidence in East Asia (24M + 8F/100 000) and *oral cancer* in South Central Asia (13M + 8F/100 000) are linked to the consumption of tobacco and beetel.

Environmental factors interact with genetic predisposition. For example, subsequent generations of people moving from countries with a low incidence to those with a high incidence of breast or colon cancer acquire the cancer incidence of the country to which they have moved while northern European stock exposed to strong UV radiation have the highest risk of developing melanoma.

THE BIOLOGY OF CANCER

Most human neoplasms are clonal in origin, i.e. they arise from genetic mutations within a single population of precursor or cancer stem cells. However, over subsequent cell divisions heterogeneity develops with the accumulation of further abnormalities. The genes most commonly affected can be characterized as those controlling cell cycle check points, DNA repair and DNA damage recognition, apoptosis, differentiation and growth signalling. Proliferation may continue at the expense of differentiation, which together with the failure of apoptosis leads to tumour formation with the accumulation of abnormal cells varying in size, shape and nuclear morphology as viewed down the light microscope.

Proliferation

The *kinetics* of cancer cell growth are exponential; however, the doubling times of human tumours are enormously variable. Mutations are common in the genes controlling a series of intracellular proteins, such as the cyclins and cyclin-dependent kinases (see p. 28), and oncogene products such as c-myc, and the ras proteins (see Cancer genetics, p. 49) that regulate proliferation. Proliferation may also be abnormal due to persistence of the nuclear enzyme telomerase which prevents senescence, loss of cell–cell adhesion signalling, abnormal stromal nutrient supply or cytokine signalling.

Epithelial growth factor (EGF), vascular endothelial growth factor and *fibroblast growth factor* and their receptors are overexpressed in many human epithelial tumours, leading to receptor tyrosine kinase signalling and unrestrained growth which has been the target of several new drugs such as imatinib. *Transforming growth factor-β (TGF-β)*, a cytokine which has effects on extracellular matrix proteins, angiogenesis (see below) and immune effector cells, is also often overexpressed in tumour cells, and defects in TGF-β signalling are often found in cancer cells. Defects in tumour suppressor genes such as p53 (encoded by the *TP53* gene located on the short arm of chromosome 17 (17p13.1)) and the less common MADH4 or their regulatory genes are common in most epithelial cancers and form the focus of much research into targeted therapies.

Apoptosis and growth

Tumour cell death may also be dysregulated. Normal cells usually die by an active and tightly regulated process known as apoptosis, or 'programmed cell death' (see p. 28). Apoptosis can occur in response to a number of physiological cell signals or pathological stimuli (tumour necrosis factor, Fas ligand and DNA-damaging cytotoxic drugs) and is mediated within the cell by a family of proteins known as caspases. Caspase activity is, in turn, regulated by intracellular inhibitors such as the Bcl-2 family of proteins and the inhibitor of apoptosis proteins (IAPs). Disturbances in the normal balance of these various proteins have been identified and targetted for treatment in the upregulation of the bcl-2 protein in follicular non-Hodgkin's lymphoma.

Tumour immunology

Tumour cells are usually not recognized and killed by the immune system. There are two main reasons. The *first* is failure to express molecules such as HLA and costimulatory B7 molecules. These are required for activation of cytotoxic, or 'killer', T lymphocytes, since expression of these 'costimulatory' molecules following gene transfection may augment an immune response. *Secondly*, tumours may also actively secrete immunosuppressive cytokines and cause a generalized immunosuppression. This can lead, for example, to the reactivation of latent herpes zoster in shingles associated with malignancy. Research has identified new strategies for antigen presentation for future vaccines (see also immunotherapy, p. 75).

Angiogenesis

For many tumours, there is a progressive slowing of the rate of growth as the tumours become larger. This occurs for many reasons, but outgrowing the blood supply is paramount. New vessel formation (angiogenesis) is stimulated by a variety of peptides produced both by tumour cells and by host inflammatory cells, such as basic fibroblast growth

factor (bFGF), angiopoietin 2 and vascular endothelial growth factors (VEGFs), which are stimulated by hypoxia. Inhibition of angiogenesis is being developed as a method of cancer therapy, to deprive the cancer of nutrients and oxygen, and inhibit haematogenous spread, or metastasis, e.g. bevacizumab, the anti-VEGFR monoclonal antibody

Invasion and metastasis

Cancers spread by both local invasion and by distant metastasis through the vessels of the blood and lymphatic systems. Infiltration into surrounding tissues is associated with loss of cell–cell cohesion which is mediated by active homotypic cell adhesion molecules (CAMs). The cadherin molecules in association with the integrins are transmembrane heterodimeric glycoproteins able to mediate cellular attachment. Epithelial cadherin (E-cadherin) is expressed by many carcinomas, and mutated in some such as familial gastric carcinoma (see p. 264).

Invasion is also determined by the balance of activators to inhibitors of proteolysis. The balance between the expression and activity of the matrix metalloproteinases (MMPs) and their tissue inhibitors (TIMPs) is involved in tumour growth, invasion, metastasis and angiogenesis, and is being targeted in new therapeutic drugs for cancer treatment.

Dissemination of tumour cells occurs through intravasation into the vascular and lymphatic vessels and dissemination to distant sites, partly by chance, but also because of specific interactions between receptors and cytokines found on stromal and tumour cells. Extravasation is followed by angiogenesis to form an established metastasis.

Bone metastases have been extensively investigated because they occur in 75% of patients with advanced breast and prostate cancer but in only 25% of patients with other solid tumours, e.g. lung, GI tract, thyroid, bladder or kidney.

Metastases are either osteolytic or osteoblastic with some patients having both. Prostate cancer is predominantly osteoblastic while most patients with breast cancer and all with multiple myeloma have osteolytic lesions.

Bone is a frequent site of metastases due to:

- high blood flow
- tumour cell production of adhesins which bind them to marrow stromal cells
- growth factors in bone, including TGFβ, insulin-like growth factor (ILG) 1 and 2, platelet-derived growth factor and fibroblastic growth factors.

Osteolytic metastases (Fig. 9.1). The destruction of bone is mediated by osteoclasts driven by the tumour cells production of parathyroid hormone-related peptide, IL-6, prostaglandin E_2, TNF and macrophage colony-stimulating factor (M-CSF). These increase the expression of receptor activity of nuclear factor κB ligand (RANKL) which directly induces formation of osteoclasts and bone resorption. Bone destruction increases calcium levels and release of matrix cytokines, which promote tumour growth and production of PTH-related peptide, so further increasing osteolytic bone destruction in many tumours. In multiple myeloma there is, additionally, inhibition of osteoblast activity (see p. 484). Bisphosphonates prevent or reverse this cascade by inhibiting osteoclasts.

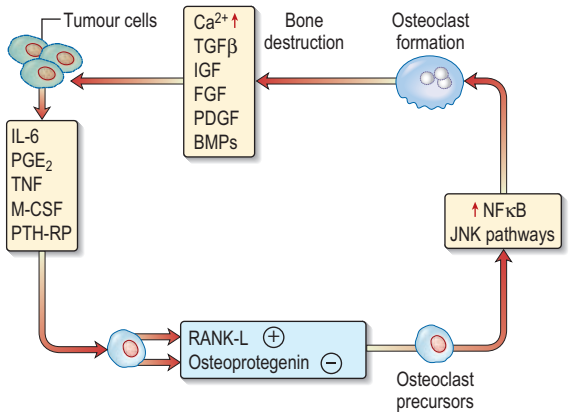


Fig. 9.1 **Mechanisms of osteolytic metastases.** Tumour cells secrete hormones and cytokines as shown. These increase various factors, e.g. RANKL, and lead to increased osteoclastic activity. Bone destruction (resorption) in turn produces factors which increase tumour growth. Osteoprotegerin (see p. 558) has an inhibitory effect on RANKL. BMPs, bone morphogenic proteins; FGF, fibroblast growth factor; IGF, insulin-like growth factor; IL-6, interleukin 6; JNK, jun N-terminal kinase; M-CSF, macrophage colony-stimulating factor; NFκB, nuclear factor kappa B; PDGF, platelet-derived growth factor; PGE_2, prostaglandin E2; PTH-RP, parathyroid hormone-related peptide; RANKL, receptor activator of NFκB ligand; TGF, transforming growth factor; TNF, tumour necrosis factor.

Osteoblastic metastases. The mechanism for this is less clear. It has been suggested that osteoclastic activity precedes osteoblastic activity and bone formation. It is also possible that the vicious circle (as in osteoclastic activity) may be in action, whereby the tumour induces osteoblastic activity and the release of growth factors for osteoblasts, which then increases the growth of tumours. Endothelin 1 has been shown to stimulate bone formation and its levels are increased in, for example, prostate and breast cancers.

Cancer genetics

The development of cancer is a multistep process associated with fundamental genetic changes within the cell. Evidence for the genetic origin of cancer is based on the following:

- Some cancers show a familial predisposition.
- Most known carcinogens act through induced mutations.
- Susceptibility to some carcinogens depends on the ability of cellular enzymes to convert them to a mutagenic form.
- Genetically determined traits associated with a deficiency in the enzymes required for DNA repair are associated with an increased risk of cancer.
- Some cancers are associated with chromosome 'instability' because of deficiencies in mismatch repair genes.
- Many malignant tumours represent clonal proliferations of neoplastic cells.
- Many malignancies contain well-described cytogenetic abnormalities, which involve mutated or abnormally regulated oncogenes and tumour suppressor genes with transforming activity in cell lines.

Table 9.3 Familial cancer syndromes

	Gene	Neoplasms
Autosomal dominant		
Retinoblastoma	*RB1*	Eye
Wilms' tumour	*WT1*	Kidney
Li–Fraumeni	*p53*	Sarcoma/brain/ leukaemia
Neurofibromatosis type 1	*NF1*	Neurofibromas
Familial adenomatous polyposis (FAP)	*APC*	Colon
Hereditary non-polyposis colon cancer (HNPCC)	*MLH1* and *MSH2*	Colon, endometrium
Hereditary diffuse gastric cancer syndrome	*E-cadherin*	Stomach
Breast ovary families	*BRCA1* *BRCA2* *p53*	Breast/ovary
Melanoma	*p16*	Skin
Von Hippel–Lindau	*VHL*	Renal cell carcinoma and haemangio-blastoma
Multiple endocrine neoplasia Type 1	*MEN1*	Pituitary, pancreas, parathyroid
Multiple endocrine neoplasia Type 2	*RET*	Thyroid, adrenal medulla
Autosomal recessive		
Xeroderma pigmentosa	*XP*	Skin
Ataxia telangiectasia	*AT*	Leukaemia, lymphoma
Fanconi's anaemia	*FA*	Leukaemia, lymphoma
Bloom's syndrome	*BS*	Leukaemia, lymphoma

Mutations may occur in the germline and therefore be present in every cell in the body, or they may occur by somatic mutation in response, for example, to carcinogens, and therefore be present only in the cells of the tumour.

Expression of the mutation and hence carcinogenesis will depend upon the penetrance (due to level of expression and presence of other genetic events) of the gene and whether the mutated allele has a dominant or recessive effect. There are a small group of autosomal dominant inherited mutations such as RB (in retinoblastoma) and a small group of recessive mutations (Table 9.3). Carriers of the recessive mutations are at risk of developing cancer if the second allele becomes mutated, leading to 'loss of heterozygosity' within the tumour, although this is seldom sufficient as carcinogenesis is a multistep process.

Malignant transformation may result from a gain in function as cellular proto-oncogenes become mutated (e.g. *ras*), amplified (e.g. *HER2*) or translocated (e.g. *BCR-ABL*). However, these mutations are insufficient to cause malignant transformation by themselves. Alternatively, there may be a loss of function of tumour suppressor genes such as P53 that normally suppress growth. Loss or gain of function may also involve alterations in the genes controlling the transcription of the oncogenes or tumour suppressor genes (e.g. p. 49) (Tables 9.3 and 9.4).

Table 9.4 Examples of acquired/somatic mutations and proto-oncogenes

Point mutation	
K-*ras*	Pancreatic cancer
DNA amplification	
myc	Neuroblastoma
HER2-neu	Breast cancer
Chromosome translocation	
BCR-ABL	CML, ALL
PML-RAR	APML
Bcl-2/IgH	Follicular lymphoma
c-myc and Ig	Burkitt's lymphoma

CML, chronic myeloid leukaemia; ALL, acute lymphoblastic leukaemia; APML, acute promyelocytic leukaemia.

DNA repair

Autosomal dominant

The following are examples of cancer syndromes that exhibit dominant inheritance (Table 9.3):

- *Retinoblastoma* is an eye tumour found in young children. It occurs in both hereditary (40%) and non-hereditary (60%) forms. The 40% of patients with the hereditary form have a germline mutation on the long arm of chromosome 13 that predisposes to retinoblastoma. In addition to the latter, children inheriting this mutation at the so-called *RB1* locus are at risk for developing other tumours, particularly osteosarcoma.
- *Breast and ovarian cancer*. Two genes have been identified – *BRCA1* and *BRCA2*. A strong family history along with germline mutation of these genes accounts for most cases of familial breast cancer and over half of familial ovarian cancers. BRCA1 and 2 proteins bind to the DNA repair enzyme Rad51 to make it functional in repairing DNA breaks. Mutations in the *BRCA* genes will lead to accumulation of unrepaired mutations in tumour-suppressor genes and crucial oncogenes.
- *Neurofibromatosis*. Inactivation of the *NF1* gene will lead to constitutive activation of ras proteins.
- *Multiple-endocrine-adenomatosis syndromes* (see p. 1026). Multiple endocrine neoplasia type 1 is associated with the *MEN1* gene and type 2 (*MEN2*) is associated with mutations in the *RET* proto-oncogene on chromosome 10 and as such are the exception to all the other syndromes which involve tumour suppressor genes.

Autosomal recessive

Some relatively rare autosomal recessive diseases associated with abnormalities of DNA repair predispose to the development of cancer (Table 9.3).

- *Xeroderma pigmentosum*. There is an inability to repair DNA damage caused by ultraviolet light and by some chemicals, leading to a high incidence of skin cancer.
- *Ataxia telangiectasia*. Mutation results in an increased sensitivity to ionizing radiation and an increased incidence of lymphoid tumours.
- *Bloom's syndrome and Fanconi's anaemia*. An increased susceptibility to lymphoid malignancy is seen.

It is not known why these chromosome-break syndromes predispose to tumours of lymphatic tissue.

FURTHER READING

Chiang AC, Massagué J. Molecular basis of metastasis. *New England Journal of Medicine* 2008; **359**: 2814–2823.

Croce CM. Oncogenes and cancer. *New England Journal of Medicine* 2008; **358**: 502–511.

Foulkes WD. Inherited susceptibility to common cancers. *New England Journal of Medicine* 2008; **359**: 2143–2153.

Finn OJ. Molecular origins of cancer: cancer immunology. *New England Journal of Medicine* 2008; **358**: 2704–2715.

Fröhling S, Döhner H. Chromosomal abnormalities in cancer. *New England Journal of Medicine* 2008; **359**: 722–734.

Kerbel RS. Molecular origins of cancer; tumor angiogenesis. *New England Journal of Medicine* 2008; **358**: 2039–2049.

Kumaran G, Clamp AR, Jayson GC. Angiogenesis as a therapeutic target in cancer. *Clinical Medicine* 2008; **8**: 455–458.

THE DIAGNOSIS OF MALIGNANCY

Most common cancers (but not the haematological cancers) start as focal microscopic clones of transformed cells, and diagnosis only becomes likely once sufficient tumour bulk has accumulated to cause symptoms or signs. In order to try to make an earlier diagnosis and increase the curative possibilities, an increasing number of screening programmes are being developed which target the asymptomatic or preinvasive stages of the cancer as in cervix, breast and colon or use serum tumour markers as in prostate and ovarian cancers.

Screening

Genetic screening can be used to target screening to groups at most risk of developing cancer, e.g. *BRCA1* positive and breast cancer. The aim of screening programmes is to improve individual and/or population survival by detecting cancer at its very early stages when the patient is asymptomatic. This strategy is dependent upon finding tests that are sufficiently sensitive and specific, using detection methods that identify cancer before it has spread, and having curative treatments that are practical and consistent with maintenance of a normal lifestyle and quality of life.

Screening is provided to populations, e.g. for breast, cervical and colon cancer in the UK, and also to individuals via annual check-ups, or opportunistic when patients see their doctor for other reasons.

Unfortunately, earlier diagnosis does not necessarily mean longer survival. The patient is merely treated at an earlier date and hence the survival appears longer, death still occurs at the same time from the point of genesis of the cancer. This is called lead time bias (Fig. 9.2). With length time bias, a greater number of slowly growing tumours are detected when screening asymptomatic individuals leading to a false impression of an improvement in survival.

An effective screening procedure should:

- be affordable to the healthcare system
- be acceptable to all social groups so that they attend for screening
- have a good discriminatory index between benign and malignant lesions
- show a reduction in mortality from the cancer.

Cervical cancer. The smear test is cheap and safe though seen as unacceptably invasive by some women. Cervical smears require a well-trained cytologist to identify the early changes (dyskaryosis and cervical intraepithelial neoplasia (CIN)). However developments in liquid cytology and DNA testing for human papillomavirus (HPV) have been introduced. The introduction of this test reduces the incidence and mortality from cervical cancer though it has not been subject to randomized trials. Future screening may be linked to the introduction of vaccination against HPV infection for women before they become sexually active.

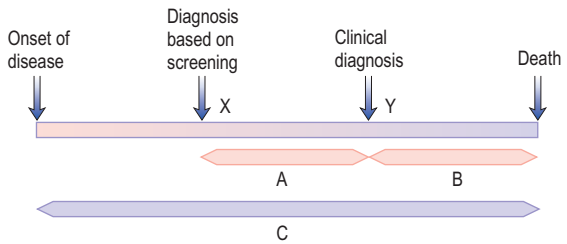


Fig. 9.2 **Lead time bias.** Earlier diagnosis, at X, made by screening tests before the clinical diagnosis, at Y, suggests an increased survival time of A + B. The actual survival time (C) remains unchanged.

Breast cancer. The UK NHS Breast Screening Programme (i.e. biplanar mammography every 3 years) for women aged 50–70 years has been shown to reduce mortality from breast cancer in randomized control studies. The test is acceptable to most women with 75% of women attending for screening.

The cost is estimated to be between £250 000 and £1.3 million per life saved, money which, according to critics of screening, could be used more appropriately in better treatment.

Women from families with *BRCA1*, *BRCA2* and *p53* mutations require intensive screening starting at an earlier age.

Colorectal cancer (CRC). Faecal occult blood is an acceptable and cheap test for the detection of CRC. The false-positive rates are high, meaning many unnecessary colonoscopies (see p. 270). The UK has recently introduced a national screening programme for colorectal cancer using faecal occult blood in patients 60–64 years. This will be evaluated over the coming years.

Other population-based screening programmes that are being used or are in trials are:

Prostate cancer. Serum prostate-specific antigen (PSA) can be used for the detection of this cancer, which is on the increase. The natural history of the disease is unclear. Many men over 70 have evidence of prostate cancer at post mortem with no symptoms of the disease and it has been suggested that over 75-year-olds should not have screening PSAs. The test must be interpreted with caution due to the natural increase in PSA with age, benign prostatic hypertrophy, and with prostatitis.

Treatment may be associated with unacceptable side-effects (see p. 646) and no overall survival advantage has yet been shown in population studies. It may be life saving for the individual diagnosed with a high grade tumour that is still amenable to curative treatment.

Epithelial ovarian cancer. Serum CA125 can be used for the early detection of this cancer and is the subject of ongoing trials. An improvement in survival of a screened population can be shown but at the cost of many unnecessary laparotomies so that further enhancements are being investigated by serial testing and in combination with transvaginal ultrasound scans.

Symptoms of cancer

Patients often present with a history of tumour site-specific symptoms, e.g. pain, and physical signs, e.g. a mass, which readily identify the primary site of the cancer. On the other hand, many seek medical attention when more systemic and non-specific symptoms occur such as weight loss, fatigue

and anorexia. These usually indicate a more advanced stage of the disease except in some paraneoplastic and ectopic endocrine syndromes (see below). Other patients are only diagnosed upon the discovery of established metastases such as the abdominal distension of ovarian cancer, the back pain of metastatic prostatic cancer or the liver enlargement of metastatic gastrointestinal cancer.

Other indirect effects of the cancer manifest as paraneoplastic syndromes (Box 9.1) that are often associated with specific types of cancer and are reversible with treatment of the cancer. The effects and mechanisms can be very variable. For example in the Lambert–Eaton syndrome (see p. 1181) there is cross-reactivity between tumour antigens and the normal tissues, e.g. the acetylcholine receptors at neuromuscular junctions.

The *coagulopathy of cancer* may present with thrombophlebitis, deep venous thrombosis and pulmonary emboli, particularly in association with cancers of pancreas, stomach and breast. Some 18% of patients with recurrent pulmonary embolus will be found to have an underlying cancer and cancer patients are also at risk of a pulmonary embolus following diagnosis.

Other symptoms are related to peptide or hormone release, e.g. carcinoid or Cushing's syndrome.

Cachexia of advanced cancer is thought to be due to release of chemokines such as tumour necrosis factor (TNF), as well as the fact that patients have a loss of appetite.

Cancer-associated immunosuppression can lead to reactivation of latent infections such as herpes zoster.

Physical examination

A general examination should be performed to include:

- main symptomatic areas, e.g. site, size of mass and associated lymphadenopathy
- precursor lesions, e.g. solar keratosis, dysplastic naevi
- general signs, e.g. jaundice, clubbing
- functional capacity (see Table 9.6).

Histology

The diagnosis of cancer may be suspected by both patient and doctor but advice about treatment can usually only be given on the basis of a tissue diagnosis. This may be obtained by surgical biopsy or on the basis of cytology (e.g. lung cancer diagnosed by sputum cytology). Malignant lesions can be distinguished morphologically from benign by the pleomorphic nature of the cells, increased numbers of mitoses, nuclear abnormalities in size, chromatin pattern and nucleolar organization, and evidence of invasion into surrounding tissues.

The degree of differentiation (or conversely of anaplasia) of the tumour has prognostic significance: generally speaking, more differentiated tumours have a better prognosis than poorly differentiated ones.

Box 9.1 Paraneoplastic syndromes

Syndrome	Tumour	Serum antibodies
Neurological		
Lambert–Eaton syndrome	Lung (small-cell) lymphoma	Anti VGLC
Peripheral sensory neuropathy	Lung (small-cell), breast and ovary lymphoma	Anti Hu
Cerebellar degeneration	Lung (particularly small-cell) lymphoma	Anti Yo
Opsoclonus/myoclonus (p. 1109)	Breast, lung (small-cell)	Anti Ri
Stiff man syndrome	Breast, lung (small-cell)	Anti-amphiphysin
Limbic, hypothalamic, brain stem encephalitis	Lung	Anti Ma protein
	Testicular	Anti-NMDAR
Endocrine/metabolic		
SIADH	Lung (small-cell)	
Ectopic ACTH secretion	Lung (small-cell)	
Hypercalcaemia	Renal, breast	
Musculoskeletal		
Hypertrophic pulmonary osteoarthropathy	Lung (non-small-cell)	
Clubbing	Lung	
Skin		
Dermatomyositis/polymyositis	Lung and upper GI	
Acanthosis nigricans	Mainly gastric	
Hyperpigmentation	Lung (small-cell)	
Pemphigus	Non-Hodgkin's lymphoma, CLL	
Haematological		
Erythrocytosis	Renal cell carcinoma, hepatocellular carcinoma, cerebellar haemangioblastoma	
Migratory thrombophlebitis	Pancreatic adenocarcinoma	
DIC	Adenocarcinoma	

SIADH, syndrome of inappropriate antidiuretic hormone secretion; ACTH, adrenocorticotrophic hormone; CLL, chronic lymphocytic leukaemia; DIC, disseminated intravascular coagulation; NMDAR, N-methyl-D-aspartate receptors.

Tissue tumour markers: Immunocytochemistry, using monoclonal antibodies against tumour antigens, is very helpful in differentiating between lymphoid and epithelial tumours and between some subsets of these, for example T and B cell lymphomas, germ cell tumours, prostatic tumours, neuroendocrine tumours, melanomas and sarcomas. However, there is much overlap in the expression of many of the markers and some adenocarcinomas and squamous carcinomas do not bear any distinctive immunohistochemical markers that are diagnostic of their primary site of origin.

There are several tests for genetic markers in tissue sections. For example, fluorescent in situ hybridization (FISH, see p. 41) can be used to look for characteristic chromosomal translocations, e.g. in lymphoma and leukaemia, deletions or amplifications, e.g. in breast cancer (see genetic basis of cancer, p. 49). Tissue microarrays can identify patterns of multiple genomic alterations and single nucleotide polymorphisms (SNPs), e.g. in breast cancer and lymphoma (see p. 35), with prognostic and predictive relevance. Increasingly genomics and proteomics are being investigated in order to target new (and expensive) therapies, e.g. imatinib in CML and GIST, trastuzumab and lapatinib in breast cancer, and erlotinib in lung cancer, but it is still unclear which of the many gene signatures are most reliable and reproducible.

Staging

Before a decision about treatment can be made, not only the type of tumour but also its extent and distribution need to be established. Various 'staging investigations' are therefore performed before a treatment decision is made. To be useful clinically the staging system must subdivide the patients into groups of different prognosis which can guide treatment selection.

The staging systems vary according to the type of tumour and may be site specific (see Hodgkin's lymphoma, p. 476), or the TNM (tumour, node, metastases) classification shown in Table 9.5 which can be adapted for application to most common cancers.

Performance status

In addition to anatomical staging, the person's age and general state of health need to be taken into account when planning treatment. The latter has been called 'performance status' and is of great prognostic significance for all tumour types (Table 9.6). Performance status reflects the effects of the cancer on the patient's functional capacity. An alternative performance rating scale is by Karnowsky. With a performance status ≥ 2, response to and survival following treatment are greatly reduced for most tumour types.

Table 9.5 TNM classification as used for lung cancer

T = extent of primary tumour; N = extent of regional lymph node involvement; M = presence of distant metastases	
Tx	Positive cytology only
T1	<3 cm diameter
T2	>3 cm/extends to hilar region/invades visceral pleura/partial atelectasis
T3	Involvement of chest wall, diaphragm, pericardium, mediastinum, pleura, total atelectasis
T4	Involvement of heart, great vessels, trachea, oesophagus, malignant effusion
N1	Peribronchial, ipsilateral hilar lymph node involvement
N2	Ipsilateral mediastinal
N3	Contralateral mediastinal, scalene or supraclavicular
M0	No distant metastases
M1	Metastases present

Table 9.6 Eastern Cooperative Oncology Group (ECOG) performance status scale

Status	Description
0	Asymptomatic, fully active and able to carry out all predisease performance without restrictions
1	Symptomatic, fully ambulatory but restricted in physically strenuous activity and able to carry out performance of a light or sedentary nature, e.g. light housework, office work
2	Symptomatic, ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours: in bed less than 50% of day
3	Symptomatic, capable of only limited self-care, confined to bed or chair more than 50% of waking hours, but not bedridden
4	Completely disabled. Cannot carry out any self-care. Totally bedridden

Serum tumour markers (Table 9.7)

Tumour markers are intracellular proteins or cell surface glycoproteins released into the circulation and detected by immunoassays. Examples are given in Table 9.7. Values in the normal range do not necessarily equate with the absence of disease and a positive result must be corroborated by histology as these markers can be seen in many benign conditions. They are most useful in the serial monitoring of response to treatment.

Cancer imaging

Radiological investigation by experts is required at various stages: at initial diagnosis and staging of the disease, during the monitoring of treatment efficacy, at the detection of recurrence, and for the diagnosis and treatment of complications.

Table 9.7 Serum tumour markers

α-Fetoprotein	Hepatocellular carcinoma, and non-seminomatous germ cell tumours of the gonads
β-Human chorionic gonadotrophin (β-HCG)	Choriocarcinomas, germ cell tumours (testicular) and lung cancers
Prostate-specific antigen (PSA)	Carcinoma of prostate
Carcinoma embryonic antigen (CEA)	Gastrointestinal cancers
CA-125	Ovarian cancer
CA-19-9	Gastrointestinal cancers particularly pancreatic cancer
CA-15-3	Breast cancer
Osteopontin	Many cancers including mesothelioma

The choice of investigations needs to be guided by the patient's symptoms and signs, site and histology of the cancer, the curative or palliative potential of treatment, and the utility of the information in guiding treatment. The investigations are described under each tumour type.

Plain X-ray allied to *ultrasonography* is quick and relatively cheap and widely available.

Cross-sectional imaging by computed tomographic scanning (*CT scan*), or magnetic resonance imaging (*MRI scan*) has greater anatomical detail, while functional imaging with radioactive nucleides for thyroid cancer is both diagnostic and therapeutic.

Contrast agents are used for increased structural discrimination and can be further enhanced with functional specificity for metabolically active tissue with ^{19}fluorodeoxyglucose uptake and CT-positron emission tomography (CT-PET scan) as used extensively in head and neck cancer, lung cancer and lymphoma. *Radionuclide imaging* of sentinel lymph nodes is used to guide lymphatic surgery in breast cancer and melanoma while paramagnetic iron particles taken up by macrophages can enable MRI lymphography for pelvic tumours. Tumour targeted contrast agents can improve detection rates such as the radiolabelled MAb rituximab for lymphoma or radiolabelled small molecules such as octreotide for neuroendocrine tumours. Research into the use of reporter agents which become visible only upon activation within the tumour environment holds the promise of greater sensitivity and specificity in the future.

CANCER TREATMENT

Aims of treatment

Cancer treatment requires the cooperation of a multidisciplinary team to coordinate the delivery of the appropriate treatment (surgery, chemotherapy, radiotherapy and biological/endocrine therapy), supportive and symptomatic care, and psychosocial support. While all members will have the patient's care as their central concern, someone, often the oncologist, has to take responsibility for the coordination of the many professionals involved. Central to this endeavour is the involvement of the patient, through education as to the nature of their disease and the treatment options available. An informed choice can then be made, even if in the end it is simply to abide by the decisions made by the professionals. Good communication embodies a humane approach which preserves hope at an appropriate level through empathy and understanding of the patient's position (see p. 12).

Curing cancer

For most solid tumours local control is necessary but not sufficient for cure because of the presence of systemic (microscopic) disease, while haematological cancers are usually disseminated from the outset. Improvement in the rate of cure of most cancers is thus dependent upon earlier detection to increase the success of local treatment and effective systemic treatment. The likelihood of cure of the systemic disease depends upon the type of cancer and its expression of appropriate treatment targets, its drug sensitivity and tumour bulk (microscopic or clinically detectable). A few rare cancers are so chemosensitive that even bulky metastases can be cured, e.g. leukaemia, lymphoma, gonadal germ cell tumours and choriocarcinoma. For most common solid tumours such as lung, breast and colorectal cancer, there is no current cure of bulky (clinically detectable) metastases, but micrometastatic disease treated by adjuvant systemic therapy (see below) after surgery can be cured in 10–20% of patients.

Adjuvant therapy for solid tumours

This is defined as treatment given, in the absence of macroscopic evidence of metastases, to patients at risk of recurrence from micrometastases, following treatment given for the primary lesion. *'Neoadjuvant' therapy* is given before primary therapy, to both shrink the tumour to improve the local excision and treat any micrometastases as soon as possible.

Micrometastatic spread by lymphatic or haematological dissemination often occurs early in the development of the primary tumour, and can be demonstrated by molecular biological methods capable of detecting the small numbers (1 in 10^6) of circulating cells. Studies correlating prognosis with histological features of the primary cancer, e.g. differentiation, or presence of early metastatic invasion of blood vessels or regional lymph nodes, have led to an increasing ability to predict which patients are at high risk of local or distant recurrence from micrometastatic disease.

Trials of treatment with local radiotherapy or endocrine, biological or chemotherapy systemic treatments have shown a significant improvement in survival in common adult cancers such as breast, bowel, prostate, head and neck, cervical cancer, choriocarcinoma and gonadal germ cell cancers. Central to these studies has been the careful selection of patients according to defined risk criteria, and the reduction of treatment toxicity to reach a balanced risk/benefit ratio. Relative risk reductions in the order of 12–25% and absolute improvements in 5- to 10-year survival of 5–10% (dependent upon the pre-existing risk) have been achieved in common epithelial cancers such as lung, bowel, breast and prostate, with greater absolute improvements of 25% in the more sensitive germ cell tumours.

While these improvements currently translate into many lives saved from common diseases at a public health level, the majority who receive such treatment do not benefit because they were already cured, or because the cancer is resistant to the treatment. Better tests in the future will identify those with the micrometastases who really need treatment. On an individual patient basis the decision on whether adjuvant treatment will be worthwhile must include consideration of other factors such as the patient's life expectancy, concurrent medical conditions and lifestyle priorities.

Treatment of metastases in the CNS

When metastatic disease has spread to the CNS it may not be accessible to conventional drug therapy. Examples of this are leukaemic infiltration of the meninges with acute lymphoblastic leukaemia, small-cell lung cancer and breast cancer. Because of the blood–brain barrier, chemotherapy agents and especially monoclonal antibodies do not enter the subarachnoid space in sufficient quantity, and are therefore ineffective in treating meningeal infiltration. In order to treat these cells, intrathecal chemotherapy and/or cranial irradiation are required for patients at risk (see p. 471).

Palliation

When cure is no longer possible, palliation, i.e. relief of tumour symptoms and prolongation of life, is possible in many cancers in proportion to their drug and radiation sensitivity. There is on average a 2- to 18-month prolongation in median life expectancy with treatments for solid tumours and up to 5–8 years for some leukaemias and lymphomas, with those with the most responsive tumours experiencing the greatest benefit. The development of more effective chemotherapeutic drugs, targeted biological agents and better supportive care such as antiemetics has done much to reduce the side-effects of systemic therapy and to improve the cost/benefit ratio for the patient receiving palliative treatment. In addition, through early assessment during treatment, it is possible to stop if there is no evidence of benefit within 6–8 weeks of starting, so as to minimize exposure to toxic and unsuccessful treatment.

Measuring response to treatment

A measurable response to treatment can serve as a useful early surrogate marker when assessing whether to continue a given treatment for an individual patient. Trials to assess response to treatment in advanced disease have identified active agents for use in the more curative setting of adjuvant treatment of early stage disease

Response to treatment can be subjective or objective. *A subjective response* is one perceived by the patient in terms of, for example, relief of pain and dyspnoea, or improvement in appetite, weight gain or energy. Such subjective response is a major aim of most palliative treatments. Quantitative measurements of these subjective symptoms form a part of the assessment of response to chemotherapy, especially in those situations where cure is not possible and where the aim of treatment is to provide prolongation of good-quality life. In these circumstances, measures of quality of life enable an estimate of the balance of benefit and side-effects to be made.

Objective response to treatment is measured either as a complete response, which is a complete disappearance of all detectable disease clinically and radiologically, or partial response, which is conventionally defined as more than a 50% reduction in the area of the tumour. The terms used to evaluate the responses of tumours are given in Box 9.2. As an alternative the RESIST convention uses a 30% reduction in the sum of all measurable lesion diameters to indicate a partial response. The term '*remission*' is often used synonymously with '*response*', which if complete means an absence of detectable disease without necessarily implying a cure of the cancer.

Box 9.2 Definitions of response

Complete response	Complete disappearance of all detectable disease
Partial response	More than 50% reduction in the product of the bidimensional diameters of the tumour
Stable disease	No change, or <50% reduction and <25% increase
Progressive disease	Increase in size of tumour by at least 25% at any site

Oncological emergencies

Superior vena caval obstruction can arise from any upper mediastinal mass but is most commonly associated with lung cancer and lymphoma. The patient presents with difficulty breathing and/or swallowing, with stridor, swollen, oedematous facies and arms with venous congestion in the neck and dilated veins in the upper chest and arms. Treatment is with immediate steroids, vascular stents, anticoagulation and mediastinal radiotherapy or chemotherapy. Some tumours, e.g. lymphomas, small-cell lung cancers and germ cell tumours, are so sensitive to chemotherapy that this is preferred to radiotherapy, as the masses are likely to be both large and associated with more disseminated disease elsewhere. An early decision is necessary on the patient's likely prognosis, as ventilatory support may be required until treatment has had time to relieve the obstruction.

Spinal cord compression (see p. 1163) needs to be rapidly diagnosed and urgent treatment arranged within 24 hours of onset of paresis to salvage as much functional capacity as possible. Early neurological clinical features may be incomplete, more subjective than objective and gradual in onset. MR scanning is the investigation of choice. Treatment should begin with high-dose steroids followed by surgical decompression and radiotherapy to the affected vertebrae to achieve the best disease control and palliation.

Neutropenic sepsis (see p. 461).

Acute lysis syndrome This occurs if treatment produces a massive breakdown of tumour cells, leading to increased serum levels of urate, potassium and phosphate. Urate deposition in the renal tubules can cause renal failure (hyperuricaemic nephropathy) requiring dialysis. The xanthine oxidase inhibitor allopurinol is given before treatment is started. Intravenous rasburicase, a recombinant urate oxidase, is occasionally used for prophylaxis and treatment but is very expensive.

Acute hypercalcaemia presents with vomiting, confusion, constipation and oliguria. Treatment is by resuscitation with intravenous fluids until a saline diuresis is established, followed by i.v. bisphosphonate, e.g. pamidronate, clodronate, or the more potent zolendronate (Emergency Box 18.2).

Raised intracranial pressure due to intracerebral metastases presents classically with headache, nausea and vomiting. There are often no localizing neurological signs and almost never papilloedema until very late in the disease. However, for many there is a slower onset with non-specific symptoms such as drowsiness or mental deterioration. Treatment is by high-dose steroids and investigation by MRI as to whether surgery is appropriate or chemotherapy and radiotherapy are required.

Hyperviscosity affects those with a very high haematocrit (>50), white cell count ($>100 \times 10^9$/L) or platelet cell count ($>1000 \times 10^9$/L) from untreated acute leukaemia, or polycythaemia. Treatment is by leucopheresis and plasmapheresis followed by chemotherapy treatment for the underlying malignancy.

Pulmonary embolus is a common complication of the coagulopathy of cancer and as a side-effect of chemotherapy.

Frequently it does not present with classical signs but as unexplained breathlessness with episodic exacerbations from multiple small emboli. A low index of suspicion should be kept in any cancer patient with hypoxia, especially if immobilized in bed in hospital. Prophylaxis with subcutaneous low molecular weight heparin is indicated in all patients if there is no contraindication from tumour bleeding.

Cardiac tamponade from a malignant pericardial effusion usually presents with a fall in cardiac output and symptoms of shortness of breath, drowsiness and hypotension that can be dramatically reversed by the ultrasound-guided percutaneous drainage of the fluid. More sustained improvement may require the surgical formation of a pericardial window and time for chemotherapy to take effect.

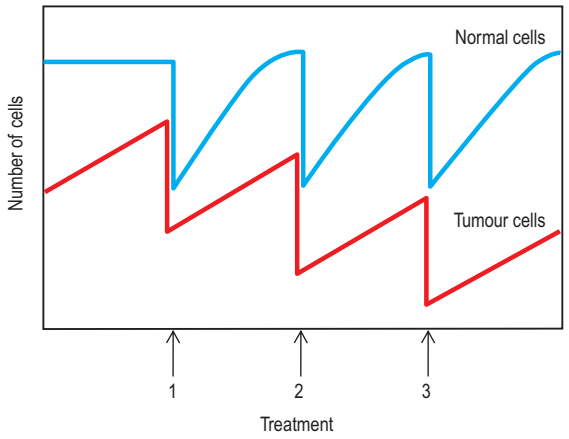


Fig. 9.3 **Effects of multiple courses of cytotoxic chemotherapy.**

PRINCIPLES OF CHEMOTHERAPY

Chemotherapy employs systemically administered drugs that directly damage cellular DNA (and RNA). It kills cells by promoting apoptosis and sometimes frank necrosis. There is a narrow therapeutic window between effective treatment of the cancer and normal tissue toxicity, because the drugs are not cancer specific (unlike some of the biological agents), and the increased proliferation in cancers is not much greater than in normal tissues (see tumour growth and failure of apoptosis, p. 27). The dose and schedule of the chemotherapy is limited by the normal tissue tolerance, especially in those more proliferative tissues of the bone marrow and gastrointestinal tract mucosa. All tissues can be affected, however, depending upon the pharmacokinetics of the drug and affinity for particular tissues (e.g. heavy metal compounds for kidneys and nerves).

The therapeutic effect on the cancer is achieved by a variety of mechanisms which seek to exploit differences between normal and transformed cells. While most of the drugs have been derived in the past by empirical testing of many different compounds, e.g. alkylating agents, the new molecular biology is leading to targeting of particular genetic defects in the cancer (see tyrosine kinase inhibitors, p. 464).

Toxicity to normal tissue can be limited in some instances by supplying growth factors such as granulocyte colony-stimulating factor (G-CSF) or by the infusion of stem cell preparations to diminish myelotoxicity. The use of more specific biological agents with relatively weak pro-apoptotic effects in combination with the general cytotoxics has also improved the therapeutic ratio (see trastuzumab and breast cancer, p. 488).

Most tumours rapidly develop resistance to single agents given on their own. For this reason the principle of intermittent combination chemotherapy was developed. Several drugs are combined together, chosen on the basis of differing mechanisms of action and non-overlapping toxicities. These drugs are given over a period of a few days followed by a rest of a few weeks, during which time the normal tissues have the opportunity for regrowth. If the normal tissues are more proficient at DNA repair than the cancer cells, it may be possible to deplete the tumour while allowing the restoration of normal tissues between chemotherapy cycles (Fig. 9.3).

In many experimental tumours it has been shown that there is a log–linear relationship between drug dose and number of cancer cells killed and that the maximum effective dose is very close to the maximum tolerated dose at which dose-limiting toxicity is reached. With a chemosensitive tumour, relatively small increases in dose may have a large effect on tumour cell kill. It is therefore apparent that where cure is a realistic option the dose administered is critical and may need to be maintained despite toxicity. In situations where cure is not a realistic possibility and palliation is the aim, a sufficient dose to exceed the therapeutic threshold, but not cause undue toxicity, is required as the short-term quality of life becomes a major consideration.

Classification of cytotoxic drugs

(Table 9.8)

DNA damaging

Alkylating agents act by covalently binding alkyl groups, and their major effect is to cross-link DNA strands, interfering with DNA synthesis and causing strand breaks. Despite being among the earliest cytotoxic drugs developed, they maintain a central position in the treatment of cancer. Melphalan is one of the original nitrogen mustards and is used in multiple myeloma. Chlorambucil is used in Hodgkin's lymphoma and chronic lymphocytic leukaemia. Other common alkylating agents include cyclophosphamide and ifosfamide,

Table 9.8 Chemotherapy: some cytotoxic drugs

DNA damaging

Free radicals – alkylators, e.g. cyclophosphamide

DNA cross-linking – platinum, e.g. cisplatin, carboplatin, oxaliplatin

Antimetabolites

Thymidine synthesis, e.g. 5-fluorouracil, methotrexate, cytarabine and mercaptopurine

DNA repair inhibitors

Topoisomerase inhibitors – epipodophyllotoxins, e.g. etoposide; campothecins, e.g. irinotecan

DNA intercalation – anthracyclines, e.g. doxorubicin

Antitubulin

Tubulin binding – alkaloids, e.g. vincristine, vinorelbine

Taxanes – e.g. paclitaxel, docetaxel

as well as the nitrosoureas, carmustine (BCNU), lomustine (CCNU) and busulfan used in chronic myeloid leukaemia. Tetrazines also alkylate DNA; dacarbazine is used in malignant melanoma and temozolomide in malignant gliomas.

Platinum compounds Cisplatin, carboplatin and oxaliplatin cause interstrand cross-links of DNA and are often regarded as non-classical alkylating agents. They have transformed the treatment of testicular cancer (cisplatin) and have a major role against many other tumours, including lung, ovarian and head and neck (cis or carbo-platin), and gastrointestinal (oxaliplatin) cancer. Toxicity, as for other heavy metals, includes renal and peripheral nerve damage.

Antimetabolites

Antimetabolites are usually structural analogues of naturally occurring metabolites that interfere with normal synthesis of nucleic acids by falsely substituting purines and pyrimidines in metabolic pathways. Antimetabolites can be divided into:

Folic acid antagonist, e.g. methotrexate. This is structurally very similar to folic acid and binds preferentially to dihydrofolate reductase, the enzyme responsible for the conversion of folic acid to folinic acid. It is used widely in the treatment of solid tumours and haematological malignancies. Folinic acid is often given to 'rescue' normal tissues from the effects of methotrexate.

Pyrimidine antagonists. 5-Fluorouracil (5-FU) consists of a uracil molecule with a substituted fluorine atom. It acts by blocking the enzyme thymidylate synthase, which is essential for pyrimidine synthesis. 5-Fluorouracil has a major role in the treatment of solid tumours, particularly gastrointestinal cancers. Oral capecitabine is metabolized to 5-FU, and tegafur with uracil and calcium folinate are used in gastrointestinal and breast cancers.

Arabinosides inhibit DNA synthesis by inhibiting DNA polymerase. Cytosine arabinoside (cytarabine) is used almost exclusively in the treatment of acute myeloid leukaemia where it remains the backbone of therapy, while its analogue gemcitabine is proving useful in a number of solid cancers such as lung, breast, pancreas and ovary. Fludarabine is used in the treatment of B cell chronic lymphocytic leukaemia; it is also used in reduced intensity stem cell transplantation (see p. 473) because of its immunosuppressive effect.

Purine antagonists, e.g. 6-mercaptopurine and 6-thioguanine, which are both used almost exclusively in the treatment of acute leukaemia.

DNA repair inhibitors

Epipodophyllotoxins. These are semisynthetic derivatives of podophyllotoxin which inhibit topoisomerase. Topoisomerases allow unwinding and uncoiling of supercoiled DNA. Etoposide is a drug used in a wide range of cancers and works by maintaining DNA strand breaks by inhibiting the enzyme topoisomerase II. Topoisomerase I inhibitors such as irinotecan and topotecan have also proved active against lung, colon, ovary and cervix cancer.

Cytotoxic antibiotics. The anthracyclines act by intercalating adjoining nucleotide pairs on the same strand of DNA and by inhibiting topoisomerase II DNA repair. They have a wide spectrum of activity in haematological and solid tumours. Doxorubicin and its cogener epirubicin are two of the most widely used of all cytotoxic drugs but have cumulative toxicity to the myocardium. Pegylated liposomal doxorubicin is used for Kaposi's sarcoma and as second-line treatment for advanced ovarian cancer with reduction of cardiotoxicity, but increased toxicity to the skin on the palms of the hands and soles of the feet. Amsacrine is a similar drug used occasionally in acute myeloid leukaemia. Bleomycin and mitomycin are also intercalating agents which promote the cleavage of DNA and RNA. Bleomycin has a particular toxicity to the lung causing interstitial fibrosis.

Antitubulin agents

Vinca alkaloids. Drugs such as vincristine, vinblastine and vinorelbine act by binding to tubulin and inhibiting microtubule formation (see p. 24). They are used in the treatment of haematological (vincristine and vinblastine) and non-haematological cancers (vinorelbine). They are associated with neurotoxicity due to their anti-microtubule effect and must never be given intrathecally.

Taxanes. Paclitaxel and docetaxel bind to tubulin dimers and prevent their assembly into microtubules. They are active drugs against many cancers such as ovarian, breast and lung cancer. Taxanes can cause neurotoxicity and hypersensitivity reactions and patients should be premedicated with steroids, H_1 and H_2 histamine antagonists prior to treatment.

Side-effects of chemotherapy

Chemotherapy carries many potentially serious side-effects and should be used only by trained practitioners. The four most common side-effects are vomiting, hair loss, tiredness and myelosuppression (Table 9.9). Side-effects are much more directly dose related than anticancer effects and it has been the practice to give drugs at doses close to their maximum tolerated dose, although this is not always necessary to achieve their maximum anticancer effect. Common combination chemotherapeutic regimens are shown in Table 9.10.

Extravasation of intravenous drugs. Cytotoxic drugs should only be given by trained personnel. They cause severe local tissue necrosis if leakage occurs outside the vein. Stop the

Table 9.9 Side-effects of chemotherapy

Common
Nausea and vomiting
Hair loss
Myelosuppression
Mucositis
Fatigue
Drug-specific
Cardiotoxicity, e.g. anthracyclines
Pulmonary toxicity, e.g. bleomycin
Neurotoxicity, e.g. cisplatinum, vinca alkaloids, taxanes
Nephrotoxicity, e.g. cisplatinum
Skin plantar–palmar dermatitis, e.g. 5-fluorouracil
Sterility, e.g. alkylating agents, anthracyclines, docetaxel
Secondary malignancy, e.g. alkylating agents, epipodophyllotoxins

Table 9.10 Some common chemotherapy regimens

Hodgkin's lymphoma	ABVD	Doxorubicin, bleomycin, vinblastine, dacarbazine
	BEACOPP	Bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisolone
Non-Hodgkin's lymphoma	CHOP	Cyclophosphamide, hydroxy-doxorubicin, vincristine, prednisolone
Breast	CMF	Cyclophosphamide, methotrexate, 5-fluorouracil
	EC-T	Epirubicin, cyclophosphamide, paclitaxel
	FEC-D	5FU, epirubicin, cyclophosphamide, docetaxel
Lung	PE	Cisplatin, etoposide
	GC	Gemcitabine, carboplatin
Stomach	ECF	Epirubicin, cisplatin, 5-fluorouracil
Colorectal	FolFOx	Oxaliplatin, 5FU, folinic acid
	OX	Oxaliplatin, capecitabine

NB. Some abbreviations are related to trade names.

infusion immediately and institute local measures, e.g. aspirate as much of the drug from the cannula, infiltrate area with 0.9% saline and apply warm compresses. Antihistamines and corticosteroids may give symptomatic relief. Dexrazoxane is used for anthracycline-induced extravasation as an antidote.

Nausea and vomiting. The severity of this common side-effect varies with the cytotoxic and it can be eliminated in 75% of patients by using modern antiemetics. Nausea and vomiting are particular problems with platinum analogues and with doxorubicin. A stepped policy with antiemetics such as metoclopramide and domperidone followed by 5-HT_3 serotonin antagonists (e.g. ondansetron, granisetron) combined with dexamethasone should be used to match the emetogenic potential of the chemotherapy. Aprepitant, a neurokinin receptor antagonist, is helpful in preventing acute and delayed nausea and vomiting. It is used with dexamethasone and a 5-HT_3 antagonist. Drugs such as cyclizine, haloperidol and levomepromazine, and benzodiazepines can be used to control persistent nausea.

Hair loss. Many but not all cytotoxic drugs are capable of causing hair loss. Scalp cooling can sometimes be used to reduce hair loss but in general this side-effect can only be avoided by selection of drugs where this is possible. Hair always regrows on completion of chemotherapy.

Bone marrow suppression and immunosuppression. Suppression of the production of red blood cells, white blood cells and platelets occurs with most cytotoxic drugs and is a dose-related phenomenon. Severely myelosuppressive chemotherapy may be required if treatment is to be given with curative intent despite the potential for rare but fatal infection or bleeding. Anaemia and thrombocytopenia are managed by red cell or platelet transfusions.

Box 9.3 Febrile neutropenia treatment

IMMEDIATE INTERVENTION IS ESSENTIAL

Resuscitation with intravenous fluids to restore circulatory function, e.g. urine output, followed by cultures of blood, urine, sputum and stool and empirical antibiotics:

- Commonly used antibiotics should include activity against pseudomonas, e.g. ceftazidime or ticarcillin with gentamicin
- May require antibiotics against *Staph. aureus* especially with indwelling venous access lines, e.g. flucloxacillin or vancomycin for MRSA.

If the patient deteriorates clinically and/or temperature still elevated after 48 hours, change antibiotics according to culture results or empirically increase Gram-negative and Gram-positive cover.

- Consider adding treatment for opportunistic infections if fever not responding to broad-spectrum antibiotics, e.g. amphotericin B
  - liposomal amphotericin B or voriconazole (latter two are very expensive) or caspofungin – fungus
  - high-dose co-trimoxazole – pneumocystis
  - clarithromycin – mycoplasma.

Neutropenic patients are at high risk of bacterial and fungal infection, often from enteric bowel flora. Patients must be warned of the possibility of neutropenia occurring. Those with a fever > 37.5°C and less than 0.5×10^9 neutrophils/L are managed by the immediate introduction of broad-spectrum antibiotics intravenously for the treatment of infection (Box 9.3). Initial empirical therapy should be reviewed following microbiological results. Prophylactic antibiotic levofloxacin during the period of myelosuppression has been shown to reduce infection in solid tumour patients but can raise the risk of *Clostridium difficile* colitis. Haemopoietic growth factors such as G-CSF and peripheral blood stem cells can reduce the duration of neutropenia significantly, benefiting patients at high risk of infectious complications.

Mucositis. This common side-effect of chemotherapy reflects the sensitivity of the mucosa to antimitotic agents. It causes severe pain in the oropharyngeal region and problems with swallowing and nutrition. Mucositis can be generalized throughout the intestinal track when it can cause life threatening diarrhoea. Treatment is with antiseptic and anticandidal mouthwash and, if severe, fluid and antibiotic support, as the mucosa is a portal for entry of enteric organisms. Folinic acid is used to prevent and treat mucositis from methotrexate. A keratinocyte growth factor (palifermin) is helpful.

Cardiotoxicity. This is a rare side-effect of chemotherapy, usually associated with anthracyclines such as doxorubicin. It is dose-related and can largely be prevented by restricting the cumulative total dose of anthracyclines within the safe range (equivalent to 450 mg/m^2 body surface area cumulative doxorubicin dose). Cardiotoxicity can also be reduced by using the analogue epirubicin or by reducing peak drug concentrations through delayed release preparations such as liposomal doxorubicin or employing fractionated administration schedules. Whether these strategies affect the antitumour efficacy is still debated. 5-Fluorouracil and its prodrug capecitabine cause cardiac ischaemia.

FURTHER READING

Hesketh PJ. Chemotherapy induced nausea and vomiting. *New England Journal of Medicine* 2008; **358**: 2482–2492.

Neurotoxicity. This occurs predominantly with the vinca alkaloids, taxanes and platinum analogues (but not carboplatin). It is dose-related and cumulative. Chemotherapy is usually stopped before the development of a significant polyneuropathy, which once established is only partially reversible. Vinca alkaloids such as vincristine must never be given intrathecally as the neurological damage is progressive and fatal.

Nephrotoxicity. Cisplatin (but not oxaliplatin or carboplatin), methotrexate and ifosfamide can potentially cause renal damage. This can usually be prevented by maintaining an adequate diuresis during treatment to reduce drug concentration in the renal tubules and careful monitoring of renal function.

Sterility and premature menopause. Some anticancer drugs, particularly alkylating agents, but also anthracyclines, and docetaxel may cause gonadal damage resulting in sterility and in women the loss of ovarian oestrogen production, which may be irreversible.

In males the storage of sperm prior to chemotherapy should be offered to the patient when chemotherapy is given with curative intent.

In females it may be possible to collect oocytes to be fertilized in vitro and cryopreserved as embryos for subsequent implantation. Cryopreservation of ovarian tissue and retrieval of viable oocytes for subsequent fertilization is still experimental. The recovery of gonadal function is dependent upon the status before treatment, and in women this is mostly related to age since menarche, with those under the age of 40 having significantly more ovarian reserve.

Secondary malignancies. Anticancer drugs have mutagenic potential and the development of secondary malignancies, predominantly acute leukaemia, is an uncommon but particularly unwelcome long-term side-effect in patients otherwise cured of their primary malignancies. The alkylating agents, anthracyclines, and epipodophyllotoxins are particularly implicated in this complication.

Drug resistance

Drug resistance is one of the major obstacles to curing cancer with chemotherapy. Some tumours have an inherently low level of resistance to currently available treatment and are often cured. These include gonadal germ cell tumours, Hodgkin's lymphoma and childhood acute leukaemia. Solid tumours such as small-cell lung cancer, breast cancer and ovarian cancer initially appear to be chemosensitive, with the majority of patients responding, but most patients eventually relapse with resistant disease. In oesophageal, gastric and colorectal cancers there is an intermediate degree of sensitivity while in other tumours such as pancreatic cancer, non-small-cell lung cancer and melanoma the disease is largely chemoresistant from the start.

Most resistance occurs as a result of genetic mutation and becomes more likely as the number of tumour cells increases. It has also been shown that anticancer drugs can themselves increase the rate of mutation to resistance. Resistance to cytotoxic drugs is often multiple and is then known as multidrug resistance (MDR), e.g. resistance to doxorubicin is often associated with resistance to vinca alkaloids and epipodophyllotoxins, and is mediated through increased expression of P-glycoprotein (a 170 kDa membrane phosphoglycoprotein), which mediates the efflux of cytotoxic drugs out of the cells. Many other mechanisms may also be involved in resistance to chemotherapy, such as the upregulation of anti-apoptotic proteins Bcl-2 and Bax and DNA repair pathways.

High-dose therapy

Most anticancer drugs have a sigmoid dose–response relationship which suggests that, up to a point, a higher dose of a cytotoxic drug will induce a greater response. However, increasing cytotoxic drug doses is often not possible, owing to toxicity. For those chemotherapeutic agents with a dose-limiting toxicity of bone marrow failure, infusion of previously harvested haemopoietic stem cells is able to restore the lymphohaemopoietic system and permit the use of higher doses to overcome resistance.

Haemopoietic stem cell transplantation

Bone marrow or peripheral blood cells may be:

- autologous – from self
- syngeneic – from identical twin
- allogeneic – from non-identical donor (matched or sometimes mismatched)
- from umbilical cord blood. This is increasingly being used for adult and childhood leukaemia.

It usually takes 2–3 weeks for engraftment to take and during this time patients need supportive care with nursing in isolated cubicles with air filtration.

Autologous stem cell transplantation

Autologous stem cells are used as rescue from myeloablative chemotherapy. Haemopoietic stem cells are collected from the patients' bone marrow, or more commonly by leucopheresis from peripheral blood following administration of the growth factor granulocyte colony-stimulating factor (G-CSF) prior to chemotherapy. They are stored by cryopreservation. These cells are then re-infused intravenously after an intensive, myeloablative chemotherapy regimen. This approach has been particularly effective in relapsed leukaemias, lymphomas, myeloma and germ cell tumours. There is no risk of graft rejection or graft-versus-host disease (GVHD) but the graft-versus-tumour effect is lost (see below).

Allogeneic stem cell transplantation

Allogeneic transplants combine the cytotoxic effect of high-dose therapy with a potent immunotherapy effect. Historically, the transplantation of donor haemopoietic cells has been combined with myeloablative chemotherapy ± radiotherapy with the dual effects of treating the malignancy as well as immunosuppression that allows the graft 'to take'. Anti-T cell antibodies are often given to reduce graft-versus-host disease and immune-related infection. The engraftment of the donor immune system, with antitumour activity (graft versus tumour), is primarily responsible for the increased effectiveness of this approach. In general, ideal donors are fully matched at the major HLA antigens. Thus siblings are more likely to be found to be potential donors than unrelated volunteers. Some degree of HLA antigen mismatch may be

tolerated in children, but is problematic in adults. Allogeneic transplantation has been successfully used in acute and chronic leukaemias, and myeloma.

Complications include '*graft-versus-host disease*', an immune reaction of the donor cells against normal host organs, which can affect 30–50% of transplant recipients and is potentially fatal in some cases. Immunosuppression, both from conditioning therapy and from the immunosuppressive drugs (ciclosporin or tacrolimus) given to prevent graft-versus-host disease, results in a high incidence of opportunistic infections. All patients receive prophylactic antibacterial, antifungal and antiviral drugs but infections still occur. *Mortality* therefore from conventional allogeneic stem cell transplantation is a major problem, with 20–40% at risk of dying from the procedure, depending on the age and status of the recipient, and the degree of HLA compatibility of the donor (see also immunotherapy, p. 75).

Non-myeloablative allogeneic stem cell transplantation. This has been developed using conditioning therapy with drugs such as fludarabine which is sufficiently immunosuppressant without radiation therapy. This maintains the anticancer effect of the stem cell transplantation without the complications of conventional allogeneic stem cell transplantation. Mortality (mainly GVHD) is lower, and the technique is being used more widely, particularly in the elderly.

PRINCIPLES OF ENDOCRINE THERAPY

Oestrogens are capable of stimulating the growth of breast and endometrial cancers, and androgens the growth of prostate cancer. Removal of these growth factors by manipulation of the hormonal environment may result in apoptosis and regression of the cancer. Endocrine therapy can be curative in a proportion of patients treated for micrometastatic disease in the adjuvant setting for breast and prostate cancer and provides a minimally toxic non-curative (palliative) treatment in advanced/metastatic disease. The presence of detectable cellular receptors for the hormone is strongly predictive of response. However this is also modified by the many molecular interactions between the activation pathways of, for example, EGFR and oestrogen receptor (ER). Figure 9.4 shows the binding of the hormone to the receptor.

Anti-oestrogens and progestogens

In premenopausal women hormonal manipulation can be achieved by the reduction of endogenous oestrogen by oophorectomy or via pituitary downregulation using a gonadotrophin-releasing hormone (GnRH) analogue such as goserelin.

Tamoxifen is a mixed agonist and antagonist of oestrogen action on the ER while the more recent drug fulvestrant is a more selective oestrogen receptor modulator (SERM). Tamoxifen is used as an adjuvant therapy in breast cancer and both are used in advanced metastatic breast disease (see p. 487).

Synthetic progestogens, e.g. medroxyprogesterone acetate and megestrol acetate, have a direct effect on breast tumour cells through progesterone receptors, as well as effects on the pituitary/ovarian (premenopausal) and adrenal/pituitary axis (postmenopausal). They can be as effective as tamoxifen in metastatic breast cancer and have a role in endometrial adenocarcinoma, with response rates of up to 25%. Trials to date with adjuvant progestogens have not been successful in increasing survival.

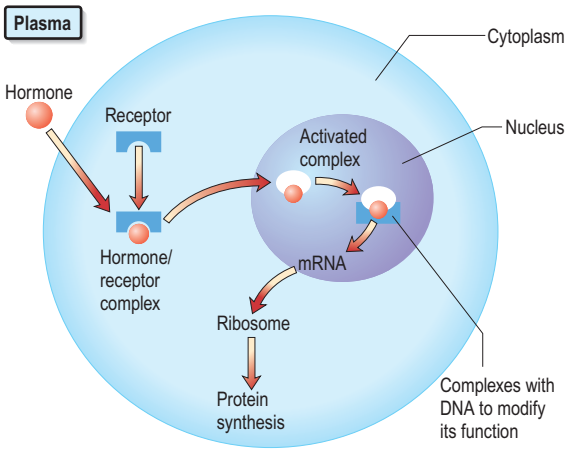


Fig. 9.4 **Mechanism of the interaction between a steroid hormone and its receptor.** This interaction modifies DNA activity and hence cell growth and replication.

In postmenopausal women, androgens are synthesized by the adrenal glands and converted in subcutaneous fat to estrone by the enzyme aromatase. *Aromatase inhibitors*, for example anastrozole, letrozole and exemestane, reduce circulating oestrogen levels and oestrogen synthesis in tumour cells and have shown greater efficacy than tamoxifen in the treatment of metastatic breast cancer (see p. 487) and equivalence in the adjuvant setting.

Side-effects are those of oestrogen deprivation hot flushes, emotional lability, vaginal dryness and loss of libido, plus specific effects such as tamoxifen coagulopathy and endometrial hyperplasia, progestogen fluid retention, and aromatase inhibitor arthralgia and osteoporosis.

Androgen deprivation

- *GnRH agonists*, e.g. goserelin and leuprorelin, and orchidectomy, are equally effective in lowering circulating androgens and inducing responses in prostate cancer. However, in the first week this produces a rise in LH and testosterone which can result in a tumour flare. Therefore goserelin is given in combination with an antiandrogen, e.g. flutamide, initially. In the adjuvant setting, the addition of androgen deprivation to prostatic radiotherapy or surgery has improved survival.
- *Androgen receptor blockers* such as flutamide and abiraterone, an inhibitor of CYP17 which is necessary for androgen production, are also effective.

PRINCIPLES OF BIOLOGICAL THERAPY

Interferons

Interferons are naturally occurring cytokines that mediate the cellular immune response. They are both antiproliferative and stimulate humoral and cell-mediated immune responses to the tumour that can result in an antitumour effect if the host effector mechanisms are fully competent.

Alpha-interferon (IFN-α) has been used to treat advanced melanoma and renal cell carcinoma. In renal cell carcinoma alpha-interferon has a low (10–15%) but significant anticancer effect and prolongation of survival.

Treatment with IFN has side-effects (see p. 341), most commonly flu-like symptoms which tend to diminish with time, and fatigue which generally does not and can be treatment limiting. IFN was given as a daily subcutaneous injection, but conjugation with polyethylene glycol (PEG interferon) has led to a reduction in frequency of injection and severity of side-effects.

Interleukins

Originally described for their activity in modulating leucocyte activation, these cytokines have widespread activity in coordinating cellular activity in many organs. Interleukin 2, a recombinant protein, is used to activate T cell responses, often in conjunction with interferon-stimulated B cell activation. Antitumour activity has been observed in renal cell carcinoma and melanoma with responses in 10–20% of patients, occasionally for prolonged periods. Toxicity is common; acutely this includes the capillary leak syndrome with hypotension, and pulmonary oedema, whilst autoimmune thyroiditis and vitiligo occur later.

Haemopoietic growth factors

Erythropoietin is given for anaemia. Granulocyte (G-CSF) or granulocyte–macrophage colony-stimulating factor (GM-CSF) are used:

- to reduce the duration of cytopenia following chemotherapy
- with or without chemotherapy, to stimulate the proliferation of haemopoietic progenitor cells in the marrow so that they enter the circulation and can be collected from the peripheral blood to support high-dose chemotherapy treatment, particularly in myeloma (see p. 486).

Monoclonal antibodies (Table 9.11)

Monoclonal antibodies (Mab) directed against tumour cell surface antigens are 'humanized' by being genetically engineered as a chimera comprising a human constant region with the murine heavy and light chains of the antigen-combining site, to reduce formation of blocking human anti-mouse antibodies when used in patients. Uses include:

- In vitro, in conjunction with complement, they are occasionally used to deplete autologous bone marrow of tumour cells in patients with leukaemia and lymphoma receiving high-dose treatment with autologous haemopoietic progenitor cell support.
- In vitro in immunoadsorption columns to select the CD34-positive (stem cell) fraction from peripheral blood progenitor cell or autologous bone marrow collections, to support high-dose treatment in haematological and other malignancies. (Used rarely.)
- As direct treatment for B cell lymphoid malignancy (e.g. rituximab anti-CD20 surface antigen). Tumour cell lysis occurs by both complement- and antibody-dependent cellular cytotoxicity.
- As a carrier molecule to target toxins or radioisotopes to the tumour cells, e.g. anti-CD20 conjugated to radioactive iodine is being used as treatment for non-Hodgkin's lymphoma.
- As anti-growth factor agents added to chemotherapy. They act by inhibiting dimerization of the extracellular receptor molecules, e.g. *trastuzumab* (a member of the epidermal growth factor receptor (EGFR) family), against the Her2/Neu or c-erbB2 antigen to increase the apoptotic response to cytotoxics in breast cancer. They are also used with *bevacizumab* anti VEGF receptor in colorectal and breast cancer, and *cetuximab* anti EGFR for head and neck and colorectal cancers.

Table 9.11 Some biological therapies

Drug	Function	Malignancy
Dasatinib	Tyrosine kinase inhibitor	AML, ALL
Imatinib	Tyrosine kinase inhibitor	Stromal cell tumour, chronic myeloid leukaemia
Sunitinib	Tyrosine kinase inhibitor	Renal cancer
Sorafenib	Tyrosine kinase inhibitor	Renal cancer
Gefitinib	EGF tyrosine kinase inhibitor	Non-small-cell lung cancer, pancreatic cancer
Bexarotene	Retinoid X receptor agonist	Cutaneous T-cell lymphoma
Erlotinib	EGF tyrosine kinase inhibitor	Non-small-cell lung cancer
Cetuximab	Anti-EGF receptor	Colorectal cancer
Bevacizumab	Anti-VEGFR	Colorectal cancer
Rituximab	Anti-CD20 on B cells	Lymphoma
Alemtuzumab	Anti-CD52 on lymphocytes	Chronic lymphocytic leukaemia
Trastuzumab	Binds to HER-2/neu receptor	Breast cancer
Bortezomib	Proteosome inhibitor	Leukaemia, myeloma

EGF, epidermal growth factor; VEGF, vascular endothelial growth factor.

Side-effects are those of hypersensitivity to the foreign protein, and specific cross-reactivities, e.g. trastuzumab for the myocardium, bevacizumab for the mucosa and renal tubule, and cetuximab for the skin follicles.

Intracellular signal inhibitors (Table 9.11)

Many cancer cells are transformed by the activity of the protein products of oncogenes that signal growth by phosphorylation of tyrosine residues on the intracellular portion of growth factor receptors. Small molecule inhibitors have many pharmacokinetic advantages over the Mab inhibitors. The first example was the tyrosine kinase inhibitor (TKi) imatinib, which specifically inhibits the BCR-ABL fusion oncoprotein cKit. This compound is an extremely effective treatment for chronic myeloid leukaemia, and gastrointestinal stromal tumours (GIST) which are also characterized by the presence of the cKit target. The less specific TKi's sunitinib and sorafenib which inhibit signalling by EGFR and VEGFR have proved effective in metastatic renal cancer, while lapatinib which inhibits Her2 has increased survival in metastatic breast cancer. Erlotinib and gefitinib on the other hand have shown only limited activity in lung cancer. Many other similar molecules are in preclinical or early clinical development.

Immunotherapy

Activation of the immune system using Bacille Calmette–Guérin (BCG) for bladder cancer induces responses in 60% of patients. Certain antigens that are specific to cancer cells, such as sequences of tumour immunoglobulin from B cell lymphomas or melanoma antigens, have been used as tumour vaccines. Antigen-presenting cells (dendritic cells) from the patient can be genetically engineered to present both antigen and cytokines such as interleukin 2 or granulocyte–macrophage colony-stimulating factor, and clinical responses have been observed. Another approach has used dendritic cells to further improve the vaccination strategy by engineering them to display the full range of HLA and B7 costimulatory molecules.

The use of non-myeloablative haemopoietic stem cell and donor lymphocyte infusions with allogeneic BMT, while losing some of the specificity, has produced the strongest evidence for the efficacy of immunotherapy via graft versus tumour activity at the risk of great toxicity.

Gene therapy

Antisense oligonucleotides are short sequences of DNA bases which specifically inhibit complementary sequences of either DNA or RNA. As a result, they can be generated against genetic sequences which are specific for tumour cells. Their clinical development has been hampered by poor uptake by tumour cells and rapid degradation by natural endonucleases. However, one antisense sequence directed against the *Bcl-2* oncogene has been shown to have an antitumour effect in patients with non-Hodgkin's lymphoma. Viral vectors for the transfection of tumour cells in vivo are being tested as a way of delivering specific replacement gene therapy in head and neck cancers.

PRINCIPLES OF RADIATION THERAPY

Radiation delivers energy to tissues, causing ionization and excitation of atoms and molecules. The biological effect is exerted through the generation of single- and double-strand DNA breaks, inducing apoptosis of cells as they progress through the cell cycle, and through the generation of short-lived free radicals, particularly from oxygen, which damage proteins and membranes. This partly explains why a lack of oxygen confers some degree of radio resistance to cells.

External beam (or teletherapy) from a linear accelerator source produces X-rays. The energy is transmitted as photons and is the most commonly used form of radiotherapy. Cobalt-60 generators can also provide gamma rays and high-energy photons, but are being gradually phased out.

Brachytherapy is the use of radiation sources in close contact with the tissue to provide intense exposure over a short distance to a restricted volume.

Systemic radionuclides, e.g. iodine-131, or radioisotope-labelled monoclonal antibodies and hormones, can be administered by intravenous or intracavitary routes to provide radiation targeted to particular tissue uptake via surface antigens or receptors.

The radiation dose is measured in Gray (Gy), where 1 Gray = 1 joule absorbed per kilogram of absorbing tissue and 1 centigray (cGy) = 1 rad. The biological effect is dependent upon the dose rate, duration, volume irradiated and the tissue sensitivity. Sensitivity to photon damage is greatest during the G2–M phase of the cell cycle and is also dependent upon the DNA repair capacity of the cell. *Fractionation* is the delivery of the radiation dose in increments separated by at least 4–6 hours to try to exploit any advantage in DNA repair between normal and malignant cells. Radiation dose is thus described by three factors:

- total dose in Gy
- number of fractions
- time for completion.

Most external beam treatments that are given with curative intent are delivered in 1.5 to 2 Gy fractions daily for 5 days per week. *Hyperfractionation* is when more than one fraction per day is given and this approach has been shown to improve outcome in head and neck and lung cancer. The treatment can also be accelerated, i.e. the total dose is given in a shorter overall time. For example a standard curative treatment taking 6.5 weeks can be accelerated so that the same dose is delivered in 5.5 weeks.

The radiation effect will also depend upon the intensity of the radiation source, measured as the linear energy transfer or frequency of ionizing events per unit of path, which is subject to the inverse square law as the energy diminishes with the distance from the source.

The generation of free radicals depends upon the degree of oxygenation/hypoxia in the target tissues. This can affect the biological effect by up to threefold and is the subject of continuing research for hypoxic cell sensitizers to overcome the reduced efficacy of radiation for hypoxic tumours. Hypoxia, however, may also drive a more malignant potential further, so reduced efficacy is only part of the solution to the hypoxia problem.

The depth of penetration of biological tissues by the photons depends upon the energy of the beam. Low-energy photons from an 85 kV source are suitable for superficial treatments, while high-energy 35 MV sources produce a beam with deeper penetration, less dose at the initial skin boundary (skin sparing), sharper edges, and less absorption by bone. Superficial radiation may be also delivered by electron beams from a linear accelerator that has had the target electrode that generates the X-rays removed.

Radiotherapy treatment planning involves both detailed physics of the applied dose and knowledge of the biology of the cancer and whether the intention is to treat the tumour site alone, or include the likely loco-regional patterns of spread. Normal tissue tolerance will determine the extent of the side-effects and therefore the total achievable dose. A balanced decision is made according to the curative or palliative intent of the treatment and the likely early or late side-effects.

The cancers for which radiotherapy is usually employed as primary curative when the tumour is anatomically localized are listed in Table 9.12, along with those in which radiotherapy has curative potential when used in addition to surgery (adjuvant radiotherapy). Palliative treatments are frequently used to provide relief of symptoms to improve quality if not duration of survival (Box 9.4). Palliative treatment is usually given in as few fractions as possible over as short a time as possible.

Radiotherapy planning, by the use of CT scanning guidance, has been complemented by the introduction of

Table 9.12 Curative radiotherapy treatment

Primary modality	Adjuvant to primary surgery
Retina	Lung
CNS	Breast
Skin	Uterus
Pharynx and larynx	Bladder
Cervix and vagina	Rectum
Prostate	Testis-seminoma
Lymphoma	Sarcoma

Box 9.4 Palliative benefits of radiotherapy

- Pain relief, e.g. bone metastases
- Reduction of headache and vomiting in raised intracranial pressure from CNS metastases
- Relief of obstruction of bronchus, oesophagus, ureter, and lymphatics
- Preservation of skeletal integrity from metastases in weight-bearing bones
- Reversal of neurological impairment from spinal cord or optic nerve compression by metastases

intensity modulated radiotherapy (IMRT) which can deliver curved dose distributions to enable an improved therapeutic ratio. This allows a greater differential in dose between the tumour and critical normal structures, in turn allowing dose escalation or a reduced risk of toxicity.

Combination chemo-radiotherapy

The local efficacy of radiotherapy can be increased by the simultaneous but not serial addition of chemotherapy with agents such as cisplatin, mitomycin and 5FU for cancers of the head and neck, lung, oesophagus, stomach, rectum, anus and cervix. Reduced local recurrence rates have translated into survival benefits and further research is investigating the concurrent use of biological agents (e.g. epidermal growth factor receptor inhibitors) with radiation.

Side-effects of radiotherapy

Early radiotherapy side-effects may occur within days to weeks of treatment when they are usually self-limiting but associated with general systemic disturbance (Table 9.13). The side-effects will depend upon tissue sensitivity, fraction size and treatment volume and are managed with supportive measures until normal tissue repair occurs. The toxicity may also be enhanced by exposure to other radiation-sensitizing agents, especially some cytotoxics, e.g. bleomycin, actinomycin, anthracyclines, cisplatin and 5-fluorouracil.

Later side-effects occur from months to years later, unrelated to the severity of the acute effects because of their different mechanism. Late effects reflect both the loss of slowly proliferating cells and a local endarteritis which produces ischaemia and proliferative fibrosis. The risks of late side-effects are related to the fraction size and total dose delivered to the tissue.

Table 9.13 Side-effects of radiotherapy

Acute temporary side-effects/dependent on region being treated	
Anorexia, nausea, malaise	
Mucositis, oesophagitis, diarrhoea	
Alopecia	
Myelosuppression	
Late side-effects	
Skin	Ischaemia, ulceration
Bone	Necrosis, fracture, sarcoma
Mouth	Xerostomia, ulceration
Bowel	Stenosis, fistula, diarrhoea
Bladder	Fibrosis
Vagina	Dyspareunia, stenosis
Lung	Fibrosis
Heart	Pericardial fibrosis, cardiomyopathy
CNS	Myelopathy
Gonads	Infertility, menopause
Second malignancies	e.g. Leukaemia, cancer, e.g. thyroid

Growth may be arrested if bony epiphyses are not yet fused and are irradiated, leading to distorted skeletal growth in later life.

Secondary malignancies following radiotherapy may appear 10–20 years after the cure of the primary cancer. Haematological malignancies tend to occur sooner than solid tumours from the irradiated tissues. The latter are very dependent upon the status of the tissue at the time of treatment, e.g. the pubertal breast is up to 300 times more sensitive to malignant transformation than the breast tissues of a woman in her thirties. Patients who smoke are more liable to develop lung cancer. Treatment of these secondary cancers can be successful providing there is normal bone marrow to reconstitute the haemopoietic system or the whole tissue at risk (e.g. thyroid after mantle radiotherapy for lymphoma) can be resected.

HAEMATOLOGICAL MALIGNANCIES

The leukaemias, the lymphomas and multiple myeloma are an interrelated spectrum of malignancies of the myeloid and lymphoid systems. They are uncommon but not rare, the lymphomas alone being the fifth commonest cancer in the UK. The aetiology of these diseases for the most part is unknown, although viruses, irradiation, cytotoxic poisons and immune suppression have been implicated in a small proportion of cases (see p. 467). The pathogenesis involves at least one or usually more molecular abnormalities, and non-random chromosomal abnormalities have been detected in several leukaemias and lymphomas. Classification has become increasingly complex, with the universally applied WHO scheme demanding morphological, cytogenetic and sometimes molecular criteria to be fulfilled. Treatment options are multiple. Patients need to be supported through treatment involving prolonged myelosuppression and immuno-

www.studentconsult.com
Ask the authors

suppression. These are potentially life-threatening but can also be curative. This has given rise to the need for highly skilled staff and specialist facilities; patients should be referred to these centres for treatment.

In the management of these diseases it is critical that patients are appraised of the natural history, its potential modification by treatment and the risks of both severe morbidity and mortality. It must be made clear from the outset whether a curative or palliative strategy is most appropriate and why. If cure is to be pursued, the patient must be appraised of the approximate probability of success and its potential price. The possibility of failure needs to be addressed at the outset and not at the last minute.

THE LEUKAEMIAS

These are relatively uncommon diseases with an incidence of about 10 per 100 000 per year; which can occur at any age. They are classified as being acute (short natural history) or chronic (long natural history), and of myeloid or lymphoid origin. More than half of the leukaemias present acutely (ALL, AML) with the remainder being chronic types (CLL, CML). The type of leukaemia varies with age; acute lymphoblastic leukaemia (ALL) is mainly seen in childhood and chronic lymphocytic leukaemia (CLL) is a disease of the elderly. The myelodysplastic syndromes are considered pre-leukaemic and are discussed on page 422. Leukaemia can be diagnosed by examination of a stained slide of peripheral blood and bone marrow, but immune phenotyping, cytogenetics and molecular genetics are essential for complete subclassification and prognostication.

General classification

The characteristics of leukaemic cells can be assessed by light microscopy, expression of cytosolic enzymes and expression of surface antigens. These will reflect the lineage and degree of maturity of the leukaemic clone. Thus, leukaemia can be divided on the basis of the speed of evolution of the disease into acute or chronic; each of these is then further subdivided into myeloid or lymphoid, according to the cell type involved.

- acute myeloid leukaemia (AML)
- acute lymphoblastic leukaemia (ALL)
- chronic myeloid leukaemia (CML)
- chronic lymphocytic leukaemia (CLL).

Aetiology

In the majority of patients this is unknown but several factors have been associated:

- *Radiation*. This can induce genetic damage to haemopoietic precursors and ALL, AML and CML have been seen in increased incidences in survivors of Hiroshima and Nagasaki and in patients treated with ionizing radiation.
- *Chemical and drugs*. Exposure to benzene used in industry may lead to marrow damage. AML occurs after treatment with alkylating agents (e.g. melphalan) and topoisomerase II inhibitors (e.g. etoposide).
- *Genetic*. Leukaemia risk is highly elevated in a number of germline conditions that result in genetic instability or bone marrow failure. These include Fanconi anaemia, ataxia telangiectasia and Li–Fraumeni syndrome. The risk is elevated some 30 times in people with trisomy 21. There is a high degree of concordance among monozygotic twins.
- *Viruses*. Leukaemias are associated with human T cell lymphotropic virus type 1 (HTLV-1), which is found particularly in Japan and the Caribbean.

Genetic abnormalities in leukaemia

Leukaemic cells often have a somatically acquired cytogenetic abnormality, which may be of prognostic, as well as diagnostic, significance.

These genetic alterations change the normal cell regulating process by interfering with the control of normal proliferation, blocking differentiation, maintaining an unlimited capacity for self-renewal and, lastly, promoting resistance to death signals, i.e. decreased apoptosis.

The first non-random chromosomal abnormality to be described was the Philadelphia (Ph) chromosome, which is associated with CML in 97% of cases. The Ph chromosome is also found in ALL, the incidence in the latter illness increasing with age. The translocation is shown schematically in Figure 9.5. The Ph chromosome is an abnormal chromosome 22, resulting from a reciprocal translocation between part of the long arm of chromosome 22 and chromosome 9. The resulting karyotype is described as t(9;22)(q34;q11). The molecular consequences of the translocation are that part of the Abelson proto-oncogene (*c-ABL*) normally present on chromosome 9 is translocated to chromosome 22, where it comes into juxtaposition with a region of chromosome 22 named the 'breakpoint cluster region' (BCR). The translocation creates a hybrid transcription unit consisting of the 5′ end of the *BCR* gene and the *c-ABL* proto-oncogene.

The new 'fusion' gene *BCR-ABL* is capable of being expressed as a chimeric messenger RNA which has been identified in cells from patients with CML. When translated, this produces a fusion protein that has tyrosine kinase

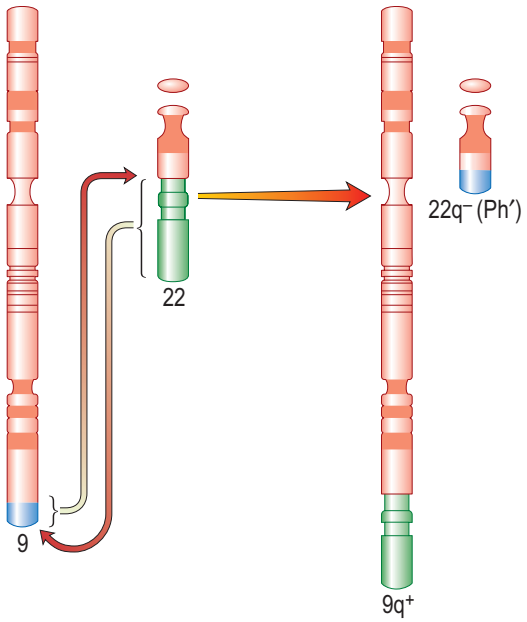


Fig. 9.5 **The Philadelphia chromosome (Ph).** The long arm (q) of chromosome 22 has been shortened by the reciprocal translocation with chromosome 9.

activity and enhanced phosphorylating activity compared with the normal protein, resulting in altered cell growth, stromal attachment and apoptosis. The breakpoint differs in CML and Ph-positive ALL, leading to the production of two different tyrosine kinase proteins with molecular weights of 210 kDa and 190 kDa respectively. It is unclear whether the presence of *BCR-ABL* is sufficient for the development of the disease. It has been shown that normal subjects can carry low levels of the *BCR-ABL* fusion gene in their blood without developing leukaemia.

Almost all patients with acute promyelocytic leukaemia (APML), a subtype of acute myelogenous leukaemia, have the t(15;17) reciprocal translocation, which occurs at the q25 band on chromosome 15 and the q22 band on chromosome 17. The breakpoint on chromosome 17 occurs in the gene encoding the retinoic acid receptor, fusing it with part of the PML gene (PML/RAR-alpha fusion gene). This is to some extent the explanation for the responsiveness of patients with APML to all-*trans*-retinoic acid (ATRA, see p. 471). Other genetic and cytogenetic abnormalities are often seen in leukaemic cells (Table 9.14).

Cell surface markers

These can be used to classify acute leukaemias. Immature myeloid cells have cell surface markers, e.g. CD13, CD14, CD33, CD34, which can be identified using monoclonal antibodies. The majority of patients with ALL (60%), for example, show the common ALL antigen (CALLA – CD10).

ACUTE LEUKAEMIAS

The acute leukaemias increase in incidence with advancing age. Acute myeloid (myeloblastic, myelogenous) leukaemia (AML) has a median age at presentation of 65 years and may arise 'de novo' or against a background of myelodysplasia, either of unknown aetiology or related to cytotoxic chemotherapy. Acute lymphoid (lymphoblastic) leukaemia (ALL) has a substantially lower median age at presentation and in addition is the commonest malignancy in childhood. The WHO classification is shown in Table 9.14.

Table 9.14 WHO classification of acute leukaemia

(a) Acute myeloid leukaemia
AML with recurrent genetic abnormalities
AML with t(8;21)(q22;q22), (AML1/ETO)
AML with abnormal bone marrow eosinophils and inv(16)(p13;q22) or t(16;16)(p13;q22), (CBFβ/MYH11)
Acute promyelocytic leukaemia with t(15;17)(q22;q12), PML/RAR-alpha and variants
AML with 11q23 (MLL) abnormalities
AML with multilineage dysplasia
Following MDS or MDS/MDP
Without antecedent MDS or MDS/MDP, but with dysplasia in at least 50% of cells in two or more myeloid lineages
AML and myelodysplastic syndromes, therapy related
Alkylating agent/radiation-related type
Topoisomerase II inhibitor-related type
Other
AML, not otherwise categorized*
AML, minimally differentiated
AML without maturation
AML with maturation
Acute myelomonocytic leukaemia
Acute monoblastic/acute monocytic leukaemia
Acute erythroid leukaemia (erythroid/myeloid and pure erythroleukaemia variants)
Acute megakaryoblastic leukaemia
Acute basophilic leukaemia
Acute panmyelosis with myelofibrosis
Myeloid sarcoma
(b) Acute lymphoid leukaemia
Precursor B cell acute lymphoblastic leukaemia
t(9;22)(q34;q11); BCR/ABL fusion gene
t(4;11)(q21;q23); MLL-AF4 fusion gene
t(1;19)(q23;p13.3); E2A/PBX1 fusion gene
t(12;21)(p13;q22); TEL/AML1
Precursor T cell acute lymphoblastic leukaemia
Burkitt-cell leukaemia

MDS, myelodysplastic syndromes; MPD, myeloproliferative diseases.
*The entities included in this group are defined almost identically to the corresponding entity in the French–American–British (FAB) classification.
Modified from Jaffe ES, Harris NL, Stein H, Vardiman JW (eds). *World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues.* Lyon: IARC Press, 2001, with permission of the World Health Organization.

Clinical features

The majority of patients with acute leukaemia, regardless of subtype, present with symptoms reflecting inadequate haematopoiesis secondary to leukaemic cells infiltration of the bone marrow:

- anaemia – shortness of breath on effort, excessive tiredness, weakness
- leucopenia – recurrent infections
- thrombocytopenia – bleeding and bruising (particularly acute promyelocytic leukaemia)
- marrow infiltration – bone pain.

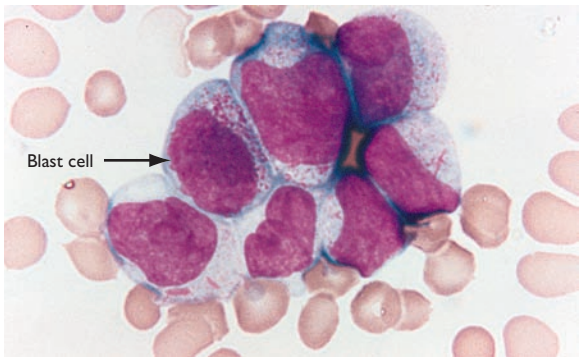
Examination may be unremarkable, but features include:

- pallor
- fever (due to infection, not the disease itself)
- petechiae, purpura, bruises, fundal haemorrhage (particularly acute promyelocytic leukaemia)
- lymphadenopathy, hepatosplenomegaly (more notable in lymphoblastic leukaemia)
- violaceous skin lesions (acute myelomonocytic leukaemia)
- testicular enlargement (acute lymphoblastic leukaemia)
- cranial nerve palsies occasionally found (acute lymphoblastic leukaemia).

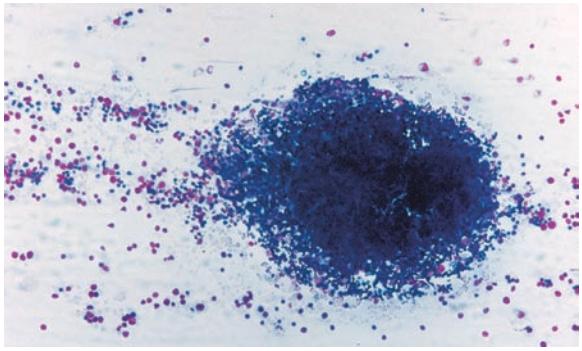
Investigations

Confirmation of diagnosis (Fig. 9.6)

- **Blood count**. Hb low, WBC raised usually (sometimes low), platelets low.
- **Blood film**. Blast cells almost invariably seen, (Fig. 9.6a), lineage identified morphologically, confirmed with immunophenotyping.
- **Bone marrow aspirate**. Increased cellularity, reduced erythropoiesis, reduced megakaryocytes, sometimes trilineage dysplasia. Replacement by blast cells > 20%



(a)



(b)

Fig. 9.6 **Acute leukaemia. (a) Peripheral blood film showing characteristic blast cells.** The arrow points to the abnormal blast cell. **(b) Bone marrow aspirate showing particle with increased cellularity.** Courtesy of Dr Manzoor Mangi.

(often approaching 100%) (Fig. 9.6b). Lineage confirmation by immunophenotyping (FISH), cytogenetic and molecular genetics, e.g. AML – CD33 or CD13, and ALL – CD10 and CD19.

- **Chest X-ray**. Mediastinal widening often present in T lymphoblastic leukaemia.
- **Cerebrospinal fluid examination**. Performed when blasts have been cleared from the peripheral blood in all patients with ALL, as the risk of CNS involvement is high. It is less critical in AML, except in cases where there is a monoblast/monocytic component.

For planning therapy

- Biochemistry, serum urate, renal and liver biochemistry.
- Cardiac function; ECG and direct tests of left ventricular function, e.g. echocardiogram/cardiac magnetic resonance (see p. 706).

Principles of management

Untreated acute leukaemia is invariably fatal, most often within months, though with judicious palliative care it may be extended to perhaps a year. Treatment with curative intent may be successful, or may fail, either because the leukaemia cannot be eradicated or because the patient cannot sustain the therapy, death occurring as early as if treatment had not been initiated.

At initial presentation, acute leukaemias range from being probably curable (most favourable risk – childhood acute lymphoblastic leukaemia) through possibly curable (de novo low-risk AML) to probably incurable (AML with adverse cytogenetic features in the elderly, secondary AML, recurrent acute leukaemia). Since curative treatment even for 'low-risk' acute leukaemia carries considerable morbidity and potential mortality and that for 'high-risk' acute leukaemia even more, it is essential that the 'risk/benefit' ratio is clearly understood by physician and patient alike.

In AML, patients with t(15;17) t(8;21) or inv(16) (or its variant t16;16), i.e. low risk, do not benefit from allogeneic stem cell transplantation during their first complete remission because the risks outweigh benefits. Patients with adverse factors (high risk) – (5/del5q), -7, abnormal 3q26 or a complex karyotype – should have stem cell transplantation because they respond poorly to conventional chemotherapy. Other poor prognostic factors (high risk) include developing the disease over 60 years of age, leukaemia following myelodysplastic syndrome (MDS), relapsed disease, secondary leukaemia and extramedullary disease.

Palliative therapy

Every attempt should be made to ensure that the patients are at home as much as possible, whilst making available the full range of supportive care. Palliation may well include both chemotherapy and irradiation in addition to blood product support. 'Moral' support is invaluable.

Curative therapy

The decision to treat with curative intent, particularly if successful, implies severe disruption of normality for the patient and family for at least 6 months and often up to a year, and, regardless of success, life is never quite the same again. In the short term, it may demand transfer to another hospital, as acute leukaemia should only be treated in units seeing at least 10 such cases per year. It is highly likely to involve admission to hospital for up to a month in the first instance, with further, partly predictable, subsequent admissions of several days' to weeks' duration, requiring discussions and decisions about work or education.

The decision to treat with curative intent implies that cure is possible, and that the chance of cure justifies the risks of the therapy. It does not imply that cure is guaranteed or even expected. The failure rate may be high, and the patient must know that he or she will be told if cure becomes an unrealistic goal. Treating with curative intent may well involve rapid decisions about resuscitation with transfer to the intensive care unit, the possibility of which is discussed in advance.

Active therapy

Supportive care

This forms the basis of treatment whether for cure or palliation:

- Avoidance of symptoms of anaemia (haemoglobin > 10 g/dL) – repeated transfusion of packed red cells (sometimes irradiation of cells is required).
- Prevention or control of bleeding (platelet count < 10×10^9/L in the uninfected, and 20×10^9/L in the infected patient).
- Treatment of infection (see Box 9.3):
  (a) *Prophylactically*. Education of patients, relatives and staff about hand washing and isolation facilities. The use of selected antibiotics and antifungals.
  (b) *Therapeutically*. Management of fever with protocol/algorithm of antibiotic and antifungal combinations.

- Control of hyperuricaemia with hydration, prophylactic allopurinol and very occasionally rasburicase (see p. 458).

Specific treatment

The initial requirement of therapy is to return the peripheral blood and bone marrow to normal (complete remission; CR). This '*induction chemotherapy*' is tailored to the particular leukaemia and the individual patient's risk factors. Since this treatment is not leukaemia specific but also impairs normal bone marrow function, it leads to a major risk of life-threatening infection, which increases the risk of early death in the short term. Since infection is the major problem, it is necessary to conduct the early therapy in hospital, the patient sleeping in a single room with en-suite lavatory and washing facilities ('some' isolation). An indwelling venous catheter inserted at the time of diagnosis secures venous access for the administration of cytotoxics, blood product support and antimicrobial therapy.

Successful *remission induction* is always followed by further treatment (*consolidation*). The details of this are determined by the type of leukaemia, the patient's risk factors and the patient's tolerance of treatment. Recurrence is almost invariable if 'consolidation' therapy is not given. This reflects the lack of sensitivity of the definition of 'complete remission', which has been solely morphological. Cytogenetics and molecular genetic techniques can however identify residual leukaemic cells not detected morphologically, and they are highly predictive of recurrence. Recommendations have recently been made to modify the definition of remission to reflect this. Failure to achieve morphological CR with two cycles of therapy carries almost as bad a prognosis as the untreated leukaemia. If CR can be achieved, e.g. by new experimental approaches, cure may still be possible with stem cell transplantation (see p. 462).

Acute myeloid leukaemia (AML; excluding APML, p. 471)

Treatment with *curative intent* is undertaken in the majority of adults below the age of 60 years, provided there is no significant co-morbidity. Risk of failure is based on the cytogenetic pattern (Fig. 9.7). Those at 'low risk' are treated with moderately intensive combination chemotherapy. This always includes an anthracycline such as daunorubicin and cytosine arabinoside (cytarabine) and consolidation with a minimum of four cycles of treatment given at 3- to 4-week intervals. Those at 'high risk' may only be treated with curative intent if an HLA-identified sibling is available for stem cell transplantation.

Those at 'intermediate risk' are a heterogeneous group. When possible they should be given consolidating chemotherapy to induce remission followed by sibling-matched allogeneic transplantation, despite its attendant risks.

For those with a favourable risk profile, allogeneic transplantation is not recommended to consolidate first remission. In this group the potential morbidity from the procedure outweighs the potential risk of relapse.

The initial treatment of the *older patient* is much more contentious. Intuitively, biological age should determine the management of the individual patient. Unfortunately 'high-risk AML' is commoner with increasing age, but is only curable with allogeneic transplantation, and the toxicity of this treatment increases dramatically with age.

Complete remission will be achieved in about three-quarters of patients under the age of 60, failure being due to either resistant leukaemia or death due to infection or (rarely) bleeding. Approximately 50% of those entering complete remission will be cured. (i.e. approximately 30% overall) (Fig. 9.8). The management of recurrence is undertaken on an individual basis, since the overall prognosis is very poor despite

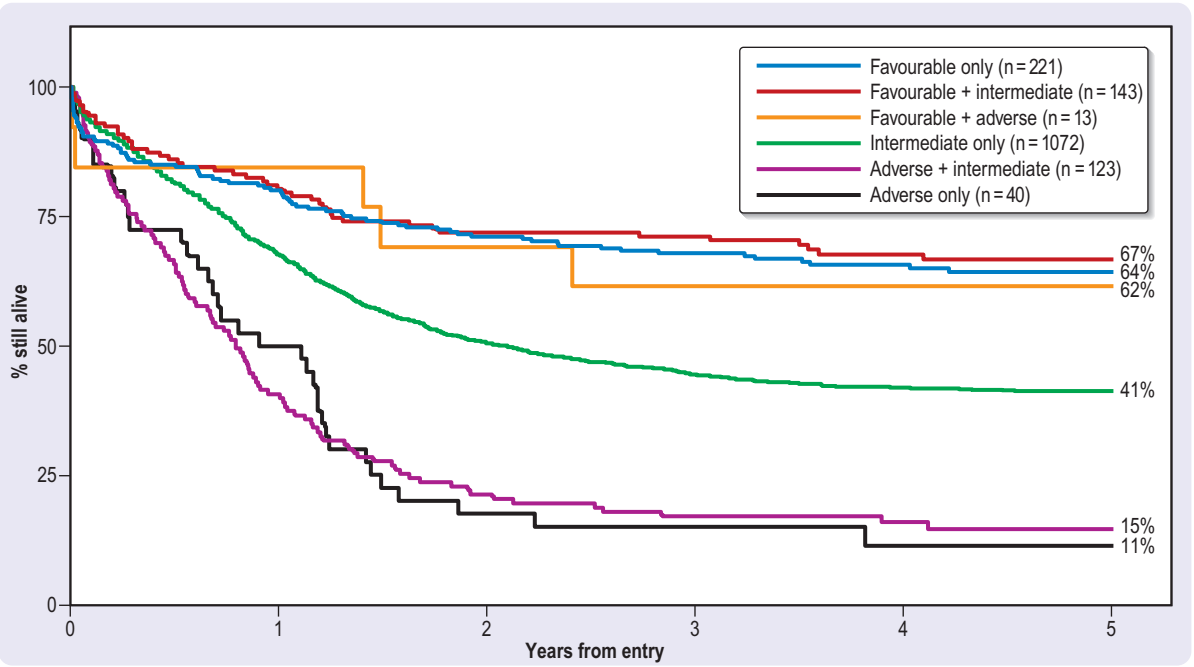


Fig. 9.7 **Influence of cytogenetic risk groups on overall survival in AML.** This research was originally published in *Blood*. Grimwade, Walker H, Oliver F et al. The importance of diagnostic cytogenetics on outcome in AML: analysis of 1612 patients entered into the MRC AML 10 Trial. *Blood* 1998; **92**: 2322–2333. © American Society of Hematology.

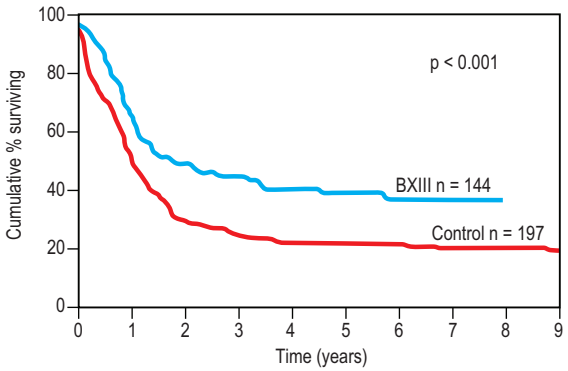


Fig. 9.8 **Acute myeloid leukaemia: overall survival with or without myeloablative therapy.** BXIII, treatment group. From Rohatiner AZS, Gregory WM, Bassan R et al. High-dose treatment with autologous bone marrow support as consolidation of first remission in younger patients with acute myelogenous leukaemia. *Annals of Oncology* 2000; **11:** 1007–1015 by permission of Oxford University Press.

the fact that second remissions may be achieved. Long survival following recurrence is rarely achieved without allogeneic transplantation. Experimental therapy should be considered.

Acute promyelocytic leukaemia (APML)

This is an uncommon variant of AML characterized by the translocation t(15;17) and with particular morphological features (see p. 467). There is an almost invariable coagulopathy, which was a major cause of death. The empirical discovery that all-*trans*-retinoic acid (ATRA) causes differentiation of promyelocytes and rapid reversal of the bleeding tendency was a major breakthrough. APML is treated with ATRA combined with chemotherapy and, following successful remission induction, with maintenance ATRA. Allogeneic transplantation may be necessary either if the leukaemia is not eliminated at the molecular level, or following a second remission after recurrence. Arsenic trioxide, which induces apoptosis via activation of the caspase cascade (see p. 28), is used with resistant or relapsed disease.

Complete remission and molecular remission occur in at least 80% of younger adults with APML (it is uncommon in the elderly). At least 60% will expect to be cured. In contrast to the other subtypes of AML, the prognosis after recurrence is quite favourable with prolonged second remissions being possible with further blocks of ATRA and chemotherapy even if allogeneic transplant is not able to be performed, although it is the treatment of choice.

Acute lymphoblastic leukaemia (ALL)

The overall strategy for the treatment of ALL differs in detail from that for AML. Remission induction is undertaken with combination chemotherapy including vincristine, a glucocorticoid, an anthracycline and asparaginase (crisantaspase). Daunorubicin is perhaps the most frequently used anthracycline, and dexamethasone is replacing prednisolone as the glucocorticoid of choice. Once remission is achieved, the details of consolidation will be determined by the anticipated risk of failure. Intensive consolidation of remission with variable numbers of chemotherapy cycles comprising cytotoxics with different mechanisms of action has been standard practice, including the administration of high-dose methotrexate. Consolidation high-dose chemotherapy and autologous stem cell rescue is probably equivalent to standard consolidation but is rarely used. In those patients with high-risk features (see below), allogeneic transplantation is recommended upon achieving first complete remission.

The other major difference between therapy for ALL and AML is the need for central nervous system directed therapy. Prophylaxis should be given with intrathecal chemotherapy under platelet cover if necessary, as soon as blasts are cleared from the blood. Depending upon risk this may be continued for up to 2 years, and complemented by high doses of systemic cytosine arabinoside (cytarabine) or methotrexate. Cranial irradiation was previously given to all patients to reduce the risk of relapse within the central nervous system; some risk adapted strategies now reserve this only for those patients at very high risk (Fig. 9.9).

After intensive induction and consolidation, maintenance therapy for 2 years is required to reduce the risk of disease recurrence in certain subtypes. This typically comprises 2 years of treatment with methotrexate and mercaptopurine, although more intensive regimens are used by some groups. During this period of time chemotherapy dosing/scheduling is adjusted to maintain a target white cell count of 3 $\times 10^9$/L.

Prognostic factors

A number of clinical and laboratory features are determinants of treatment response and survival in ALL. Increasingly therapeutic strategies, based upon prognostic risk, are being used in the management of the disease.

- *Age*. Overall survival decreases with increasing age in adult ALL.
- *White cell count*. High white cell counts at the time of diagnosis (>30 000 $\times 10^9$/L); associated with high risk of CNS relapse.
- *Immunophenotype*. The immunophenotype is characterized at presentation and can be used to divide ALL patients into risk groups. B cell ALL represents 75% of ALL, with the remainder of T cell lineage. In the former, Pro-B ALL is considered a poor prognostic group compared to the more common Pre B-ALL form. T-ALLs are associated with the presence of a mediastinal mass although they may have superior outcomes to those of B-cell lineage.
- *Cytogenetic aberrations*. Chromosomal abnormalities are detected in 60% of adult ALL patients. The presence of t(9;22)or t(4;11) is associated with a poor prognosis, the former being particularly unfavourable. It is unclear yet what the role of imatinib will be in this setting.
- *Time to response*. Early clearance of blasts is favourable, with failure to achieve a CR within 3–4 weeks being adverse
- *Minimal residual disease*. Highly sensitive techniques have been employed to detect low level leukaemic populations in the morphologically normal bone marrow. These include flow cytometry, PCR quantitative assessment of immunoglobulin gene rearrangement and of chromosomal translocation. This identifies patients with high levels of minimal residual disease who may be at high risk of relapse and thus suitable for therapeutic intensification.

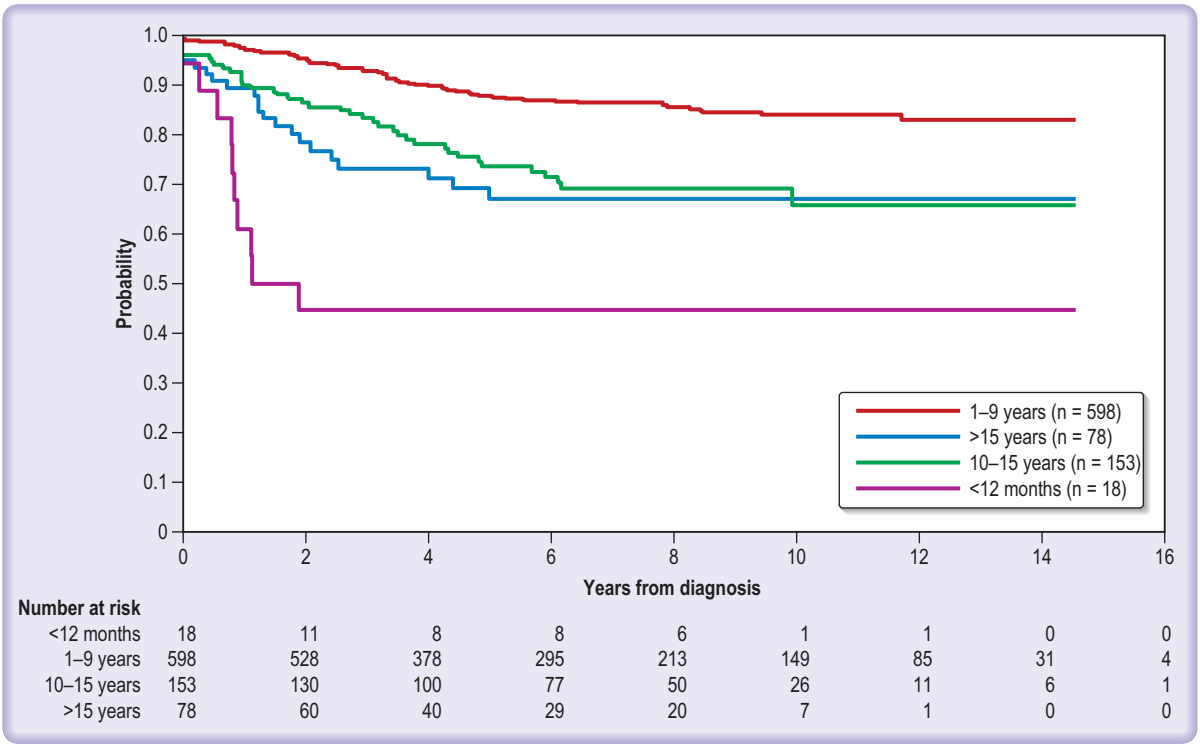


Fig. 9.9 **Overall survival in 2628 children with newly diagnosed ALL** participating in consecutive studies conducted at St Jude Children's Research Hospital from 1962 to 2005. From Pui C-M, Robinson LL, Lock T. Acute lymphoblastic leukaemia. *Lancet* 2008; **371:** 1030–1043, figure 3 with permission from Elsevier.

Prognosis (Fig. 9.9)

The prognosis of ALL in childhood is now excellent: complete remission is achieved in almost all, with up to 80% being alive without recurrence at 5 years. Failure occurs most frequently in those with high blast count and t(9;22) translocation.

The situation is far less satisfactory for adults, the prognosis getting worse with advancing years. Co-morbidity and t(9;22) translocation increases in frequency with age. Overall the complete remission rate is 70–80%, failure being due partly to resistant leukaemia and partly to failure of supportive care. Failure to achieve complete remission with first-line therapy carries a very poor prognosis. If CR can be achieved with new therapies, it should be consolidated with sibling or possibly even unrelated donor transplantation despite the high risk of graft-versus-host disease. Between 30% and 40% of patients continue in durable first remissions, resulting in approximately 25–30% overall patient cure.

As with AML, most recurrences occur within the first 3 years and the outcome is extremely poor. Second remissions, though usually achieved, are rarely durable except following allogeneic transplantation. Isolated extramedullary recurrences, however, may be cured.

CHRONIC LEUKAEMIAS

Chronic myeloid leukaemia (CML)

Chronic myeloid leukaemia (CML), which accounts for about 14% of all leukaemias, is almost exclusively a disease of adults with the peak of presentation being between 40 and 60 years and is characterized by the presence of the Philadelphia chromosome (Fig. 9.10). Unlike the acute leukaemias which are either rapidly reversed or rapidly fatal,

Fig. 9.10 **Philadelphia chromosome.** This is formed by a reciprocal translocation of part of the long arm (q) of chromosome 22 to chromosome 9. It is seen in 90–95% of patients with chronic myeloid leukaemia. The karyotype is expressed as 46XX, (9;22)(q34;q11).

CML has a more slowly progressive course which if not initially cured will be followed eventually by blast crisis (90% myeloid, 20% lymphoid) or myelofibrosis and death after 3–4 years.

Clinical features

CML usually presents in the chronic phase and some patients have no symptoms. Symptoms include:

- symptomatic anaemia (e.g. shortness of breath)
- abdominal discomfort due to splenomegaly

- weight loss
- fever, sweats, in the absence of infection
- headache (occasionally) due to hyperleucocytosis
- bruising, bleeding (uncommon), priapism.

Signs include:

- pallor
- splenomegaly, often massive
- lymphadenopathy (uncommon, when found suggests blast crisis)
- retinal haemorrhage due to leucostasis.

Investigations

- **Blood count**. Hb low (normochromic and normocytic) or normal, WBC raised (usually $>100 \times 10^9$/L), platelets low, normal or raised.
- **Blood film**. Neutrophilia with the whole spectrum of myeloid precursors including occasional blasts. Elevated basophils and eosinophils.
- **Bone marrow aspirate**. Increased cellularity, increased myeloid precursors. Cytogenetics reveals t(9;22) translocation (the Philadelphia chromosome) (Fig. 9.10).
- **Fluorescein-in-situ hybridization (FISH)** or reverse transcriptase polymerase chain reaction (RT-PCR) are used to demonstrate the cytogenetic/molecular abnormality. These are also used to quantitatively monitor response to therapy.
- **Leukocyte alkaline phosphatase** is usually reduced.

Management

Imatinib, a tyrosine kinase inhibitor that specifically blocks the enzymatic action of the BCR-ABL fusion protein, is first-line treatment for the chronic phase. It has replaced alpha-interferon. Imatinib produces a complete haematological response in over 95% of patients, and 70–80% of these have no detectable BCR-ABL transcripts in the blood. Event-free, and overall, survival appear to be better than for other treatments. Imatinib can be continued indefinitely.

In the acute phase (blast transformation) most patients have only a short-lived response to imatinib, and other chemotherapy as for acute leukaemia is used in the hope of achieving a second chronic phase.

Side-effects of imatinib, which usually are well tolerated, include nausea, headaches, rashes and cytopenia. Resistance to imatinib as a single agent may develop as a result of secondary mutations beyond the t(9;22). The use of second generation tyrosine kinase inhibitors, dasatinib and nilotinib, may restore haematological or molecular remissions in those patients in the chronic phase that have primary or acquired resistance to imatinib.

Stem cell transplantation (SCT)

Allogeneic haemopoietic stem cell transplantation can cure approximately 70% of chronic phase CML patients. SCT was choice therapy for young patients with an HLA matched donor, but this has changed in the light of the success of imatinib therapy. It is now only used in those with an inadequate response to imatinib or those that have disease progression on therapy. There are many potential complications with an allogeneic transplantation approach; death may occur as a result of graft-versus-host disease (GVHD) or opportunistic infection, and the decision to proceed is based upon a balance of risk.

Factors making complications more likely include:

- increasing age
- SCT in acute phase
- prolonged interval from diagnosis to transplantation
- degree of histocompatibility between donor and recipient.

Graft-versus-leukaemia effect plays a role in the increased survival following SCT so that reduced-intensity transplantation is being more frequently used.

Chronic lymphocytic leukaemia (CLL)

This is the commonest leukaemia, occurring predominantly in later life and increasing in frequency with advancing years (median age of presentation between 65 and 67 years). It results from the clonal expansion of small lymphocytes and is almost invariably (95%) B cell in origin. The majority of patients are asymptomatic, identified as a chance finding on a blood count performed for another indication. Other patients, however, present with the features of marrow failure or immunosuppression. The median survival is about 10 years, and prognosis correlates with various clinical features at presentation (Table 9.15). These clinical features simply represent differences in the biology of the disease, and a number of cytogenetic and molecular abnormalities are now recognized as being of prognostic significance (see below). A pre-malignant condition, *monoclonal B cell lymphocytosis* (MBL), occurs where there are less than the 5×10^9/L B cells required for a diagnosis of CLL. Some of these have CLL phenotype and may progress to CLL.

Clinical features

The majority of patients are asymptomatic at presentation. Common symptoms are:

- recurrent infection because of (functional) leucopenia and immune failure (reduced immunoglobulins)
- anaemia due to haemolysis or marrow infiltration
- painless lymphadenopathy
- left upper quadrant discomfort (from splenomegaly).

The commonest findings on examination are:

- anaemia
- fever (due to infection)
- generalized lymphadenopathy (may involve single area)
- hepatosplenomegaly, sometimes massive.

However, none of these may be present.

Investigations

- **Blood count**. Hb normal or low; WBC raised, and may be very high; with lymphocytosis (criteria for diagnosis $> 5 \times 10^9$/L), platelets normal or low.
- **Blood film**. Small or medium sized lymphocytes. May see smudge cells in vitro.
- **Bone marrow**. Reflects peripheral blood, often very heavily infiltrated with lymphocytes.
- **Immunophenotyping** shows mainly $CD19^+$, $CD5^+$, $CD23^+$ with a weak expression of CD20 and CD79b and surface immunoglobulin (kappa and lambda light chains).
- **Cytogenetics/FISH analysis** are not essential for diagnosis but may help in the assessment of prognosis.

Table 9.15 The Rai and Binet staging systems for chronic lymphocytic leukaemia*

System and stage	Risk	Manifestations	Percent of patients	Median survival (years)	Recommended treatment
Rai staging system					
0	Low	Lymphocytosis	31	>10	Watch and wait
I	Intermediate	Lymphadenopathy	35	9	Treat only with progression†
II	Intermediate	Splenomegaly, lymphadenopathy, or both	26	7	Treat only with progression†
III	High	Anaemia, organomegaly	6	5	Treatment indicated in most cases
IV	High	One or more of the following: anaemia, thrombocytopenia and organomegaly	2	5	Treatment indicated in most cases
Binet staging system					
A	Low	Lymphocytosis, <3 lymphoid areas enlarged‡	63§	>10	Watch and wait
B	Intermediate	≥3 Lymphoid areas enlarged‡	30	7	Treatment indicated in most cases
C	High	Anaemia, thrombocytopenia or both	7	5	Treatment indicated in most cases

*Lymphocytosis is present in all stages of the disease.
†Progression is defined by weight loss, fatigue, fever, massive organomegaly and a rapidly increasing lymphocyte count.
‡Enlarged lymphoid areas may include the cervical, axillary and inguinal lymph nodes; the spleen or liver may be enlarged.
§Stage A includes all patients with Rai stage 0 disease, two-thirds of patients with Rai stage I disease and one-third of those with Rai stage II.
From Dighiero G, Binet JL. When and how to treat chronic lymphocytic leukaemia. *New England Journal of Medicine* 2000; **343**: 1800. Copyright © Massachusetts Medical Society. All rights reserved.

- **Coombs' test**. May be positive if there is haemolysis.
- **Immunoglobulins**. Low or normal.

Prognostic factors

The clinical course of CLL is variable. Several serum markers, e.g. β_2 microglobulin, soluble CD23 and thymidine kinase, have been shown to predict progression and survival. Variations in predictor cut-off levels have limited their widespread application. Cytogenetic abnormalities are detected in >90% of cases. Patients with an isolated deletion at 13q have an excellent prognosis, in contrast to those with either 11q deletions or 17p deletions (the sites of the tumour suppressor genes *ATM* and *TP53* respectively) who tend to have a rapidly evolving clinical course. Trisomy of 12 is frequently observed and similarly conveys risk of progression. In those tumours that demonstrate a high level of mutation within the variable region of the rearranged immunoglobulin heavy chain (*IgVH*) the clinical course is more indolent than those where the *IgVH* sequence more closely resembles that of the germ line. Such assessments are technically demanding; expression of ZAP70, a 70 kDa tyrosine kinase protein, correlates well with mutational status. Patients with <20% expression of ZAP70 have median 10-year survival of ≥50%; in >20% expression the median survival was <5 years. High expression of CD38 on leukaemic cells may also indicate adverse prognosis.

Management

In CLL, the major consideration is when to treat, indeed 30% of patients will never require intervention. Treatment depends on the 'stage' (Table 9.15) of the disease and the prognostic biomarkers. Choice of therapy will depend upon patient-related factors such as age and co-morbidity, adverse prognostic features and anticipated response and toxicities to therapy. Intervention, when indicated, usually causes improvement in symptoms and in the blood count. The effect on survival is unclear. More aggressive treatments, particularly combinations of cytotoxic chemotherapy with antibody therapy, result in better quality remission of longer duration. These improvements may translate into a survival advantage, to accompany the improvement in quality of life afforded by good supportive care.

Early-stage disease is usually managed expectantly, advanced-stage disease is always treated immediately and the approach to the intermediate stage is variable. The absolute indications for treatments are:

- marrow failure manifest by worsening anaemia and/or thrombocytopenia
- recurrent infection
- massive or progressive splenomegaly or lymphadenopathy
- progressive disease manifest by doubling of the lymphocyte count in 6 months
- systemic symptoms (fever, night sweats or weight loss)
- presence of haemolysis or other immune mediated cytopenias.

General/supportive treatment

Anaemia due to haemolysis is treated with steroids. If it is refractory or recurrent, or if splenic discomfort is a problem, a splenectomy is performed. Anaemia and thrombocytopenia due to marrow infiltration is treated with chemotherapy and, when necessary, transfusion. Erythropoietin (see p. 390) may avoid the need for transfusions, particularly in patients receiving chemotherapy.

Infection is treated with antibiotics, with prophylactic therapy being given during periods of chemotherapy. Immunoglobulin replacement may be helpful.

Allopurinol is given to prevent hyperuricaemia.

Specific treatment

- Chlorambucil, given in modest doses, usually reduces the blood count and decreases lymphadenopathy and splenomegaly, and successfully palliates the disease. The bone marrow rarely returns to normal. Treatment is usually limited to a few months' duration and then withheld until progression. In asymptomatic patients without an indication for therapy, early use of chlorambucil does not provide a survival advantage over expectant management.
- *Purine analogues, fludarabine* alone or in combination with *cyclophosphamide* or *mitoxantrone* (with or without steroids), have had a much greater impact on the bone marrow and can induce complete or molecular complete remission although they are not helpful in 17p-deletion or p53 mutations.
- Combination therapy with *rituximab* (relatively ineffective alone) shows a dramatic improvement in the response rate and has become *standard choice first-line therapy*. Alemtuzumab, a humanized monoclonal antibody targeting CD52 which is highly expressed on B-CLL, may be used in those patients that progress after fludarabine.
- Allogeneic stem cell transplantation with non-myeloablative conditioning regimens is undergoing investigation, particularly for the younger patient.

Lymphomatous transformation

CLL may undergo lymphomatous (Richter's) transformation in 5–10% of cases, most typically to diffuse large B-cell lymphoma, although Hodgkin's like transformation is recognized. In the main, response to cytotoxic chemotherapy is unsatisfactory and survival short.

Hairy cell leukaemia (HCL)

HCL is a clonal proliferation of abnormal B (or very rarely T) cells which, as in CLL, accumulate in the bone marrow and spleen. It is a rare disease, median age at presentation is 52 years old and the male to female ratio is 4:1. The bizarre name relates to the appearance of the cells on a blood film and in the bone marrow – they have an irregular outline owing to the presence of filament-like cytoplasmic projections. They show a strong acid phosphatase reaction that is resistant to tartaric acid. The cells express many cellular differentiation markers including CD19, 20 and 103 but not CD21 or 5.

Clinical features

Clinical features include anaemia, fever and weight loss. Splenomegaly occurs in 80%, lymphadenopathy is uncommon. Anaemia, neutropenia, thrombocytopenia and low monocyte counts are found.

Treatment

The purine analogues 2-chloroadenosine acetate (2-CDA) (cladribine) and pentostatin have specific activity in this condition; complete remission is achieved in 90% with just one cycle of treatment. The remissions sometimes last for several years and patients can be retreated. Rituximab is used in cases who do not respond to the above drugs.

Prolymphocytic leukaemia

Prolymphocytic leukaemia is another rare disorder, often mistaken for CLL. It may be of B or of T cell lineage. It is characterized by bone marrow failure (anaemia, neutropenia and thrombocytopenia) and – as in HCL – splenomegaly. Treatment generally comprises chlorambucil as for CLL, although splenectomy may be indicated and fludarabine can be useful.

FURTHER READING

British Committee for Standards in Haematology, Milligan DW, Grimwade D, Cullis JO et al. Guidelines on the management of acute myeloid leukaemia in adults. *British Journal of Haematology* 2006; **135**: 450–474.

Copelan EA. Hematopoietic stem-cell transplantation. *New England Journal of Medicine* 2006; **354**: 1813–1826.

Deininger MW. Management of early stage disease (chronic myeloid leukaemia). *Hematology (American Society of Hematology Education Program)* 2005; 174–182.

Dighirro G, Hambun TJ. Chronic lymphatic leukaemia. *Lancet* 2008; **371**: 1017–1029.

Etsey E, Döhner H. Acute myeloid leukaemia. *Lancet* 2006; **368**: 1894–1907.

Jaffe ES, Harris NL, Stein H et al *(eds). World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues.* Lyon: IARC Press, 2001.

Lowenburgh B. Diagnosis and prognosis in acute myeloid leukaemia. *New England Journal of Medicine* 2008; **358**: 1960–1967.

Pui Ching- Hon, Evans WE, Relling MV. Acute lymphoblastic leukaemia. *Lancet* 2008; **371**: 1030–1043.

THE LYMPHOMAS

The lymphomas are commoner than the leukaemias and are increasing in incidence for reasons which are unclear. They arise as the result of abnormal proliferation of the lymphoid system, and hence occur at any site where lymphoid tissue is found. Most commonly they are manifest by the development of lymphadenopathy at single or multiple sites, although primary extranodal presentations account for up to 20% of non-Hodgkin's lymphoma. The prognosis is determined by the specific subtype of lymphoma and the anatomical extent of disease and its bulk, the clinical course ranging from months to years.

The guiding principles of management are broadly the same as for the leukaemias. The precise diagnosis is established, appropriate further investigation is conducted to allow a management plan to be formulated, both for the short and long term, and the situation is clearly explained to the patient.

Lymphomas are currently classified on the basis of histological appearance into:

- Hodgkin's lymphoma
- non-Hodgkin's lymphoma.

The distinction between lymphoid leukaemia and lymphoma is not always clear.

HODGKIN'S LYMPHOMA (HL)

This is a rare disease involving primarily the lymph nodes. The incidence in the UK and North America is approximately 3/100 000, although this is less in populations derived from Eastern Asia. It occurs slightly more frequently in males than females with a ratio of 1.3:1. Over 90% will occur in adults between 16 and 65 years with the peak incidence in the third decade. The incidence is stable.

Aetiology

There is epidemiological evidence linking previous infective mononucleosis with HL and up to 40% of patients with HL have increased EBV antibody titres at the time of diagnosis and several years prior to the clinical development of HL. EBV DNA has been demonstrated in tissue from patients with HL. These data suggest a role for EBV in pathogenesis. Other viruses have not been detected. Other environmental and occupational exposures to pathogens have been postulated.

Pathology

The hallmark of HL is the presence of the clonal malignant Reed–Sternberg cell in a lymph node biopsy or (rarely) extranodal tissue (Fig. 9.11). There is a paucity of such malignant cells, surrounded by abundant bystander cells. The Reed–Sternberg cells are usually derived from germinal centre B cells or, rarely, peripheral T cells. CD30 and CD15 are almost always expressed in the majority of cases of classical HL. It is clear that several pathways may lead to HL but the central issue must be that lymphocytes of the B-cell lineage not expressing immunoglobulins somehow escape apoptosis.

The WHO classification of HL is shown in Table 9.16. Classical HL (representing 95% of cases) can be divided into:

- *Nodular sclerosing HL* (70% of cases) demonstrates a nodular growth pattern with many fibrotic bands present. This type is typically seen in young adults, without sex predominance. It involves particularly cervical and supraclavicular lymph nodes and the anterior mediastinum.
- *Lymphocyte-rich HL* appears in 5% and is characterized by an infiltrate of many small lymphocytes and Reed–Sternberg cells. It often occurs in peripheral lymph nodes. It is often an indolent disease which presents at a higher median age.
- *Mixed cellularity HL*. Approximately 25% of cases have mixed cellularity with lymphocytes, eosinophils, neutrophils and histiocytes. Reed–Sternberg cells are present but no fibrotic bands. It is more common in men and is associated with B symptoms (see below).
- *Lymphocyte-depleted HL* is rare and there is lack of cellular infiltrate with numerous Reed–Sternberg cells. It typically presents with advanced stage and B symptoms. It is seen in HL associated with HIV.

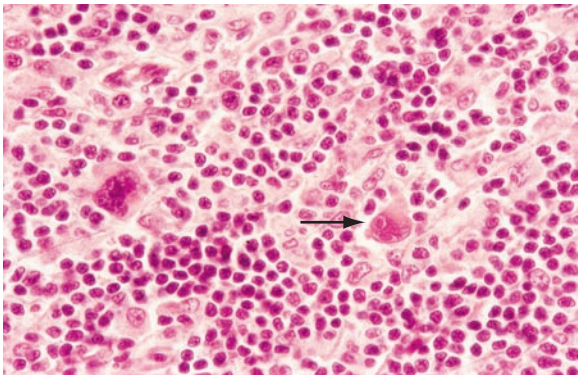


Fig. 9.11 **Histological appearance of Hodgkin's lymphoma.** There is a background rich in benign small lymphocytes and histiocytes together with scattered mononuclear Hodgkin's cells and a classical malignant binucleate Reed–Sternberg cell (arrow) to the right of centre. Courtesy of Dr AJ Norton.

Table 9.16 Hodgkin's lymphoma – pathological classification

Nodular lymphocyte-predominant Hodgkin's lymphoma
Classical Hodgkin's lymphoma:
Nodular sclerosis HL
Lymphocyte-rich HL
Mixed cellularity HL
Lymphocyte-depleted HL

From Harris NL et al. *Journal of Clinical Oncology* 1999; **17**: 3835–3849.

Table 9.17 Cervical lymph node enlargement – differential diagnosis

Infections

Acute
Pyogenic infections
Infective mononucleosis
Toxoplasmosis
Cytomegalovirus infection
Infected eczema
Cat scratch fever
Acute childhood exanthema

Chronic
Tuberculosis
Syphilis
Sarcoidosis
HIV infection

Autoimmune rheumatic disease
Rheumatoid arthritis

Drug reactions
Phenytoin

Primary lymph node malignancies
Hodgkin's lymphoma
Non-Hodgkin's lymphoma
Chronic lymphocytic leukaemia
Acute lymphoblastic leukaemia

Secondary malignancies
Nasopharyngeal
Thyroid
Laryngeal
Lung
Breast
Stomach

Miscellaneous
Kawasaki's syndrome
Kikuchi's disease (histiocytic necrotizing lymphadenitis)

In addition to classical HL, nodular lymphocyte-predominant HL (5% of cases) contains malignant L and H cells (lymphocytic and/or histiocytic Reed–Sternberg cell variants, also called 'popcorn' cells) which are positive for CD20, CD45, BCL6, CD79a without expressing CD15 or CD30.

Clinical features

- Lymph node enlargement, most often of the cervical nodes (other causes are shown in Table 9.17); these are usually painless and with a rubbery consistency. The pattern of spread is usually contiguous.
- Enlargement of the spleen/liver.
- Systemic 'B' symptoms: fever, (25%) drenching night sweats, weight loss of >10% bodyweight (see Table 9.18).

- Other constitutional symptoms, such as pruritus, fatigue, anorexia and, occasionally, alcohol-induced pain at the site of enlarged lymph nodes.
- Symptoms due to involvement of other organs (e.g. mediastinum – cough and breathlessness).

Investigations

(i) Standard

- **Lymph node biopsy** is required for a definitive diagnosis (Fig. 9.11).
- **Blood count** may be normal, or there can be a normochromic, normocytic anaemia. Lymphopenia and occasionally eosinophilia are present.
- **Erythrocyte sedimentation rate (ESR)** is usually raised and is an indicator of disease activity.
- **Liver biochemistry** is often abnormal, with or without liver involvement.
- **Serum lactate dehydrogenase**; raised level is adverse prognostic factor.
- **Uric acid** is normal or raised.
- **Chest X-ray** may show mediastinal widening, with or without lung involvement (Fig. 9.12).
- **CT scans** show involvement of intrathoracic nodes in 70% of cases. Abdominal or pelvic lymph nodes are also found. It is the investigation of choice for staging (Table 9.18).
- **Bone marrow aspirate and trephine biopsy** are seldom done but show involvement in patients with advanced disease. This is unusual at initial presentation (5%).

(ii) Almost standard

- **Positron emission tomography (PET)** is increasingly being used for staging, assessment of response and direction of therapy (Fig. 9.13). Despite the fact that the evidence base is still quite small, and many hypotheses remain to be validated, there can be no doubt that it is a major advance in the management of lymphoma.

Prognostic features

The presence of B symptoms confers an adverse prognosis as does advanced stage. The Hasenclever score, based on seven clinical parameters, is used to determine prognosis in advanced stage disease (Box 9.5).

Table 9.18 Cotswolds modification of Ann Arbor staging classification

Stage	Description
Stage I	Involvement of a single lymph-node region or lymphoid structure (e.g. spleen, thymus, Waldeyer's ring) or involvement of a single extralymphatic site
Stage II	Involvement of two or more lymph-node regions on the same side of the diaphragm (hilar nodes, when involved on both sides, constitute stage II disease); localized contiguous involvement of only one extranodal organ or site and lymph-node region(s) on the same side of the diaphragm (IIE). The number of anatomic regions involved should be indicated by a subscript (e.g. II_3)
Stage III	Involvement of lymph-node regions on both sides of the diaphragm (III), which may also be accompanied by involvement of the spleen (IIIS) or by localized involvement of only one extranodal organ site (IIIE) or both (IIISE)
III1	With or without involvement of splenic, hilar, coeliac or portal nodes
III2	With involvement of para-aortic, iliac and mesenteric nodes
Stage IV	Diffuse or disseminated involvement of one or more extranodal organs or tissues, with or without associated lymph-node involvement
Designations applicable to any disease state	
A	No symptoms
B	Fever (temperature > 38°C), drenching night sweats, unexplained loss of more than 10% of body weight within the previous 6 months
X	Bulky disease (a widening of the mediastinum by more than one-third of the presence of a nodal mass with a maximal dimension greater than 10 cm)
E	Involvement of a singe extranodal site that is contiguous or proximal to the known nodal site

From Diehl V, Thomas RK, Re D et al. Hodgkin's lymphoma – diagnosis and treatment. *Lancet Oncology* 2004; **5**: 19–26 with permission from Elsevier.

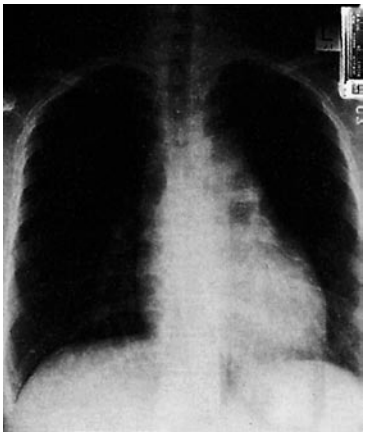


Fig. 9.12 **Chest X-ray of a large mediastinal mass that is due to Hodgkin's lymphoma.**

Management

Treatment is almost always recommended and undertaken with curative intent and considerable expectation of success (Fig. 9.14). Expectant management may be reasonable in some cases of lymphocyte-predominant Hodgkin's lymphoma, although the rationale for this must be made clear to the patient and there needs to be close early surveillance.

Specific treatment is based otherwise on the anatomical distribution of disease, its 'bulk' and the presence or absence of 'B' symptoms. ('stage': Table 9.18).

'Early stage' (IA, IIA no bulk)

The treatment of choice is brief chemotherapy followed by involved field irradiation. Large field irradiation has come under criticism because of a significantly increased incidence

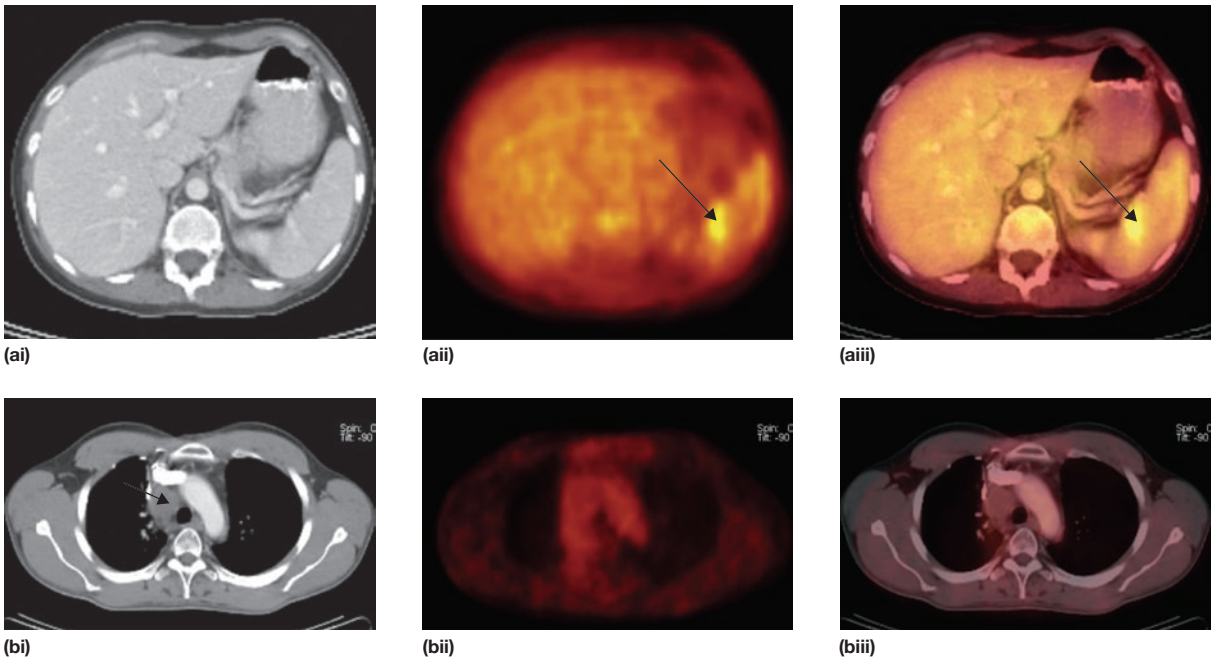


Fig. 9.13 **(a) Lymphoma in spleen detected on PET and CT/PET but not on CT. (b) Malignant lymphoma: mediastinal mass on CT scan shown to be metabolically inactive on PET, PET/CT.** Courtesy of Dr N Avril.

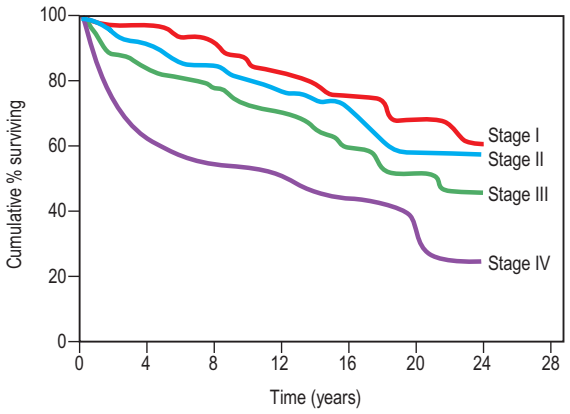


Fig. 9.14 **Survival in Hodgkin's lymphoma related to Ann Arbor stage at presentation.**

FURTHER READING

Connors JM. *Evolving Approaches to Primary Treatment of Hodgkin Lymphoma.* Washington DC: *American Society of Hematology* 2005; 239–244.

Diehl V, Thomas RK, Re D. Hodgkin's lymphoma – diagnosis and treatment. *Lancet Oncology* 2004; **5**: 19–26.

of breast cancer in young women, lung cancer in smokers and cardiac disease following supradiaphragmatic 'mantle' field irradiation. 'Moderate' chemotherapy ABVD (Table 9.10, p. 461), 2–4 cycles (i.e. non-sterilizing and low secondary cancer risk) followed by involved field irradiation (20–30 Gy) has become standard care. Current trials are evaluating the role of PET scanning to see if patients who become 'PET' negative can be spared irradiation altogether.

Advanced disease

This is also curable for a significant proportion of patients, the median survival exceeding 5 years. Cyclical combination chemotherapy with irradiation at sites of bulk disease has been standard at many centres for years. Increasingly, the data from therapeutic strategies incorporating PET scanning during and after therapy have suggested that PET-negative residual masses are likely to represent fibrous tissue, and that irradiation may be omitted (Fig. 9.13). This is a major advance. Thus it may be that the 50–60% cure rate with this

Box 9.5 Advanced stage Hodgkin's lymphoma

Clinical prognostic factors

Score	0	1
Serum albumin	≥40 g/L	<40 g/L
Haemoglobin	≥10.5 g/dL	<10.5 g/dL
Age	≤45	>45
Sex	Female	Male
Stage	<IV	IV
Leucocytosis	$<15 \times 10^9$/L	$\geq 15 \times 10^9$/L
Lymphopenia	$>0.6 \times 10^9$/L	$\leq 0.6 \times 10^9$/L

Cumulative number of points associated with increased adverse prognosis, freedom from progression at 5 years

Score 0	84%	7% (of all patients)
Score 1	77%	22%
Score 2	67%	29%
Score 3	60%	28%
Score 4	51%	12%
Score 5	42%	7%

approach may be achieved with a considerable reduction in the number of patients receiving irradiation. The major short-term toxicity of the therapy relates to myelosuppression, the mortality being around 1%, and the long-term risks being to the heart and lungs. Infertility and second malignancy are uncommon.

The above approach fails for about 25% of patients. More intensive treatment programmes have been tested, e.g. BEACOPP (Table 9.10), with better results, but with greater toxicity profiles (and greater expense), and are clearly not necessary for the majority of patients. The challenge remains to identify those who need it prior to initiation of therapy, which has proven difficult. An alternative is to develop a 'risk-adapted' therapy relying on an early assessment of response to 'minimal' therapy (ABVD) and escalating as appropriate. Trials are in prospect to direct therapy as a consequence of the result of the PET scan after 1 or 2 cycles of therapy.

Recurrent Hodgkin's lymphoma, certainly after 'conventional'-dose chemotherapy, is potentially very serious though not necessarily fatal. Hopefully, it is a declining problem. The median survival from the first recurrence is more than 10 years; it may be influenced by the duration of first remission: it may not be so good if failure occurs after more intensive initial therapy. Second and third remissions are achieved, more often than not with 'appropriate' re-induction chemotherapy. It is conventional to consolidate remission in this group when possible, with high-dose therapy and peripheral blood cell progenitor rescue (PBPCR). Registry data suggest that this may be curative in up to 50%, although follow-up does not extend beyond 15 years.

Experimental approaches for recurrent and refractory disease include new cytotoxic drugs, monoclonal antibodies and reduced-intensity allogeneic transplantation. Limited success with myeloablative chemotherapy with haemopoietic stem cell rescue has been reported in those with 'refractory disease'.

NON-HODGKIN'S LYMPHOMA (NHL)

These are malignant tumours of the lymphoid system classified separately from Hodgkin's lymphoma. Most (70%) are of B cell origin with 30% of T cell origin.

The incidence of these tumours is approximately 15/100 000 per year in developed countries, an incidence which has increased over the last 20–30 years for reasons largely unknown. There is a slight male predominance.

Aetiology

The cause is unknown. There is wide geographical variation which probably reflects different environmental factors. A number of infective agents are associated with the development of NHL.

- EBV virus DNA is observed in between 10% and 30% of tumours and is strongly associated with endemic Burkitt's lymphoma of sub-Saharan Africa, but is present in only about 20% of the sporadic cases observed in the Western world.
- The human T cell lymphotropic virus which is prevalent in Japan, Africa, South America and the Caribbean is a major risk factor for adult T cell lymphoma/leukaemia.
- Herpesvirus 8 is associated with primary effusion lymphomas and Castleman's disease.
- There is an increase in lymphoma in patients with HIV infection and more specifically primary brain lymphoma, immunoblastic diffuse large B-cell lymphoma and Burkitt's lymphoma are AIDS defining illnesses.
- *Helicobacter pylori* is an aetiological factor in gastric MALT lymphoma.

Epidemiological studies have documented an association between farming and NHL which may be a result of pesticide exposure. In addition organic solvents and hair dyes used prior to 1980 have been implicated.

Lymphomas also occur in a number of congenital immunodeficiency states. Acquired immunodeficiency, as a result of organ transplantation, is strongly associated with NHLs. These are commonly extranodal and most frequently occur in the first year after transplantation. In autoimmune disorders an elevated risk of NHL is reported, attributed to disturbances in immune function (e.g. Sjögren's syndrome and extranodal marginal zone lymphoma). A number of familial cancer syndromes are also associated with NHL (Table 9.3, p. 453).

Pathogenesis

There is a malignant clonal expansion of lymphocytes which occurs at different stages of lymphocyte development. In general, neoplasms of non-dividing mature lymphocytes are indolent whereas those of proliferating cells (e.g. lymphoblasts, immunoblasts) are much more aggressive. This malignant transformation is usually due to errors in gene rearrangements which occur during the class switch, or gene recombinations for immunoglobulins and T cell receptors. Thus, many of the errors occur within immunoglobulin loci or T cell receptor loci. For example, an abnormal gene translocation may lead to the activation of a proto-oncogene next to a promoter sequence for the immunoglobulin heavy chains (Ig-H).

Cytogenetic features (Table 9.19)

Burkitt's lymphoma was the first tumour in which a cytogenetic change was shown to involve the translocation of a specific gene. The most frequent change is a translocation between chromosomes 8 and 14 in which the *myc* oncogene moves from chromosome 8 to a position near the constant region of the immunoglobulin heavy chain gene on chromosome 14, resulting in upregulation of myc. Similar rearrangements involving the light chain loci are seen in the alternative Burkitt's lymphoma translocations between chromosome 8 and either chromosome 2 or 22. Other somatic cytogenetic abnormalities associated with human lymphoma are the t(14;18) in follicular lymphoma, involving upregulation of the Bcl-2 gene, or the upregulation of the cell cycle regulator cyclin D1 as a result of t(11;14) in mantle cell lymphoma. Gene expression profiling and other molecular techniques are increasingly revealing aberrations or patterns of aberration which are associated specifically with one or other subtype of lymphoma and have led to the identification of

Table 9.19 Chromosome translocations in non-Hodgkin's lymphoma

Type	Translocation	Genes	Function
Follicular	t(14;18)	*Bcl-2/Ig*	Suppresses apoptosis
Lymphoplasmacytic	t(9;14)	*PAX5*	Transcription factor
Mantle cell	t(11;14)	*Bcl-1* (cyclin D1)	Cell cycle regulator
Diffuse large B cell	t(3;4)	*Bcl-6*	Cell cycle regulator
Burkitt's	t(8;14) t(2;8)	*c-myc* and Ig	Transcription factor
Anaplastic	t(2;5)	*NPM1/ALK*	Tyrosine kinase
MALT	t(11;18)	*BIRC3/MALT1 fusion protein*	Suppresses apoptosis

MALT, mucosal associated lymphoid tissue.

new molecular subclasses of lymphoma. Such studies have also resulted in the identification of genes of prognostic significance.

Immunophenotypes

All NHL B cells express CD20 and surface immunoglobulin. Individual lymphomas vary in their expression, e.g. follicular lymphomas express CD10, mantle cell CD43, while the diffuse large B cell lymphoma expresses both CD10 and CD43. They can be used in the classification. T cell lymphomas do not express CD20 but variably express CD3, 4, 8 and 30.

Classification

The WHO classification (2001) which is based on the Revised European American Lymphoma (REAL) system is shown in Table 9.20. Previous classifications have divided lymphomas into indolent (or low grade) and aggressive (or high grade), and the WHO classification has been modified to include aggressive or highly aggressive lymphomas. NHL can also be staged in a similar way to Hodgkin's lymphoma (Table 9.18).

Clinical features

- *Peripheral lymphadenopathy*. Most patients present with painless, superficial lymph node enlargement.
- *Systemic symptoms (B symptoms).* Fever, sweats and weight loss.
- *Extranodal presentation*. This is more common than in HL and may involve the gastrointestinal tract, lung, brain, testes, thyroid and skin. Abdominal involvement may reveal hepatosplenomegaly. Skin involvement (T cell lymphomas) presents as mycosis fungoides (see p. 1260) and Sézary syndrome (see p. 1260).
- *Oropharyngeal involvement* occurs rarely.

Investigations

- **Full blood count**. Normochromic, normocytic anaemia, an elevated white cell count or neutropenia and thrombocytopenia are suggestive of bone marrow infiltration.
- **ESR** may be elevated.
- **Urea and electrolytes**. Patients may have renal impairment as a consequence of ureteric obstruction secondary to intra-abdominal or pelvic lymph node enlargement.
- **Serum uric acid** level may be raised.

Table 9.20 Modified WHO classification of lymphoid neoplasms other than ALL (2001)

B cell lymphomas	
Precursor B cell lymphoma	Precursor B lymphoblastic lymphoma/ leukaemia (*highly aggressive*)
Mature B cell lymphoma	Chronic lymphocytic leukaemia/small lymphocytic lymphoma Lymphoplasmacytic lymphoma Splenic marginal zone lymphoma Extranodal marginal zone B cell lymphoma of mucosa-associated lymphoid tissue (MALT-lymphoma) Nodal marginal zone B cell lymphoma Follicular lymphoma (*aggressive*) Mantle cell lymphoma Diffuse large B cell lymphoma (*aggressive*) Mediastinal (thymic) large B cell lymphoma Intravascular large B cell lymphoma Primary effusion lymphoma Burkitt's lymphoma/leukaemia (*highly aggressive*)
T/NK cell lymphomas	
Precursor T cell lymphoma	Precursor T cell lymphoblastic leukaemia/lymphoma (*highly aggressive*) Blastic NK cell lymphoma
Mature T/NK cell lymphoma	Adult T cell leukaemia/lymphoma (*very aggressive*) Extranodal NK/T cell lymphoma, nasal type Enteropathy-type T cell lymphoma Hepatosplenic T cell lymphoma Subcutaneous panniculitis-like T cell lymphoma Mycosis fungoides Sézary syndrome Primary cutaneous anaplastic large cell lymphoma Peripheral T cell lymphoma, unspecified (*aggressive*) Angioimmunoblastic T cell lymphoma Anaplastic large cell lymphoma (*aggressive*)

NK, natural killer.

Modified from Jaffe ES, Harris NL, Stein H, Vardiman JW (eds). *World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues.* Lyon: IARC Press, 2001 with permission of the World Health Organization.

- **Liver biochemistry**. This may be abnormal if there is hepatic involvement.
- **Serum lactate dehydrogenase and β_2-microglobulin** are prognostic indicators.
- **Serum immunoglobulins.** Decreased levels may occur with paraproteinaemia.
- **Chest X-ray, CT scans** of chest, abdomen and pelvis. PET and gallium scans help in staging.
- **Bone marrow aspirate and trephine biopsy** are always performed.
- **Lymph node biopsy** (or Trucut needle biopsy, often under radiological guidance, in the case of surgically inaccessible nodes). Immunophenotyping and cytogenetic/molecular analysis (DNA microarray analysis), to distinguish type of NHL.

At the completion of staging investigations, it is possible to assign a clinical prognostic index score, at least for follicular lymphoma (FLIPI, see later) and 'large cell lymphoma' (IPI).

Follicular lymphoma

These comprise 20% of all B cell lymphomas. Most patients with follicular lymphoma present feeling well but with painless lymphadenopathy. Investigation usually reveals multiple sites of disease: involvement of the bone marrow is common. Managed conservatively it is a remitting and recurring disease with a clinical course running over a median of 10 (1–20 years) years during which there will be about three 'episodes' of relapse. Death occurs because of resistant disease, transformation to diffuse large B cell lymphoma (LBCL) or the effects of therapy.

The 'well' patient should be managed with no specific therapy until progression is documented. Repeat biopsy should be performed at this time in case there has been histological transformation to LBCL as this has specific implications for therapy and prognosis.

The indications for the initiation of therapy are:

- *Stage I presentation* single lymph node field (10–15%). This is treated with 'involved' field mega voltage irradiation, which almost invariably induces 'complete remission'. The median time to progression is 10–15 years. Some patients may be cured. The therapy has a low morbidity and mortality.
- Advanced disease with:
  - constitutional 'B' symptoms
  - 'organ impairment', i.e. bone marrow failure
  - 'bulky' disease, i.e. lymph node mass > 10 cm
  - progressive disease after expectant management, documented if necessary on two scans 3 months apart
  - histological transformation.
- Philosophy of patient and physician.

A specific follicular lymphoma international prognostic index (FLIPI) has been formulated, defining 'low', 'intermediate' and 'high' risk groups. This may provide a more objective basis for the decision to begin therapy, and with what. Box 9.6 shows the prognostic factors in non-Hodgkin's lymphoma, which are very similar.

A study has shown that survival can be predicted by the different genes expressed on accompanying T cells and dendritic cells (not tumour cells) identified by gene expression microarray analysis of tumour tissue.

Management

Follicular lymphoma has been a recurring remitting disease, treated intermittently, but with a high risk over time of transformation to diffuse large B-cell lymphoma, the median survival being about 10 years. An increase in the therapeutic options has coincided over the past 10 years with an apparent improvement in the prognosis. There is now a possibility of cure, particularly after allogeneic stem cell haematopoietic transplantation.

Treatment

Chemo-immunotherapy

This has become the treatment of choice for follicular lymphoma since the demonstration in large randomized phase II clinical trials that the anti-CD20 antibody, *rituximab*, added to cyclical combination chemotherapy, improves the response rate, freedom from progression and overall survival. There is only circumstantial evidence that one chemotherapy combination is better than another. Indeed, for the long term, it has not been shown that there is any advantage of combination over single agent chemotherapy. The duration of antibody therapy after chemotherapy has been completed remains under investigation.

Other options

- **Rituximab alone**. This is being investigated in the setting of the patient 'not needing therapy', and may also be valuable in patients with progressive disease, for whom chemotherapy is inappropriate. It induces remission, mainly partial, in 30–60% of patients depending on how much previous therapy has been received.
- **Chemotherapy alone.** Both mono- and combination chemotherapy are now only being used alone when there is evidence that antibody therapy has failed.
- **Myeloablative chemo-(immunotherapy) with peripheral blood progenitor cell rescue.** This has been shown to result in very long freedom from recurrence in phase II trials when performed in second remission, and to have a survival advantage also in a phase III trial. For many physicians, it is the treatment of choice in second remission, although there are some alternative compelling data for 'maintenance rituximab'.

Box 9.6 Prognostic factors in non-Hodgkin's lymphoma

Adverse factors:

- Age > 60 years
- Stage III or IV, i.e. advanced disease
- High serum lactate dehydrogenase level
- Performance status (ECOG 2 or more)
- More than one extranodal site involved

ECOG, Eastern Cooperative Oncology Group.

Ask the authors
www.studentconsult.com

- **Irradiation for stage 1 disease** (see p. 465). Targeted irradiation, ^{131}I or ^{90}Y being delivered with anti-CD20 as a carrier is being used.
- **Allogeneic haematopoietic stem cell transplantation (HSCT).** Data suggests that this may be curative. Reduced intensity conditioning HSCT is the subject of many phase II trials. The design of phase III trials is difficult.

Lymphoplasmacytic lymphoma

This is an uncommon B cell lymphoma often presenting with extensive bone marrow infiltration, and is almost the only lymphoma to be diagnosed on bone marrow biopsy alone. There is frequently splenomegaly and anaemia, and in some an associated paraprotein. This IgM (a large pentameric molecule which at high concentrations increases plasma viscosity) is associated with immune paresis (*Waldenström's macroglobulinaemia*; WM). Patients in this group are usually older and commonly present with the symptoms of bone marrow failure or hyperviscosity. It may be a chance diagnosis.

Management may be expectant, the indications for treatment being:

- symptomatic anaemia
- recurrent infection
- symptoms of hyperviscosity, e.g. headache, visual disturbance, mucosal haemorrhage, heart failure
- progression.

Treatment is supportive in the first instance. Transfusion and/or erythropoietin is given for the anaemia, particularly if chemotherapy is being given. Plasmapheresis is an excellent means of controlling the paraproteinaemia both in the short and longer term.

Chlorambucil is the conventional treatment, the 'response criteria' for WM being different from those of the rest of the lymphomas. The paraprotein will be reduced by 50% in 50–70% of cases in the short term. Progression is the rule but as with follicular lymphoma there may be further response to treatment. The purine analogue fludarabine is also 'effective'. Rituximab is being used increasingly.

It is not clear how much the use of chemotherapy has actually influenced overall survival, and quality of life is poorly recorded. The median survival is several years.

A heightened awareness of this relatively uncommon but symptomatic lymphoma has led to more enthusiastic exploration of the new therapies being tested on the other B cell lymphomas (see above).

Mantle cell lymphoma

This is an uncommon lymphoma (6% of non-Hodgkin's lymphoma) with a median survival of 3–4 years. It occurs predominantly in the elderly and is almost always widely disseminated at presentation, with lymphadenopathy and frequent involvement of the bone marrow and gastrointestinal tract. Treatment is often compromised by co-morbidity.

Combination chemotherapy is the most usual first-line therapy, if possible: the response rate may be above 50% but the complete remission rate is low. The addition of rituximab has little effect and has not increased survival. Intensifying chemotherapy does not appear to improve survival and bortezomib, the proteasome inhibitor, is being used after failure of other therapy.

Diffuse large B cell lymphoma (DLBCL)

This is the commonest lymphoma and is almost invariably fatal without therapy within months, and was previously classified as aggressive or high-grade lymphoma. Now >50% of young patients are cured. The only indications for a palliative approach at the initial presentation are extreme co-morbidity and the will of the patient. Expectant management is inappropriate. Patients present with rapidly progressive lymphadenopathy and progressive infiltration of many organs, e.g. spinal cord, gastrointestinal tract.

Treatment

Treatment decisions are based on the stage and may be tailored by the International Prognostic Index (Box 9.6). PET scanning is useful in judging the response to treatment.

In the absence of relevant co-morbidity all patients should receive cyclical combination chemo-immunotherapy, the gold standard being cyclophosphamide, hydroxydaunorubicin, vincristine, prednisolone and rituximab (CHOP + R). Between 60% and 70% of those with early stage disease (I, IIA without bulk) may expect to be cured either with 6 cycles or 3 cycles followed by involved field irradiation. Those with more extensive disease conventionally receive 6–8 cycles. There has been a suggestion that decreasing the interval between cycles may be feasible (with growth factor support), and this may improve the results. There are conflicting reports about the advantages of increasing the intensity of initial treatment for those perceived to be at 'high risk'. There is also controversy about the indications for central nervous system prophylaxis and which form it should take.

For those with advanced disease the cure fraction, which was about 30%, has been increased with the incorporation of rituximab as standard therapy, by 15%.

Progression during therapy or failure of the initial treatment to achieve complete remission has a very poor prognosis. Second-line therapy, e.g. platinum compounds or gemcitabine, should be initiated with a view, if possible, to high-dose therapy and peripheral blood progenitor cell rescue (PBPCR) if further response is achieved. This is not usually possible and long-term success is rare. Proper consideration should be given to palliation or experimental therapy. By contrast, recurrence after a disease-free interval is potentially still curable. The complete remission rate after a second treatment is 40–50%: it is conventional to consolidate such remissions with high-dose therapy which will cure perhaps half (i.e. 25% 'cure' overall of the selected younger patients in whom curative therapy is attempted).

Burkitt's lymphoma

This is an uncommon lymphoma in the western world. It is endemic to Africa in the mosquito belt: there is a close association with the Epstein–Barr virus and it is a disease with a very high proliferative index which is very rapidly fatal without therapy.

The clinical presentation worldwide is usually that of lymphadenopathy, often with an abdominal mass and fre-

quently with bone marrow infiltration (and 'leukaemia'). Central nervous system involvement is common, up to 30% having meningitis at the time of presentation. In Africa, by far the two commonest presentations are the abdominal mass or a large tumour involving the jaw (Fig. 9.15); the bone marrow and central nervous system are also frequently involved.

Treatment

Treatment, whilst undoubtedly toxic with both morbidity and mortality, is much less dangerous than the disease. It relieves symptoms and is potentially curative in a high proportion of cases.

Supportive care is with hydration and prevention of hyperuricaemia with allopurinol associated with rapid tumour lysis. Rasburicase (see p. 485) is a major advance.

Drug therapy

Cyclical combination chemotherapy, incorporating at least high doses of cyclophosphamide and methotrexate as well as vincristine and doxorubicin, is followed by further cycles including high-dose cytarabine depending on the extent of disease at presentation. Up to 6 cycles must be given at about monthly intervals. While both the high-dose systemic methotrexate and cytarabine cross the blood–brain barrier in tumoricidal doses, many prefer to supplement them with intrathecal chemotherapy also. The role of cranial irradiation is unclear. In Africa, treatment can be difficult to coordinate but is based on the same principles.

This treatment will be tolerated quite well by the majority of patients, who tend to be young. The reported complete remission rate is very high, 70–90% depending on age (and excluding those with concurrent HIV infection), with very few recurrences occurring after 1 year. Hence 60–70% may be cured. Demonstrating improvements on this will be difficult but rituximab may have a role.

The management of progression despite treatment is difficult and rarely successful. Remission induction with alternative therapy, possibly including cisplatin, is the first line of attack. If achieved, it is consolidated with high-dose therapy or allogeneic transplantation.

T cell lymphomas

These are much less common than the B cell counterparts. In the main, the overall treatment strategies are the same, but success is much more limited.

Fig. 9.15 **A child with Burkitt's lymphoma.**

Primary extranodal lymphoma

The WHO classification does not distinguish between primarily nodal or extranodal at the time of presentation if the histological picture is the same.

Primary cerebral lymphoma

This is a very aggressive disease with survival untreated measured in months. It is particularly dangerous within the setting of HIV infection. Irradiation plus corticosteroids has, until recently, been the treatment of first choice. It is rarely curative, and may be associated with severe cerebral toxicity when given in high enough doses to eliminate lymphoma, particularly in the elderly. Even following irradiation, the median time to progression is less than a year. Subsequent management is palliative.

A very aggressive approach, involving sequential very high doses of methotrexate followed by cytarabine, may obviate the need for irradiation and be curative in a proportion of cases. Followed by irradiation this treatment may eliminate lymphoma in an even higher proportion of cases, but the potential central nervous system toxicity is considered by most to be unacceptable.

Primary gastric lymphoma

In a high proportion of cases, particularly in Northern Italy, lymphoma is associated with *Helicobacter pylori* infection (see p. 262). Biopsy of the gastric lesion usually shows lymphoma ('low-grade' B cell pathologically, of extranodal marginal zone type) and *H. pylori* is usually detected. Treatment to eradicate *H. pylori* is with antibiotics (and a proton pump inhibitor) for 2 weeks (see p. 261). Symptomatic relief is usually rapid. Provided there is no evidence of disease outside the stomach, further treatment is not given but surveillance endoscopy is performed, first at 3 months then 6-monthly. Partial or complete remission is the rule, but may take many months to achieve. Treatment of progression (always confirmed by biopsy) is with more antibiotics initially. Both irradiation and alkylating agent therapy are effective. Rituximab is being investigated. There is no role for surgery in the management of gastric lymphoma.

Primary cutaneous lymphoma (see p. 1260)

This is much commoner than generally perceived. It must be carefully distinguished from cutaneous infiltration with lymphocytes in a patient with nodal disease. It is usually of T cell origin (*Sézary syndrome, mycosis fungoides*). It is often multifocal and responds well to local therapy even when it appears histologically aggressive. Survival may be very long without chemotherapy, and in many cases treatment may only serve to disrupt the quality of life.

Lymphoma in HIV disease

NHL is the most common AIDS defining malignancy, and with the declining incidence of opportunistic infection due to highly active antiretroviral therapies has become the most common AIDS defining illness. In the main they are of B cell origin, have an aggressive course and present at unusual extranodal locations (e.g. body cavity, anogenital, sinus). The most common are Burkitt's lymphoma and the immunoblastic subtype of large B cell lymphoma; the indolent lymphomas are much less common. Those patients presenting with advanced stage disease, low CD4 counts at diagnosis, aged

> 35 or with an elevated serum lactate dehydrogenase have an inferior prognosis.

Primary CNS lymphoma, a common site, may be difficult to distinguish from toxoplasmosis on radiological assessment.

Primary effusion lymphoma (PEL), a lymphoma of the coelomic cavities, presents with pericardial and pleural effusions or ascites. Median survival even with chemotherapy is short.

Multicentric Castleman's disease is a polyclonal lymphoproliferative disorder presenting with fever, lymphadenopathy, organomegaly that frequently progresses to malignant lymphoma. It is associated with HIV and HHV-8 infection.

Treatment is difficult but CHOP therapy is usually the first-line option. This is followed by rituximab with reports of great efficacy. Antivirals, e.g. ganciclovir, are used but are palliative only.

Hodgkin's disease is increased in HIV-infected patients, although this is not an AIDS defining illness. Patients are more likely to present with B symptoms and marrow involvement.

FURTHER READING

Canellos GP, Lister TA, Young BD (eds). *The Lymphomas,* 2nd edn. Philadelphia: Saunders, 2006.

Cheson BD, Leonard JP. Monoclonal antibody therapy for B-cell non-Hodgkin's lymphoma. *New England Journal of Medicine* 2008; **359**: 613–626.

Czuczman MS. Controversies in follicular lymphoma: 'who, what, when, where, and why?' (not necessarily in that order!). *Hematology (American Society of Hematology Education Program)* 2006: 303–310.

Hennessy BT, Hanrahan EO, Daly PA. Non-Hodgkin lymphoma: an update. *Lancet Oncology* 2004; **5**: 341–353.

Sehn LH. Optimal use of prognostic factors in non-Hodgkin lymphoma. *Hematology (American Society of Hematology Education Program)* 2006: 295–302.

MYELOMA

Myeloma is a malignant disease of the plasma cells of bone marrow, accounting for 1% of all malignant disease. There is a clonal expansion of abnormal, proliferating plasma cells producing a monoclonal paraprotein, mainly IgG (55%) or IgA (20%) and rarely IgM and IgD. The paraproteinaemia may be associated with excretion of light chains in the urine (Bence Jones protein), which are either kappa or lambda. In approximately 20% there is no paraproteinaemia, only light chains in the urine. Rarely, patients produce no paraprotein or light chains (<10%).

Clinicopathological features

Myeloma is a disease of the elderly, the median age at presentation being over 60 years. It is rare under 40 years of age. The annual incidence is 4 per 100000 and it is commoner in males and in black Africans but less common in Asians. There is:

- *Bone destruction*, often causing fractures of long bones or vertebral collapse (which can cause spinal cord compression) and hypercalcaemia. Soft tissue plasmacytomas also occur and they are the usual cause of spinal cord compression in myeloma.
- *Bone marrow infiltration* with plasma cells, resulting in anaemia, neutropenia, thrombocytopenia, together with production of the paraprotein which may (rarely) result in symptoms of hyperviscosity.
- *Renal impairment* (see p. 422) owing to a combination of factors – deposition of light chains in the renal tubules, hypercalcaemia, hyperuricaemia, use of NSAIDs and (rarely), deposition of amyloid.

In addition there is a reduction in the normal immunoglobulin levels (immuneparesis), contributing to the tendency for patients with myeloma to have recurrent infections, particularly of the respiratory tract.

Cytogenetics

With fluorescent in situ hybridization and microarray techniques abnormalities are found in most cases of myeloma. Abnormalities of chromosome 13 and hypodiploidy (<45 chromosomes) have been shown to be associated with poor survival, as have t(4;14), t(14;16) and p53 (17p) deletions. t(11:14) and hyperdiploidy (>50 chromosomes) are associated with a better prognosis.

Bone disease

There is dysregulation of bone remodelling which leads to the typical lesions, usually seen in the spine, skull, long bones and ribs. In myeloma there is increased osteoclastic activity with no increased osteoblast formation of bone. Bisphosphonates that inhibit osteoclast activity are useful in myeloma but surprisingly there is no increase in bone deposition (see below).

Adhesion of stromal cells to myeloma cells stimulates the production of RANKL, IL-6, and also VEGF (which plays a role in angiogenesis). RANKL also stimulates osteoclast formation and the lytic lesions (see Fig. 9.1). Myeloma cells also produce dickkopf-1 (DKK1) which inhibits osteoblast activity and therefore production of new bone. This occurs because DKK1 binds to the Wnt co-receptor, lipoprotein receptor-related protein 5 (LRP5), inhibiting Wnt signalling and osteoblast differentiation.

Symptoms

- Bone pain – most commonly backache owing to vertebral involvement (60%)
- symptoms of anaemia
- recurrent infections
- symptoms of renal failure (20–30%)
- symptoms of hypercalcaemia
- rarely, symptoms of hyperviscosity and bleeding due to thrombocytopenia.

Patients can be asymptomatic, the diagnosis being suspected by 'routine' abnormal blood tests. Life-threatening complications are shown in Box 9.7.

Investigations

- *Full blood count*. Hb, WCC and platelet count are normal or low.
- *ESR*. This is almost always high.
- *C-reactive protein* is almost always raised.

Fig. 9.16 **Myeloma affecting the skull.** Note the rounded lytic translucencies produced by infiltration of the skull with myeloma cells.

Box 9.7 Life-threatening complications of myeloma

- Renal impairment – often a consequence of hypercalcaemia – requires urgent attention and patients may need to be considered for long-term peritoneal or haemodialysis.
- Hypercalcaemia should be treated by rehydration and use of bisphosphonates such as pamidronate.
- Spinal cord compression due to myeloma is treated with dexamethasone, followed by radiotherapy to the lesion delineated by a magnetic resonance imaging (MRI) scan.
- Hyperviscosity due to high circulating levels of paraprotein may be corrected by plasmapheresis.

- *Blood film*. There may be rouleaux formation as a consequence of the paraprotein.
- *Urea and electrolytes*. There may be evidence of renal failure (see above)
- *Serum β_2-microglobulin* > 2.5 mg/L – useful in prognosis.
- *Serum lactate dehydrogenase* (LDH) raised – useful in prognosis.
- *Serum calcium* is normal or raised.
- *Serum alkaline phosphatase* is usually normal.
- *Total protein* is normal or raised.
- *Serum albumin* is normal or low – useful in prognosis.
- *Serum protein electrophoresis* and immunofixation characteristically shows a monoclonal band.
- *Serum free light chain assay.*
- *Uric acid* is normal or raised.
- *Skeletal survey*. This may show characteristic lytic lesions, most easily seen in the skull (Fig. 9.16). CT, MRI and PET are used in plasmacytomas (bone or soft tissue deposits). MRI spine is useful if there is back pain – may show imminent compression/collapse.
- *24-hour urine electrophoresis and immunofixation* is used for assessment of light-chain excretion.
- *Bone marrow aspirate* or trephine shows characteristic infiltration by plasma cells (Fig. 9.17). Amyloid may be found.

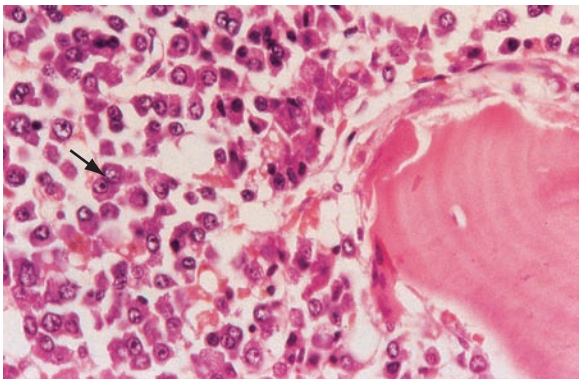


Fig. 9.17 **Multiple myeloma.** Histology shows replacement of the medullary cavity by abnormal plasma cells with some binucleate forms (arrow). A residual bony trabeculum is present towards the right. Courtesy of Dr AJ Norton.

Box 9.8 International prognostic index based on serum albumin + β_2 microglobulin (β_2M)

Stage	Median survival (months)
Stage 1 – β_2M < 3.5 mg/L and s albumin ≥ 35 g/L	62
Stage 2 – not stage 1 or 3	44
Stage 3 – β_2M ≥ 5.5 mg/L	29

Diagnosis

Two out of three diagnostic features should be present:

- paraproteinaemia or Bence Jones protein
- radiological evidence of lytic bone lesions
- an increase in bone marrow plasma cells.

For symptomatic myeloma, evidence of end organ failure should also be present, i.e. anaemia, lytic lesions, renal impairment, hypercalcaemia, recurrent infections.

An international prognostic index based on serum albumin + β_2 microglobulin (β_2M) at diagnosis is used for prognostic information (Box 9.8).

Monoclonal gammopathy of unknown significance (MGUS). MGUS describes an isolated finding of a monoclonal paraprotein in the serum, usually in the elderly; 20–30% go on to develop multiple myeloma over a 25-year period.

Treatment

With good supportive care and chemotherapy with autologous or allogeneic stem cell transplantation, median survival is now 5 years with some patients surviving to 10 years. Young patients receiving more intensive therapy may live longer. The conversion of asymptomatic to symptomatic myeloma is related to the level of myeloma protein in the blood and number of plasma cells in the bone marrow.

Supportive therapy

- Anaemia should be corrected; blood transfusion may be required. Erythropoietin often helps. Transfusion should be undertaken cautiously in patients with hyperviscosity.
- Infection should be treated promptly with antibiotics. Give yearly flu vaccinations.

- Bone pain can be helped most quickly by radiotherapy and systemic chemotherapy or high-dose dexamethasone. NSAIDs are usually avoided because of the risk of renal failure. Bisphosphonates, e.g. zoledronate, which inhibit osteoclast activity, reduce progression of bone disease.
- Pathological fractures may also be prevented by prompt orthopaedic surgery (kyphoplasty) with pinning of lytic bone lesions seen on the skeletal survey. Kyphoplasty and vertebroplasty may be useful in treating vertebral fractures.

Specific therapy

Myeloma remains incurable. Therapy is aimed at treatment of specific complications, prevention of these and prolongation of overall survival. In older or less fit patients, melphalan and prednisolone is the standard of care, with a median survival of 29–37 months; complete remissions are rare. Recent phase III studies suggest that this combined with thalidomide results in improved response rates and overall survival, albeit with increased toxicity. In younger patients (<65–70 years) a high-dose dexamethasone based induction followed by high-dose melphalan with peripheral blood stem cell rescue (autotransplantation) has a significantly higher response rate with 40% of patients achieving a CR with median survival increasing to 6 years. The role of allogeneic transplant is currently unclear.

Thalidomide has activity as a single agent and is now widely used in relapsed myeloma. Lenalidomide is a thalidomide analogue, which has recently obtained marketing authorization in the USA and Europe for relapsed myeloma. It has greater potency than thalidomide with less toxicity. Bortezomib is a proteosome inhibitor which is also licensed for relapsed myeloma. Thalidomide, lenalidomide and bortezomib all show synergy with dexamethasone, and chemotherapy and several phase II studies have shown very promising activity in newly diagnosed patients. One phase III study has shown that bortezomib plus melphalan and prednisolone is superior to the latter two alone.

FURTHER READING

Bruno B. New drugs for treatment of multiple myeloma. *Lancet Oncology* 2004; **5**: 430–442.

Sirohi B, Powles R. Multiple myeloma. *Lancet* 2004; **363**: 875–887.

COMMON SOLID TUMOUR TREATMENT

Common solid cancer mortality is listed in Table 9.1; the improvements in treatment over the past 10 years have come from advances in both prevention, diagnosis and treatment. The presentation, diagnosis and natural history of the common cancers are described in the relevant organ system chapters. In this section the systemic therapy of the common cancers is described, and in addition the treatment of germ cell tumours illustrates what can be achieved when chemotherapy resistance is overcome. The decision to treat and the aim of that treatment, whether for palliation or cure, require knowledge of the natural history of the disease, prognostic and predictive factors, the patient's performance status, and the potential efficacy of treatment. Management should be carried out by multidisciplinary teams, usually led by an oncologist.

Lung cancer (see p. 880)

Presentation, diagnosis and surgery are discussed on page 881.

Prognostic factors

Lung cancer histology is divided into two main types: *small-cell (neuroendocrine) lung cancers* (SCLC), which tend to disseminate early in their development, and *non-small-cell lung cancers* (NSCLC), which are more likely to be diagnosed in a localized form. Tumour stage and patient performance status are used in selecting treatment and predicting response and prognosis. While overall 5-year survival has remained approximately 15%, treatment is beginning to have an impact in selected groups and the multidisciplinary team can greatly aid in the appropriate application of treatment and the avoidance of nihilism.

Non-small-cell lung cancer

The staging is classified according to the TNM system (Table 9.5), by which the disease can be divided into local, locally advanced and advanced stages with 5-year survival varying from 55–67%, to 23–40%, to 1–3% respectively. The addition of CT and PET scanning has increased the accuracy of staging and improved the selection of patients for surgery and adjuvant therapy.

Treatment

In operable disease stages T1N0 to T3N2 (stage I to IIIa) adjuvant radiotherapy and chemotherapy following surgery can improve prognosis in patients of good performance status as shown by the international adjuvant lung cancer trial and a meta-analysis of 12 randomized controlled trials. Cisplatin-based combination chemotherapy induced a response in 60% and produced a relative risk reduction of 11% with an absolute improvement in 5-year survival for stage II and IIIa disease of 4% from 40.4% to 44.5%.

For unresectable disease, the combination of concurrent cisplatin with radiotherapy (chemoradiation) when compared with radiotherapy alone has increased the resection rate and 3-year survival from 11% to 23% at the expense of greater oesophageal toxicity.

In advanced disease cisplatin or carboplatin in combination with one other drug such as paclitaxel or gemcitabine for 12 weeks produces a symptomatic improvement in 40%, and increases median survival from 6 to 10 months compared with best supportive care, with 10–20% alive at 1 year. Erlotinib inhibits the EGFR tyrosine kinase and can induce responses in 10% of NSCLC particularly in adenocarcinoma in non-smoking women. However gene profiling has yet to be developed to be able to predict responses and target the use of this drug to those patients most likely to benefit.

Small-cell lung cancer

Prognostic factors

The staging of small-cell lung cancer is divided into limited and extensive disease according to whether or not it is confined to a single anatomical area or radiation field. Systemic therapy is the primary therapeutic modality because of the usually disseminated nature of the disease.

Treatment

Limited disease is present in approximately 30% of patients and is best treated with concurrent chemo- and radiotherapy using a combination of cisplatin and etoposide or irinotecan, which increases the survival at 5 years from 15% to 25% compared with radiotherapy alone. A similar degree of

Table 9.21 Breast cancer histology

Non-invasive
Ductal cancer in situ
Lobular cancer in situ
Invasive
Infiltrating ductal cancer
Infiltrating lobular cancer
Mucinous cancer
Medullary cancer
Papillary cancer
Tubular cancer
Other
Adenoid cystic, secretory, apocrine cancers
Paget's disease of the nipple
Phylloides tumour

Table 9.22 Increased risk of breast cancer

Increased for women with a first degree affected relative with:	
Premenopausal breast cancer	3.0 × average
Postmenopausal breast cancer	1.5
Postmenopausal bilateral breast cancer	5.0
Premenopausal and bilateral	9.0

Box 9.9 Poor prognostic factors for breast cancer

- Young age
- Premenopausal
- Large tumour size
- High tumour grade
- Oestrogen and progesterone receptor negative
- Positive nodes

improvement can also be achieved with hyperfractionated radiotherapy. Prophylactic whole-brain radiation to prevent cerebral metastases can reduce symptomatic CNS disease and improve overall survival by 5%.

Extensive disease can be palliated with the combination of carboplatin and etoposide or irinotecan, which when compared with best supportive care can increase median survival from 6 months to 9–13 months and 2-year survival to 20%.

Breast cancer

Breast cancer is the most common cancer in women who do not smoke. The screening programme in the UK, with biplanar digital mammography every 3 years in women aged 50–70 and improvements in multimodality treatment, have improved overall survival and rates of cure, while breast-conserving surgery has greatly ameliorated the psychosexual impact of the disease.

Aetiology and pathology

The majority of breast cancers arise from the epithelial cells of the milk ducts and reproduce their histological features in a variety of patterns (Table 9.21), of which the most common is an infiltrating ductal carcinoma. For many cancers it is thought that there is an identifiable precursor in situ stage which is confined within the basement membrane and is still truly localized and detectable by its trade mark microcalcification on a screening mammogram. For others this stage may be so brief or non-existent as to not be detectable and invasive disease is present from very early in development with a consequently worse prognosis. Approximately 10% of women with breast cancer have detectable mutations in the *BRCA 1* and *2* genes and TP53, which account for < 5% of breast cancer (Table 9.22). The hormonal environment exerts a major effect on the expression of the breast cancer potential and is related to reproductive behaviour, diet, exercise, weight, and exogenous hormones from oral contraception and postmenopausal hormone replacement therapy.

Symptoms and signs

Most women with symptomatic rather than screen-detected breast cancer present with a painless increasing mass which may also be associated with nipple discharge, skin tethering, ulceration and, in inflammatory cancers, oedema and erythema. In developing countries, 80% are likely to present with advanced disease and metastases.

Investigations

The triple assessment of any symptomatic breast mass by palpation, radiology (mammography, ultrasound and MRI scan) and fine-needle aspiration cytology is the most reliable way to differentiate breast cancer from the 15 times more common benign breast masses. Large bore core needle biopsy should follow to provide histological confirmation and predictive factors such as grade and oestrogen, progesterone and Her2 receptor status to inform the subsequent decision-making process. Assessment should be carried out in a dedicated one-stop clinic able to provide the appropriate support and referral accordingly. Staging is both surgical with respect to tumour size and axillary lymph node status and, in advanced disease, by investigation of common sites of metastasis by chest X-ray, CT/MRI, bone and liver scan. At present, only 20% of patients are diagnosed with no evidence of microscopic nodal metastases.

Prognostic factors

The following are all significant independent predictors of risk of recurrence: size of the primary tumour; the histological subtype (most are infiltrating ductal carcinoma); histological grade/differentiation; oestrogen and progesterone receptor (ER, PR) status; patient age and menopausal status. Expression of Her2 like ER and PR is a predictor of treatment response.

Gene expression profiles have identified sets of up to 50 genes, the pattern of expression of which can identify low- and high-risk subsets on which to decide adjuvant therapy. However the definitive trial to test the gene-directed therapy has yet to be conducted, in part through the continuing difficulties with the technology and inability to provide reproducible results across different laboratories.

Early breast cancer

The prognosis (10-year survival) can be predicted from the combined effect of the independent prognostic factors in Box 9.9. Survival probability and benefit from adjuvant treatment can be calculated using the website www.adjuvantonline.com that is based on the American Surveillance Epidemiology and End Results (SEER) database and has been validated on independent data sets from British

Columbia and Finland. The estimated 10-year survival probability will vary from 93–99% for small (<1 cm) low-grade node-negative tumours to only 10% for large high-grade tumours with more than nine axillary nodes involved without adjuvant therapy.

Local treatment

Surgery may vary from wide local excision or segmental mastectomy and breast conservation for masses < 4 cm in diameter, to simple mastectomy with or without reconstruction. The choice is dictated by the location and extent of the breast mass in relation to the breast size, and patient preferences. Surgery of the axilla is by sentinel lymph node guided sampling (after dye injection) in the absence of clinical or radiological (usually ultrasound) evidence of lymphadenopathy, or full dissection to level 3 if there are clinically involved nodes in order to gain local control and provide prognostic information to guide adjuvant treatment. The greater the amount of axillary surgery the more the risk of post-operative lymphoedema.

Radiotherapy is given to the conserved breast after wide local excision to reduce local recurrence, and to the chest wall after mastectomy if there are risk factors such as proximity to surgical margins or lymph node metastases, to complete the local control measures. Radiotherapy to the axilla and supraclavicular fossa can be added after sampling but not full dissection of the axilla because the combination raises the risk of severe lymphoedema to 30%. Adjuvant radiotherapy reduces the risk of local recurrence by 25% and improves 10-year survival by 3%.

Recent data suggest that women over 70 years with oestrogen receptor positive cancers up to 2 cm diameter may be offered surgery and tamoxifen alone without radiotherapy, without compromising outcome.

Adjuvant systemic treatment

Endocrine

In about one-third of patients the breast cancer will express receptors for oestrogen and progesterone. Tamoxifen (see p. 463) adjuvant therapy immediately following surgery for *receptor-positive disease* reduces the 10-year relative risk of women dying from breast cancer by about 25% and the absolute 10-year death rate by 12% for all stages.

For *postmenopausal women* with oestrogen and/or progesterone receptor-positive disease, adjuvant tamoxifen or aromatase inhibitors (AIs) such as anastrozole given for 5 years reduces the risk of death from breast cancer by a similar 25%.

The choice between AIs and tamoxifen is the subject of continuing controversy and trials. AIs avoid the adverse effects of tamoxifen on the uterus and venous thromboembolism but add risks of arthralgia and osteoporosis (Table 9.23). Aromatase inhibitors achieve a greater reduction in contralateral breast cancers than tamoxifen and in some trials a reduction in distant metastases contributing to an overall improvement in relapse-free survival which has not as yet produced a survival advantage except in one study of sequential tamoxifen and exemestane. Both tamoxifen and aromatase inhibitors cause symptoms of oestrogen deprivation such as hot flushes and vaginal dryness, but whereas tamoxifen has a beneficial effect upon serum lipid profiles the reverse is true of AIs leading to concerns over their long-term effects on women's cardiovascular health. It is recommended that the choice should be discussed for each patient taking the individual co-morbidities into account.

Table 9.23 Side-effects of endocrine agents for breast cancer in order of frequency

Tamoxifen	Aromatase inhibitors
Hot flushes	Hot flushes
Weight gain	Vaginal dryness
Mood changes	Arthralgia
Vaginal discharge	Skin rash
Thromboembolism	Osteoporosis
Endometrial hyperplasia and neoplasia	Adverse lipid profile

Chemotherapy and biological therapy

A meta-analysis of all randomized trials of adjuvant therapy in breast cancer has shown that for women with high-risk features (Box 9.9) (including receptor-negative), adjuvant chemotherapy with first generation regimens such as CMF (cyclophosphamide, 5-fluorouracil plus methotrexate) for 6 months, reduces the absolute 10-year death rate by about 10% and the relative risk of death by 20%. Ovarian ablation by GnRH analogue for 2 years is equally as effective as CMF chemotherapy in premenopausal women. More effective second generation regimens with epirubicin, (Epi-CMF – cyclophosphamide, methotrexate, 5-fluorouracil) or FEC100 (5-fluorouracil, epirubicin, cyclophosphamide) increase this to 25%. A third generation regimen with a taxane (AC-T (doxorubicin, cyclophosphamide, paclitaxel), FEC-D (fluorouracil, epirubicin, cyclophosphamide followed by docetaxel), TAC (paclitaxel, doxorubicin, carboplatin)) (Table 9.10) gives a 33% relative risk reduction but has increased toxicity. The effects of endocrine therapy are additive to those of chemotherapy and most effective given following, not concurrently with, the chemotherapy.

Menopausal status does not affect the relative efficacy of chemotherapy; however, since the risk of recurrence is lower after the menopause, the absolute improvement in survival is less. Toxicity may also be higher in this age group so that treatment decisions may need to be more individualized in discussion between the patient and her doctors. The combined effect of radiotherapy, chemotherapy and tamoxifen or aromatase inhibitor approximately halves the risk of dying of breast cancer for appropriately selected patients.

Her2/cerbB2 targeted therapy

Three recent trials of the addition of adjuvant IV trastuzumab for a further year to chemotherapy for the treatment of the 25% of patients in whom the breast cancer overexpresses Her2 have all shown a significant reduction in risk of mortality: 25% for trastuzumab alone, and 33% when administered concurrently with a taxane. Trastuzumab has been rapidly adopted as standard of care for these patients but many questions remain to be asked about the schedule of administration, toxicity to the myocardium and comparison with the oral small molecule inhibitors of EGFR such as lapatinib. Trastuzumab has a direct toxic effect upon the myocardium that is additive to pre-existing myocardial damage, especially that caused by anthracyclines, and cannot be given concurrently with them. Left ventricular ejection fraction must be monitored before and during treatment to avoid potentially severe congestive heart failure.

Neoadjuvant and primary systemic treatment

The application of systemic treatment with either endocrine or chemotherapeutic agents as above before surgery has been investigated, on the premise that it would improve survival by preventing the peri-operative release of viable tumour cells into the circulation. However randomized trials have failed to demonstrate a survival advantage when compared with postoperative treatment, i.e. the advantage over surgery alone was the same, but have identified the benefit in either rendering inoperable tumours operable (called primary systemic therapy) or large tumours smaller and suitable for breast conserving surgery in about one-third of such patients (called neoadjuvant therapy).

Advanced breast cancer

Patients with established metastatic disease may require endocrine therapy, chemotherapy and radiotherapy. The treatment is not curative but may be of great palliative benefit and consistent often with many years of good-quality life. Little additional benefit has been gained by adding endocrine and chemotherapy together, although the addition of anti-HER2 (antibodies) to chemotherapy has produced a modest survival advantage. In general, therefore, the serial use of intermittent courses of the different endocrine and chemotherapies starting with the least toxic most effective treatment seems most consistent with maintaining a good quality of life for as long as possible.

Endocrine therapy (see p. 463)

Women who have high levels of oestrogen receptors (ER) and progesterone receptors (PR) in their tumour have a greater chance of responding to endocrine treatments (i.e. 60% versus 10% for ER negative disease).

Endocrine therapy is usually tried first in those patients who have characteristics suggesting they are likely to respond and who do not have immediately life-threatening disease. Remission lasts on average 2 years and is consistent with an excellent quality of life. When relapse occurs, further treatment with alternative agents may produce another remission.

Chemotherapy (see p. 459)

Chemotherapy is used for patients who are unlikely to respond to hormonal treatment or who fail to respond to endocrine therapy or who require a rapid response if at risk of, for example, liver or respiratory failure. If chosen carefully, chemotherapy can provide good-quality palliation and prolongation of life. The drugs listed below are all able to induce objective responses in metatastatic disease and patients in whom the disease responds are likely to experience further serial responses to subsequent treatment at relapse. There is no advantage in combining more than two drugs at a time, and considerable uncertainty over the advantages of combinations with higher response rates but only rarely better survival (e.g. DC vs D) compared with single agent regimens which have the advantage of preserving more options for future use. There is very little difference in efficacy between the different regimens for metastatic disease, with response rates varying from 40% to 60% for median duration of 8–10 months. The most common regimens used include:

- MM – mitoxantrone and methotrexate
- AC/EC – doxorubicin or epirubicin and cyclophosphamide
- DC – docetaxel and capecitabine
- PG – paclitaxel and gemcitabine
- VC – vinorelbine and capecitabine.

The addition of trastuzumab and more recently lapatinib to the cytotoxic drugs (except the anthracyclines – see above) has significantly improved survival for those women whose tumour overexpresses the *c-erbB2/Her2* oncogene. For the remainder current research has identified responses to the inhibition of EGFR and VEGFR by bevacizumab and sunitinib when added to chemotherapy.The regimens do differ in toxicity, with MM being one of the least toxic. The multiple regimens provide the possibility of avoiding drug resistance over several episodes of treatment interspersed with treatment-free periods so that the disease can be palliated, often for several years.

Bisphosphonate therapy

Bone metastases are a common problem in the management of breast cancer, and the bisphosphonates have a major role in reducing the incidence of osteolytic deposits, bone pain and fracture when used preventatively, and in treating pain and hypercalcaemia from established metastases. Second generation bisphosphonates pamidronate and oral clodronate are often sufficient. The more potent third generation drugs such as zoledronate and oral ibandronate may be more effective and require less frequent administration (once every 6 months) but can be associated with impaired bone healing and osteonecrosis of the jaw (see p. 565).

Gastrointestinal cancer

Presentation, diagnosis and local treatments are described in Chapter 6.

Oesophageal cancer

Prognostic factors

Histology, stage, age and performance status are critical prognostic factors for treatment decisions, which should be made by a multidisciplinary team in designated units. The prognosis for the majority of symptomatic patients is poor, 50% have distant metastases at the time of diagnosis and the majority of the remainder will have loco-regional spread into adjacent mediastinal structures. CT-PET scans have improved the selection of patients with truly localized disease for whom curative treatment may be attempted.

Treatment

Neoadjuvant therapy for potentially resectable squamous carcinomas with cisplatin, 5-fluorouracil and concurrent radiotherapy achieves complete remission in 20–40% with a median survival of 19 months and 25–35% of patients alive 5 years after surgery. There is an increased perioperative mortality. Pre- and postoperative chemotherapy with epirubicin, cisplatin and 5FU for adenocarcinomas at the oesophago-gastric junction has improved overall survival for those patients with resectable disease and good performance status with a relative risk reduction of 25% and absolute improvement in 5-year survival by 13% from 23% to 36%.

Locally advanced or metastatic disease can be palliated with 5-fluorouracil or capecitabine chemotherapy in approximately 30%, increasing to 45–55% with the addition of oxaliplatin or irinotecan for a median duration of 6–8 months. Distressing symptomatic problems with dysphagia can be

partially relieved by endoscopic insertion of expanding metal stents to support enteral feeding or percutaneous endoscopic gastrostomy tubes and endoscopic ablation to help control bleeding. In deciding upon such measures the patient and their family understand that it does not improve survival beyond that dictated by the underlying cancer.

Gastric cancer

Presentation and diagnosis and local treatment are described on page 265.

Prognostic factors

The histological grade and staging with respect to the presence of serosal involvement (T3) and nodal involvement (N1–2), and performance status, are the main factors in prognosis. For the majority of patients with node-positive disease the 5-year survival is only 20% following surgery alone.

Treatment

Adjuvant postoperative chemo-radiotherapy with cisplatin and 5-fluorouracil compared with surgery alone showed significantly increased median survival from 28 to 35 months and 3-year survival from 41% to 50%. This was achieved mainly through improvement in loco-regional control but is suitable only for good performance status patients.

Advanced disease may be palliated with chemotherapy such as epirubicin or docetaxel combined with cisplatin and infusional 5-fluorouracil, or oxaliplatin and capecitabine in 40–50% of patients for a median of 8–10 months in good performance status patients. Supportive care for patients with upper GI cancers more than any other site must include careful attention to nutrition with the use of endoscopic stents to relieve obstruction, nasojejunal and percutaneous gastrostomy feeding tubes, and occasionally parenteral nutrition. When the disease progresses beyond active anticancer treatment management of the distressing obstructive symptoms can be helped with palliative medication to reduce secretions and at times a venting gastrostomy.

Colorectal cancer

Presentation and diagnosis and local treatment are described on page 304.

Prognostic factors

The site of the disease, above or below the pelvic peritoneal reflection, TNM stage and performance status are the main prognostic factors (see Table 6.20).

Treatment

Neoadjuvant chemo-radiation treatment of rectal cancers using cisplatin and 5-fluorouracil with radiotherapy has increased the proportion of locally advanced tumours able to be resected with clear surgical margins and reduced local relapse rate but has not improved overall survival compared with postoperative adjuvant treatment. However the preoperative use has much reduced toxicity to the other pelvic structures due to the preservation of the normal anatomical relations.

Adjuvant chemotherapy with 5-fluorouracil and folinic acid or with capecitabine for rectal and colonic adenocarcinoma significantly increases 5-year survival for node-positive disease stage III (Dukes' C) by 5% from 40% to 45%. A further 6% improvement in a recent trial was achieved by the addition of oxaliplatin, and trials are examining the addition of monoclonal antibody to the EGF receptor (cetuximab) or VEGF receptor (bevacizumab).

Advanced colorectal cancer is successfully palliated with little toxicity by 5-fluorouracil and folinic acid regimens in approximately 30% of patients for a median of 12–14 months. The addition of irinotecan or oxaliplatin increases the proportion who benefit to 55% but with increased toxicity. The monoclonal antibody bevacizumab (and possibly cetuximab) increases the response rate to chemotherapy and in one trial the median survival from 14 to 21 months.

Liver metastases are a common problem with colorectal cancers and local treatment can be accomplished with a variety of methods from surgical resection to radiofrequency, and cryo-ablation, or hepatic artery embolization or chemoembolization with chemotherapy laden microspheres. Appropriate selection of the patients who are most likely to benefit through careful restaging by MRI and PET-CT scans to exclude those with extra-hepatic disease and poor performance status is required. Small lesions can be ablated, larger lesions are best managed by partial hepatectomy, or a combination approach so that embolization is followed by hepatic regeneration before final resection. Long-term survival without recurrence is reported in up to 20% of patients at 5 years with a single < 4 cm lesion amenable to resection presenting more than a year from initial diagnosis. Current trials are examining preoperative chemotherapy before hepatic resection.

Epithelial ovarian cancer

Aetiology and pathology

There is uncertainty over the *epithelial tissue* of origin that gives rise to the 80% of all ovarian cancers that are epithelial (Table 9.24). The ovarian surface epithelium of the serosal peritoneum is most likely but others include the epithelial lining of the fallopian tubes. There is a consistent relationship between the risk of epithelial ovarian cancer (EOC) and the frequency and duration of ovulation. While not mechanistically explained this has provided a successful rationale for reducing risk of EOC by up to one-third through early pregnancy and the use of the oral contraceptive pill. The remainder are of germ cell or stromal origin although molecular

Table 9.24 Ovarian cancer pathology

Serous cystadenocarcinoma
Papillary cystadenocarcinoma
Endometrioid cancer
Adenocarcinoma
Mucinous cancer
Clear cell cancer
Mixed mesodermal Müllerian tumours
Germ cell cancers
Dysgerminoma
Embryonal cancer
Endodermal sinus tumour
Choriocarcinoma
Teratoma immature/mature
Granulosa cell tumours
Brenner, Sertoli–Leydig tumour, carcinoid tumours
Other stromal cell tumours

biological markers have shown that the category of mixed müllerian sarcoma is misnamed being an entirely epithelial tumour with metaplastic stromal elements.

Symptoms and signs

Ovarian cancer typically causes few specific symptoms, sometimes there is a sensation of a pelvic mass, which may become (acutely) painful, often there is only vague abdominal distension and epigastric discomfort. On examination the majority of patients present with a pelvic mass and advanced stage III (spread within the peritoneal cavity) or IV (extraperitoneal) disease. Screening (see p. 454) by serum CA125 tumour marker and transvaginal ultrasound scan can detect some early cancers and improve survival but at the cost of too many negative laparotomies and is thus still considered a research tool.

Investigation

Pelvic examination should be complemented by a transvaginal ultrasound and serum CA125. MRI is the definitive imaging technique for the pelvis while CT-PET scans assist in staging the patient.

Prognostic factors

Histological subtype (clear cell and mucinous are worse), grade/differentiation, stage, extent of residual disease following surgery, and performance status are all significant independent prognostic factors for survival.

Treatment

Surgery (with total abdominal hysterectomy, bilateral salpingo-oophorectomy and omentectomy) has a major role in the treatment of ovarian cancer. For patients in whom the disease is confined to the ovary, i.e. stage I, the surgery can be curative in 80–90% if the histology is well to moderately differentiated. For patients with poorly differentiated or more advanced disease, with spread throughout the peritoneal cavity, surgery still has a major role in staging the patient and improving survival when it is possible to debulk optimally to <1 cm residual. *Primary chemotherapy and delayed surgery* is an alternative approach currently under investigation and is able to render approximately one-third of inoperable patients fit for optimal debulking surgery.

Carboplatin, which is associated with fewer side-effects than cisplatin, has become the mainstay of epithelial ovarian cancer treatment. Response is achieved in approximately two-thirds of patients. Paclitaxel has been shown to improve the survival of many patients when added to a platinum-based treatment.

Adjuvant treatment with carboplatin and paclitaxel for stage I high-risk disease increases the absolute 5-year survival by 9% from 70% to 79%, and for combined stages I–III completely debulked disease by 19% from 60 to 79%, a relative risk reduction of 29%.

In more advanced disease 75% of patients will respond to combination chemotherapy and the median survival is approximately 3 years. At recurrence further palliative responses can be achieved with carboplatin and paclitaxel if the treatment free interval is >6 months, or liposomal doxorubicin, etoposide and gemcitabine are all active drugs in recurrent disease. Up to 30% of those with metastatic disease may be alive after 5 years, although this falls to 5–10% if the cancer is not able to be debulked at operation or has spread outside the peritoneal cavity. Promising new drugs being investigated are bevacizumab and erlotinib in combination with carboplatin.

Epithelial ovarian cancer tends to remain within the peritoneal cavity often exclusively which has lead to new intraperitoneal routes of administration of treatment with potentially improved survival. However at recurrence bulk disease in the peritoneum commonly causes progressive bowel obstruction which may require palliative operation or expert palliative care support to manage the terminal phases of the illness.

Prostate cancer

Early prostate cancer

Presentation and diagnosis are described on page 645.

Prognostic factors

Diagnosis is usually made on a raised serum PSA (prostate-specific antigen) followed by a transrectal ultrasound guided needle biopsy. The histological appearances are graded and accorded a Gleason Score which together with the height of the serum PSA plus accurate staging of the local extent of disease with pelvic MRI and transrectal ultrasound can identify prognostic groups (Table 9.25). This allows the selection of patients with good prognosis for no active treatment who may reasonably choose to be kept under surveillance and, like 75% of men over the age of 80, die with, but not because of, their prostate cancer.

Treatment

Patients with localized disease requiring treatment can be managed by curative surgery (radical prostatectomy), or external beam radiotherapy, or brachytherapy implants which can achieve equivalent survival rates but differ in the spectrum of unwanted side-effects with respect to incontinence and sexual dysfunction. In appropriately selected series of patients a 5-year survival of 85% can be achieved. Discussion between patient and clinician is vital to enable a treatment choice that is most appropriate to the patient's circumstances.

Adjuvant androgen deprivation treatment such as monthly depot goserelin has not improved the survival following surgery but when given before and during radiotherapy can improve the overall survival at 3 years from 62% to 78%.

Table 9.25 Prostate cancer prognostic factors

At initial diagnosis
Clinical stage
Biopsy Gleason grade
Serum PSA level
Post surgery
Surgical/pathological stage
Surgical margins
Extracapsular spread, extension to seminal vesicles, or lymph nodes
Metastatic hormone-resistant
Performance status
Serum PSA
Hb, albumin, alkaline phosphatase and LDH level

Advanced prostate cancer

Treatment

Metastatic prostate cancer with either local or most often osteoblastic skeletal spread is rapidly and effectively palliated in 70% of patients by androgen deprivation through orchidectomy, monthly depot injection of GNRH analogues goserelin or leuprorelin, or androgen receptor blockade with oral flutamide or abiraterone (see p. 646). The median duration of response is 2 years. However, on the development of androgen insensitivity, prostate cancer management must be further tailored to the patient's age, performance status and bone marrow function. These measures include simple steroids and bisphosphonates or docetaxel chemotherapy which can be guided by PSA response to provide some further palliation. Radiotherapy provides very effective palliation for the common problem of painful skeletal metastases and can be delivered by external beam therapy or systemically by intravenous bone-seeking strontium-labelled bisphosphonate for patients with multiple affected sites.

Bone pain can also be reduced with bisphosphonates such as zolendronate, as the metastatic sites comprise a combination of increased osteoclastic and osteoblastic activity.

Testicular and ovarian germ cell tumours

Germ cell tumours are the most common cancers in men aged 15–35 years but comprise only 1–2% of all cancers. They are much less common in women. There are two main histological types, seminoma (dysgerminoma in women) and teratoma. Teratomas may comprise varying proportions of mature and immature elements. Mature teratomas in women present as dermoid cysts with low malignant potential. Germ cell tumours may rarely occur in extragonadal sites in the midline from pituitary, mediastinum or retroperitoneum but should be treated in a similar manner (Table 9.26).

Symptoms and signs

Most men present with a testicular mass which is often painful, some with symptoms of metastases to the para-aortic lymph nodes with back pain. In women the mass presents with vague pelvic symptoms but at a younger age than the more common epithelial ovarian cancers.

Investigations and surgery

- Ultrasound or MRI scanning of the testicle or ovary is required.
- Assay of serum tumour markers, α-fetoprotein (AFP) and β-human chorionic gonadotrophin (β-HCG), and lactic dehydrogenase (LDH).
- CT or MRI scan for distant metastases.
- Surgery for men is by the inguinal approach to avoid spillage of highly metastatic tumour in the scrotum.
- Surgery in women for diagnosis and staging should always be conservative compared to the approach in epithelial ovarian cancer with preservation of fertility because of the efficacy of chemotherapy.

Table 9.26 Germ cell histology: WHO classification

Seminoma
Embryonal cancer
Teratoma mature, immature, or with malignant differentiation
Choriocarcinoma
Endodermal sinus tumour
Mixed teratoma

Treatment

Seminomas

Seminomas are the least common of these tumours and are very radiosensitive and chemosensitive. Seminomas are associated with a raised serum LDH but only rarely a mildly raised HCG and never a raised AFP. Stage I disease limited to the gonad is associated with a 30% risk of recurrence with surgery alone. Adjuvant therapy with either chemotherapy or radiotherapy to the para-aortic lymph nodes leads to greater than 95% cure in early-stage disease but chemotherapy with single-agent cisplatin or carboplatin does not have the long-term risks of secondary malignancy associated with radiotherapy. Combination chemotherapy (e.g. cisplatin, etoposide and bleomycin) will cure 90% of those with visible metastatic disease.

Teratomas

The risk of relapse with stage I disease varies from 5% to 40% depending upon the prognostic factors of histological differentiation and extent of local invasion.

Adjuvant chemotherapy for those at moderate to high risk (e.g. cisplatin, etoposide and bleomycin) leads to a 95% cure rate.

Metastatic disease commonly involves para-aortic lymph nodes and lungs but may spread rapidly (especially if there are trophoblastic (HCG-producing) elements present) and cause life threatening respiratory or other organ failure. HCG is associated with gynaecomastia and should be tested by the urinary pregnancy test in any young male to enable rapid institution of potentially life-saving treatment. About 80% of teratomas will express either HCG or AFP and almost all metastatic disease will be associated with an elevation of the less-specific serum marker LDH.

Chemotherapy cure for metastatic teratoma varies from over 90% for those with small-volume to 40% for those with large-volume metastases and associated rises of AFP > 10 000 and β-HCG > 100 000 IU/L.

Although approximately 20% of men will be infertile due to azoospermia at the time of diagnosis, the majority of the remainder will retain their fertility after chemotherapy and be able to father normal children. Similarly, most women retain their fertility, although less is known about the association with infertility at presentation owing to the much lower frequency of germ cell tumours in women.

Metastatic cancer of unknown primary

Patients presenting with symptoms of their metastases or with an incidental finding on imaging without a clinically obvious primary after investigation represent a common clinical problem and comprise 5–10% of patients in a specialist oncological centre. As a result of several systematic studies, some with post-mortem follow-up, the following guidance

Box 9.10 Adenocarcinoma with unknown primary: primary sites with major treatment benefits

- Breast, e.g. isolated axillary lymphadenopathy
- Ovary, e.g. peritoneal carcinomatosis
- Prostate, e.g. pelvic lymphadenopathy

should aid the choice of appropriate investigation and treatment.

Diagnosis

Diagnosis requires histology first and foremost, as it will lead to the identification of several distinct groups.

1. *Squamous cancers* – mostly presenting in the lymph nodes of the cervical region, 80% will be associated with an occult head and neck primary, the remainder arising from the lung. Inguinal nodes point usually to a primary of the genital tract or anal canal. Treatment with radiotherapy and chemotherapy may have curative potential especially in the head and neck area even in the absence of an identifiable primary on pan endoscopy.
2. *Poorly differentiated or anaplastic cancers* – this group will contain the majority of the curable cancers such as high-grade lymphomas and germ cell tumours identifiable by their immunocytochemistry and tumour markers. Gene markers such as *ip*12 for germ cell tumours and Bcl-2 for lymphomas are increasingly available to aid this diagnosis. Treatment and prognosis is as outlined in the respective primary sites.
3. *Adenocarcinomas* form the majority, and their investigation should be guided by the desire to identify the most treatable options and the knowledge that the largest proportion will have arisen from the lung or pancreas, with relatively poor treatment prospects (Box 9.10). Tissue tumour markers can be helpful (Table 9.27). Investigations should therefore comprise a chest X-ray and abdominal CT scan with, in men, serum PSA and rectal ultrasound to identify prostate cancers, and in women, mammography to identify occult breast cancer, and pelvic CT or MRI to identify ovarian cancer. For good prognosis patients wishing to have palliative chemotherapy, investigations such as endoscopy to identify lung, colon or stomach primaries are indicated to guide the choice of chemotherapy agents, although the diagnostic yield of 4–5% must be set against the discomfort and risks. Serum tumour markers for other solid cancers, although highly sensitive, are too non-specific to be useful as diagnostic aids in this situation.

Further investigation may require MRI for breast and ovarian masses, PET for head and neck, lung and possibly other primaries, and radioisotope scans for thyroid and carcinoid tumours.

Prognosis

The histological type and extent of the disease, and performance status of the patient are the key factors. Most large series report an overall median survival of 12 weeks but considerably better survival amongst the special subgroups such as patients presenting with isolated nodal metastases who have a significantly better prognosis than the majority with visceral and/or bone metastases and may warrant more extensive investigation.

Table 9.27 Adenocarcinoma of unknown primary (ACUP): immunohistochemistry markers of most probable but not exclusive tissue of origin

EMA (epithelial membrane antigen)	Epithelial
LCA (leucocyte common antigen)	Lymphoid
CEA (carcino-embryonic antigen)	Gastrointestinal cancer
CA125 peritoneal antigen	Ovarian cancer
Cytokeratin 7	Positive in breast and ovary, negative in colorectal cancer
Cytokeratin 20	Positive in colorectal, negative in breast and ovary cancer
TTF-1 (thyroid transcription factor)	Lung cancer and thyroid cancer
Thyroglobulin	Thyroid cancer
ER and PR & Her2 receptors	Breast cancer
S100, melanin and HMB45	Melanoma
Myosin, desmin and factor VIII	Soft tissue sarcoma
Chromogranin and NSE (neurone specific enolase)	Neuroendocrine cancers
AFP (alpha fetoprotein)	Germ cell tumours and hepatocellular carcinoma
βHCG (beta human chorionic gonadotrophin)	Germ cell and trophoblastic tumours
CD117	Gastrointestinal stromal tumours

Treatment

In women, an isolated axillary lymph node metastasis should be treated as for lymph-node-positive breast cancer with a similar prospect for long-term cure. Malignant ascites in women should have a trial of chemotherapy as for epithelial ovarian cancer. The prognosis for those responding to the therapeutic trial is similar to the disease of known primary origin. Primary chemotherapy that achieves an excellent response by imaging and CA125 criteria should be followed by debulking surgery with, if successful, a median survival in excess of 4 years.

For men, the occasional occult prostatic cancer found from a raised serum PSA offers some palliative treatment prospects. For the patient presenting with hepatomegaly due to metastases there is an increasing choice and efficacy of chemotherapy agents for gastrointestinal cancers that have the potential to improve their palliation unless due to melanoma.

If there is an excellent response, suitable patients may even be considered for hepatic ablation or resection. If after all efforts no primary has been identified, palliative chemotherapy treatment can achieve responses in 20–40% in highly selected series, with median survivals of 9–10 months and 5–10% surviving to 5 years.

FURTHER READING

Bradbury PA, Shepherd FA. Chemotherapy for operable NSCLC. *Lancet* 2007; **369**: 1903–1904.

Cannistra SA. Cancer of the ovary. *New England Journal of Medicine* 2004; **351**: 2519–2529.

Cunningham D, Allum WH, Stenning SP et al. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *New England Journal of Medicine* 2006; **355**: 11–20.

Geh JL, Bond SJ, Bentzen SM et al. Systematic overview of preoperative (neoadjuvant) chemoradiotherapy trials in oesophageal cancer: evidence of a radiation and chemotherapy dose response. *Radiotherapy Oncology* 2006; **78**: 236–244.

Punglia RS, Morrow M, Winer EP et al. Local therapy and survival in breast cancer. *New England Journal of Medicine* 2007; **356**: 2399–2405.

Veronesi U, Boyle B, Goldhirsch A et al. Breast cancer. *Lancet* 2005; **365**: 1727–1741.

Wagner AD, Grothe W, Haerting J et al. Chemotherapy in advanced gastric cancer: meta-analysis. *Journal of Clinical Oncology* 2006; **24**: 2903-2909.

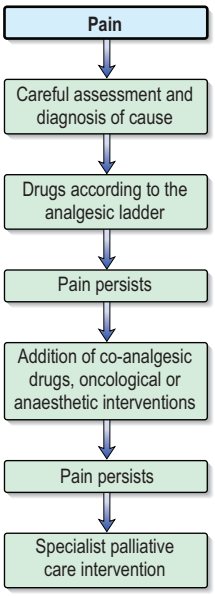


Fig. 9.18 **Management of cancer pain.**

PALLIATIVE MEDICINE AND SYMPTOM CONTROL

Palliative care may be defined as the active total care of patients with advanced, progressive life-shortening disease. The goal of care is to achieve the best possible quality of life for patients and their carers by managing not only physical symptoms, but also psychological, social and spiritual problems. Death is accepted as a normal process and the need to provide support for the family and friends in bereavement is also recognized.

Basic palliative care should be a core skill for all healthcare providers. There is increasing recognition that good palliative and supportive care can be given alongside disease-specific treatments, for example, radio/chemotherapy in patients with cancer or cardiac medication in patients with heart failure. An 'either/or' approach should be avoided. A problem-based approach that would ensure patients' access to support services, including the multi-professional specialist palliative care services for symptoms with complex aetiology, is preferred. This seems to be a major factor in appropriate supportive and palliative care in patients with non-malignant disease where referral is often left too late if palliative care is only seen as relevant for the immediate end-of-life.

In malignant disease, there is good evidence that integration of palliative care and antitumour management earlier on in the course of disease reduces long-term distress.

Good communication between doctor, patient, carer and members of the healthcare team is vital, with a shared understanding of the problems that underpin the whole approach. Assessment of the patient's understanding of disease and their wishes for future care, with acknowledgment of their concerns, will help the team plan effective support for the future. The aim is to enable the patient to be cared for and to die in the place of their choice. Good liaison between hospital, primary care and hospice is needed to allow this to happen. It is also necessary to recognize that patients will have differing needs for information, and will deal with bad news in different ways, some ways more helpful than others. A sensitive approach respecting the individual is key.

The importance of symptom assessment

A patient's symptoms are often multifactorial and a holistic assessment is imperative in order to manage them properly. The following section deals primarily with just the physical treatments.

Pain

Pain is a feared symptom and at least two-thirds suffer significant pain throughout the course of their disease. Pain may be due to a number of different causes and not all pains respond equally well to opioid analgesics (Fig. 9.18). The pain may not be directly related to the tumour, but rather indirectly, for example due to weight loss or pressure sores, or due to a co-morbidity such as arthritis. Remember that emotional and spiritual distress may be expressed as, or exacerbate, physical pain, and therefore are unlikely to respond well to opioid analgesics.

The concept of 'total pain' indicates that pain can be due to or made worse by various factors:

- Biological – from the cancer itself, therapy (drugs, surgery, radiotherapy).
- Social – financial problems from loss of job, anxiety in family and friends, loss of independence.
- Psychological – fear of dying, pain or being in hospital; anger that you are dying or at the process of diagnosis and delays; depression from all of above.

The WHO analgesic ladder

The cancer pain relief ladder guides the choice of analgesic according to pain *severity*:

Box 9.11 Drugs used for adjuvant pain relief

Drugs	Indication
NSAIDs, e.g. diclofenac	Bone pain, inflammatory pain
Anticonvulsants, e.g. gabopentin/pregabalin	Neuropathic pain
Tricyclic antidepressants, e.g. amitriptyline	Neuropathic pain, depression
Bisphosphonates, e.g. disodium pamidronate	Metastatic bone disease
Dexamethasone	Headache from cerebral oedema due to brain tumour

- *Mild pain*: non-opioid, e.g. paracetamol, ± adjuvant (Box 9.11)
- *Moderate pain*: simple analgesic, e.g. paracetamol, + weak opioid, e.g. codeine or tramadol.
- *Severe pain*: strong opioids, e.g. morphine, ± non-opioid ± adjuvant.

If regular use of optimum dosing (such as paracetamol 1 g × 4 daily for step 1) does not control the pain then an analgesic from the next step should be given. As pain may be due to different physical *aetiologies*, an appropriate adjuvant analgesic may be needed in addition to, or instead of, an analgesic ladder drug, for example, neuropathic pain may need a tricyclic antidepressant adjuvant (see below) such as amitriptyline (Box 9.11).

Strong opioid drugs

Morphine is the drug of choice and in most circumstances should be given regularly by mouth. The dose should be tailored to the individual's needs by allowing 'as required' doses; morphine does not have a 'ceiling' effect. If a patient has needed additional doses to the regular doses, then the amount can be included in the following day's regular dose until that requirement is stable – a process called 'titration'. When the stable daily dose requirement is established, the morphine can be changed to a sustained-release preparation. For example:

20 mg morphine elixir 4-hourly
=120 mg morphine per day
=60 mg twice-daily of a 12-hour preparation
or 120 mg daily of a 24-hour preparation

The starting dose of morphine will be between 5 and 10 mg every 4 hours depending on patient size, renal function and whether they are already taking a weak opioid. If there is significant renal dysfunction, morphine should be used in low doses because of the risk of metabolite accumulation (it is renally excreted). In this circumstance an alternative opioid such as fentanyl can be given by the buccal or transdermal route, e.g. 72 hour self-adhesive patches. This can be used for breakthrough pain.

If a patient is unable to take medication due to weakness, swallowing difficulties or nausea and vomiting, the opioid should be given parenterally. For cancer patients who are likely to need continuous analgesia, continuous subcutaneous infusion is the preferred route. Diamorphine is used most commonly in the UK because its high solubility enables syringe volumes to be kept to a convenient level even with high doses.

Doctors are often worried about addiction to opioids and therefore do not prescribe adequate analgesia for pain. In fact, addiction is very rare with the risk of iatrogenic addiction being <0.01%.

Side-effects and toxicity

The most common side-effects are of *nausea and vomiting*. In the majority of patients this can be managed or prevented with antiemetics (such as metoclopramide). Some of these can be given mixed with the diamorphine, e.g. haloperidol, metoclopramide, cyclizine in the same syringe, although precipitation can occur with high doses. Constipation is common and nearly always occurs with opioids and should be anticipated with lactulose, senna or co-danthramer administration. Unacceptable side-effects may persist in a small proportion of patients, and a change to a different opioid may be needed.

Confusion, persistent and undue *drowsiness, myoclonus, nightmares* and *hallucinations* indicate opioid toxicity. Sometimes this is due to over-rapid dose escalation and usually responds to dose reduction and slower re-titration. It may be due to poorly opioid responsive pain and indicates a need for adjuvant analgesics. Antipsychotics such as haloperidol may help settle the patient's distress whilst waiting for resolution of toxicity. Some patients will tolerate an alternative opioid better, e.g. oxycodone, or, if necessary by subcutaneous injection.

Adjuvant analgesics (Box 9.1)

Not all pains respond completely to opioids and additional or alternative approaches are required. The most commonly used adjuvants are:

- Non-steroidal anti-inflammatory drugs (NSAIDs) used in addition to opioids in bone pain.
- Tricyclic antidepressants (e.g. amitriptyline 10–75 mg daily) and anticonvulsants (e.g. gabapentin 600–2400 mg daily or pregabalin 150 mg at start, increasing up to 600 mg daily) for neuropathic pain.
- Steroids for nerve and inflammatory pains such as liver capsule pain. It may also be used to help the headache of raised intracranial pressure.
- Other headache treatments such as radio/chemotherapy, anaesthetic or neurosurgical intervention, acupuncture and TENS may be useful in selected patients.

Regular review is necessary to achieve optimal pain control, including regular assessment to distinguish pain severity from pain distress.

Gastrointestinal symptoms

- ***Anorexia, weight loss, malaise and weakness*** form the troublesome cancer-cachexia syndrome of advanced disease which carries a poor prognosis. Although attention to nutrition is necessary, the syndrome is mediated through chronic stimulation of the acute phase response, and specifically secreted tumour substances (e.g. lipid mobilizing factor and proteolysis inducing factor). Thus calorie-protein support alone gives very limited benefit, and parenteral feeding has

Ask the authors
www.studentconsult.com

been shown to make no difference to patient survival or quality of life. There is a small and evolving evidence base for specific therapies such as eicosapentanoic acid (EPA) fish oil, COX inhibition with an NSAID and antioxidant treatment, but currently, unless the patient is fit enough for, and responds to antitumour therapy, management is supportive. Some patients may benefit from a trial of a food supplement that contains EPA and antioxidants. Until very recently progesterones or corticosteroids were recommended and are still commonly used as an appetite stimulant, but it is increasingly recognized that weight gain is usually fluid, and particularly in the case of corticosteroids, muscle catabolism is accelerated. Also any benefit in appetite stimulation tends to be short-lived. Thus their use should be limited to short-term only, if at all, for this indication.

- ***Nausea and vomiting***. This is a common symptom, often due to opioid therapy being given without an antiemetic such as haloperidol 1.5 mg × 1–2 daily or metoclopramide 10–20 mg × 3 daily. When associated with chemotherapy, patients should have an antiemetic, starting with metoclopramide, but if the risk of nausea and vomiting is high, give a specific $5HT_3$ antagonist, e.g. ondansetron 8 mg orally or by slow i.v. injection. For most chemical causes of these symptoms, e.g. due to hypercalcaemia, haloperidol 1.5–3 mg daily is the first choice. Metoclopramide is also a prokinetic and therefore useful for emesis due to gastric stasis or constipation. Cyclizine 50 mg × 3 daily is also used.
- ***Gastric distension*** due to pressure on the stomach by the tumour (squashed stomach syndrome) is treated with an antacid and an antiflatulent, e.g. simeticone or with domperidone 10 mg × 3 daily.
- ***Persistent hiccups*** are treated with metoclopramide.
- ***Bowel obstruction***. Metoclopramide should be avoided in complete bowel obstruction where an antispasmodic such as hyoscine butylbromide (60–120 mg/24 hours) is the preferred choice. Control of vomiting in bowel obstruction may also need the addition of octreotide (a somatostatin analogue) to dry up gut secretions thus reducing the volume of vomitus, or physical measures such as defunctioning colostomy or placing of a venting gastrostomy. Occasionally, a lower bowel obstruction may be amenable to resolution by insertion of a stent, or transrectal resection of tumour. Steroids may shorten the length of episodes of obstruction if resolution is possible. In a palliative situation the patient should be allowed to drink and take small amounts of soft diet as they wish. With good mouth care, the sensation of thirst is often avoidable, thus preventing the need for parenteral fluids.

For any cause of vomiting, it may be necessary to start antiemetic therapy parenterally by continuous subcutaneous infusion to gain control of the symptom, and if the patient has gastrointestinal obstruction, this route may need to be continued.

Respiratory symptoms

Breathlessness remains one of the most distressing symptoms in palliative care. Full assessment and active treatment of all reversible conditions, such as drainage of pleural effusions, optimization of treatment for heart failure or chronic pulmonary disease is mandatory. There is often a panic–breathlessness cycle and many patients fear they will die in an acute episode. Reassurance that this is unlikely should be given. Non-pharmacological approaches such as pacing, prioritizing, breathing training, anxiety management and the use of a hand-held fan are helpful. There is little evidence to suggest that correction of hypoxaemia helps the *sensation of breathlessness* and oxygen should only be used with careful consideration of an individual patient rather than used as a blanket panacea. Opioid drugs have been shown to be useful although current evidence suggests there is no advantage with nebulized opioids. Likewise, if panic/anxiety plays a significant part, a quick-acting benzodiazepine such as lorazepam (used sublingually for rapid absorption) may be useful. Care should be taken because of the risk of poor memory and falling. The risk of respiratory depression appears to be less of a concern than previously thought.

Persistent unproductive *cough* is another troublesome symptom that can be helped by opioids (codeine or morphine elixir) due to an antitussive effect. Methadone linctus should be avoided because of its long duration of action. Excessive respiratory secretions can be treated with hyoscine hydrobromide 400–600 µg every 4–8 hours but does give a dry mouth. Glycopyrronium is also useful by subcutaneous infusion of 0.6–1.2 mg in 24 hours.

Other physical symptoms

Patients with cancer may develop other physical symptoms that are caused directly by the tumour (e.g. hemiplegia due to brain secondaries) or indirectly (e.g. bleeding or venous thromboembolism due to disturbances in coagulation). Symptoms may also result from treatment, such as lymphoedema following treatment for breast or vulval cancer, or heart failure secondary to anthracycline chemotherapy. The principles of holistic assessment, reversal of reversible factors and appropriate involvement of the multiprofessional team should be applied.

Non-physical issues

Depression is a common feature of life-limiting and disabling illness and is often missed or dismissed as ‘understandable’. However, it may well respond to usual drug and/or non-drug measures such as cognitive behavioural therapy, and make a big difference to quality of life and the ability of the patient to cope with their situation.

Extending palliative care to patients with non-malignant disease

Specialist palliative care services have traditionally grown up around cancer patients. There is increasing recognition that patients with chronic non-malignant disease such as organ failures (heart, lung and kidney), degenerative neurological disease and infectious disease (HIV) have a similar need, and benefit from a palliative care approach and access to specialist services for complex problems. There may be a less clear end-stage of disease, but the principles of symptom control are the same: holistic assessment,

reversal of reversible factors and multiprofessional support. The management as outlined above is largely transferable with some exceptions (e.g. amitriptyline and NSAIDs are contraindicated in heart failure). Joint working with the relevant consultant involved, e.g. cardiologist, allows optimization of tolerated disease-directed medication as well as palliative care. Clear communication about the stage of disease allows patients the ability to make choices about the end-of-life and to access supportive and palliative care. Discussion of difficult choices is particularly helpful for patients with neurological disease where the question of gastrostomy feeding in the event of loss of swallowing is easier to answer if this is addressed before the patient loses the ability to speak clearly, or loses mental competence to make decisions about their own healthcare.

The patient-orientated principles of palliative medicine can be usefully applied throughout all medical practice so that all patients, irrespective of care setting (home, hospital or hospice) receive appropriate care from the staff looking after them, and have access to specialist palliative services for complex issues.

In summary:

- Patients should always be involved in decisions about their care.
- The quality of life is increased when the goals of treatment are clearly understood by everyone including the patient and carer.
- The multidisciplinary team can provide a high standard of care but there must be realism and honesty at what can be achieved.
- Hospitalization is sometimes necessary but end-of-life care is often delivered in hospices.
- However care at home should be encouraged for as long as possible.

FURTHER READING

Ellershaw J, Wilkinson S. *Care of the Dying; A Pathway to Excellence.* Oxford: Oxford University Press, 2003.

Goldstein NE, Fischberg D. Update in palliative medicine. *Annals of Internal Medicine* 2008; **148**: 135–140.

Qaseem A, Snow V, Shekelle P et al. Evidence based interventions to improve palliative care of pain, dyspnea and depression at the end of life. *Annals of Internal Medicine* 2008; **148**: 141–146.

Quigley C. The role of opioids in cancer pain. *British Medical Journal* 2005; **331**: 825–829.

SIGN. *Control of Pain in Patients with Cancer. Publication number 44.* Edinburgh: SIGN, 2000. http://www.sign.ac.uk/guidelines/fulltext/44/section3.html

Twycross R. *Introducing Palliative Care.* Oxford: Radcliffe Medical, 2003.

SIGNIFICANT WEBSITES

http://www.cancer.org
US cancer organization

http://www.cancerbacup.org.uk
UK patient organization

http://www.cancerresearchuk.org/
UK charity (formed from merged Imperial Cancer Research Fund and Cancer Research Campaign)

http://www.palliativemedjournal.com
Journal of Palliative Medicine

This page intentionally left blank

10 Rheumatology and bone disease

RHEUMATOLOGICAL AND MUSCULOSKELETAL DISORDERS

The normal joint 499
Clinical approach to the patient 501
Common regional musculoskeletal problems 505
Chronic pain syndromes 516
Osteoarthritis (OA) 518
Inflammatory arthritis 522
Rheumatoid arthritis (RA) 523
Seronegative spondyloarthropathies 532
Crystal arthritis 536
Infections of joints and bones 539
Autoimmune rheumatic diseases 541
Systemic inflammatory vasculitis 548
Differential diagnosis of rheumatic complaints in the elderly 552
Arthritis in children 553
Rheumatological problems seen in other diseases 554
Miscellaneous arthropathies 556

DISEASES OF BONE

Structure and physiology 557
Calcium homeostasis and its regulation 558
Investigation of bone and calcium disorders 559
Osteoporosis 560
Osteonecrosis 565
Paget's disease 565
Rickets and osteomalacia 567
Skeletal dysplasias 568

RHEUMATOLOGICAL AND MUSCULOSKELETAL DISORDERS

Many common locomotor problems are short-lived and self-limiting or settle with a course of simple analgesia and/or physical treatment; for example, physiotherapy or osteopathy. However, they represent 20–30% of the workload of the primary care physician, where non-inflammatory problems predominate. Recognition and appropriate early treatment of many painful rheumatic conditions may help reduce the incidence of chronic pain disorders. Early recognition and subsequent treatment of inflammatory arthritis by specialist multidisciplinary teams leads to better symptom control and prevents long-term joint damage and disability. A wide selection of pamphlets and websites offer helpful advice for patients, and their use should be encouraged. Most of the musculoskeletal diseases are seen world-wide, although the prevalence of individual conditions varies.

THE NORMAL JOINT

There are three types of joints – fibrous, fibrocartilaginous and synovial.

Fibrous and fibrocartilaginous joints

These include the intervertebral discs, the sacroiliac joints, the pubic symphysis and the costochondral joints. Skull sutures are fibrous joints.

Synovial joints (Fig. 10.1)

These include the ball-and-socket joints (e.g. hip) and the hinge joints (e.g. interphalangeal).

They possess a cavity and permit the opposed cartilaginous articular surfaces to move painlessly over each other. Movement is restricted to a required range, and stability is maintained during use. The load is distributed across the surface, thus preventing damage by overloading or disuse.

Synovium and synovial fluid. The joint capsule, which is connected to the periosteum, is lined with synovium which is a few cells thick and vascular. Its surface is smooth and non-adherent and is permeable to proteins and crystalloids. As there are no macroscopic gaps, it is able to retain normal joint fluid even under pressure. Macrophages and fibroblast-like synoviocytes form the synovial layer by cell-to-cell interactions mediated by cadherin-II. The synoviocytes release hyaluronan into the joint space, which helps to retain fluid in the joint. Synovial fluid is a highly viscous fluid secreted by the synovial cells and has a similar consistency to plasma. Glycoproteins ensure a low coefficient of friction between the cartilaginous surfaces. Tendon sheaths and bursae are also lined by synovium.

Juxta-articular bone

The bone which abuts a joint (epiphyseal bone) differs structurally from the shaft (metaphysis) (see Fig. 10.31). It is highly

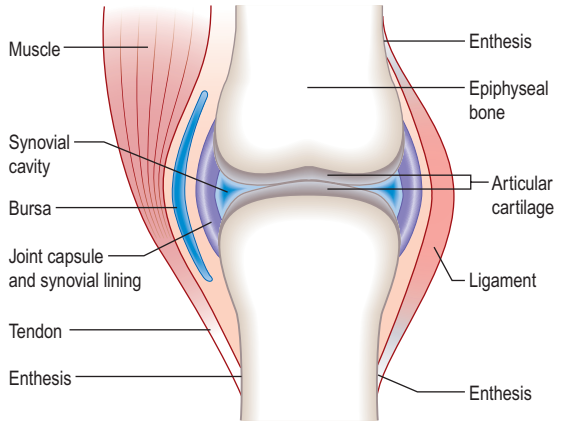


Fig. 10.1 **Normal synovial joint.**

vascular and comprises a light framework of mineralized collagen enclosed in a thin coating of tougher, cortical bone. The ability of this structure to withstand pressure is low and it collapses and fractures when the normal intra-articular covering of hyaline cartilage is worn away – as, for example, in osteoarthritis (OA). Loss of surface cartilage also leads to the abnormalities of bone growth and remodelling typical of OA (see p. 518).

Hyaline cartilage

This forms the articular surface and is avascular. It relies on diffusion from synovial fluid for its nutrition. It is rich in type II collagen that forms a meshwork enclosing giant macromolecular aggregates of proteoglycan. These heterogeneous macromolecules comprise protein chains (aggrecans) to which are attached side-chains of the carbohydrates keratan and chondroitin sulphate. These molecules retain water in the structure by producing a dynamic tension between the retaining force of the collagen matrix and the expansive effect of osmotic pressure. Intermittent pressure from 'loading' of the joint is essential to normal cartilage function and encourages movement of water, minerals and nutrients between cartilage and synovial fluid. Chondrocytes secrete collagen and proteoglycans and are embedded in the cartilage. They migrate towards the joint surface along with the matrix they produce.

Ligaments and tendons

These structures stabilize joints. Ligaments are variably elastic and this contributes to the degree of stiffness or laxity of joints (see p. 554). Tendons are inelastic and transmit muscle power to bones. The joint capsule is formed by intermeshing tendons and ligaments. The point where a tendon or ligament joins a bone is called an *enthesis* and may be the site of inflammation.

Components of extracellular matrix

All connective tissues contain an extracellular matrix of macromolecules – collagens, elastins, non-collagenous glycoproteins and proteoglycans – in addition to cells, e.g. synoviocytes. There are several different types of cell surface receptors that bind extracellular matrix proteins including the integrins, CD44 and the proteoglycan family of receptors, e.g. *syndecans*.

Collagens. Collagens consist of three polypeptide chains (alpha chains) wound into a triple helix. These alpha chains contain repeating sequences of Gly–x–y triplets, where x and y are often prolyl and hydroxyprolyl residues. Collagen fibres show considerable genetic heterogeneity, with genes on at least 12 chromosomes. The majority of collagen in the body is type I – the major component of bone, tendon, ligament, skin, sclera, cornea, blood vessels and the hollow organs. Types III, V and VI are also present in most tissues although little collagen type III is found in bone or cartilage. Other types of collagen are tissue specific, i.e. types II, IX, X and XI are found in hyaline cartilage, type IV and XVII in basement membrane and type VII in anchoring structures at junctions between epithelium and mesenchyme. There are several classes of collagen genes, based on their protein structures, and abnormalities of these may lead to specific diseases (see p. 568).

Elastin is an insoluble protein polymer and is the main component of elastic fibres. Tropoelastin, its precursor, is synthesized by vascular smooth muscle cells and skin fibroblasts. Cross-linkages with desmosine and isodesmosine are specific to elastin fibres.

Glycoproteins. Fibronectin is the major non-collagenous glycoprotein in the extracellular matrix. Its molecule contains a number of functional domains, or cell recognition sites that bind ligands and are involved in cellular adhesion. A peptide sequence (*Arg–Gly–Asp*), which mimics some of the functions of fibronectin, is also found in other adhesion proteins (e.g. vitronectin, laminin and collagen type VI). Fibronectin plays a major role in tissue remodelling. Its production is stimulated by interferon-gamma and by transforming growth factor-beta and inhibited by tumour necrosis factor and interleukin-1.

Proteoglycans. These proteins contain glycosaminoglycan (GAG) side-chains and are of variable form and size. Many have been identified at different sites in connective tissue, e.g. aggrecan, biglycan, fibromodulin, decorin (in extracellular matrix), syndecan, CD44, fibroglycan (on cell surfaces), cerebroglycan (in brain), serglycan (in intracellular tissues) and perglycan (in basement membranes). Their function is to bind extracellular matrix together, retain soluble molecules in the matrix and assist with cell binding. Abnormalities of any of these structures may lead to periarticular or articular symptoms and/or predispose to the development of arthritis.

Joint sensation

The ligaments, periosteum, synovial tissue and capsule of the joint are richly supplied by blood vessels and nerves. Pain usually derives from inflammation of these sites because the synovial membrane is relatively insensitive.

Connective tissue degradation

Connective tissue undergoes repair and re-modelling by enzymes (mainly matrix metalloproteinases, MMPs) which require zinc and act at a neutral pH. There are several MMPs which act on different collagens, e.g. the gelatinases (MMP-2 and -9), which degrade denatured collagen. MMPs also act on non-collagen proteins, e.g. the stromelysins (MMP-3, -10, -11), which degrade proteoglycans and fibronectin. Other members of the MMP group include ADAMs (**a** **d**isintegrin **a**nd a **m**etalloproteinase) which are localized to the cell surface and play a role in matrix degradation.

The turnover of normal collagen is initiated by cytokines, e.g. interleukin-1 synthesized by chondrocytes. Activation of latent MMPs and tissue plasminogen activator then occurs.

Two inhibitors, TIMP (**t**issue **i**nhibitor of **m**etallo**p**roteinase) and plasminogen activator inhibitor-1 (PAI-1), inhibit degradation during matrix remodelling.

Skeletal muscle

This consists of bundles of myocytes containing actin and myosin molecules. These molecules interdigitate and form myofibrils which cause muscle contraction in a similar way to myocardial muscle (p. 683). Bundles of myofibrils (fasciculi) are covered by connective tissue, the perimysium, which merges with the epimysium (covering the muscle) and forms the tendon which attaches to the bone surface (enthesis).

CLINICAL APPROACH TO THE PATIENT

TAKING A MUSCULOSKELETAL HISTORY

The following questions are helpful in assessing the problem and making a diagnosis. A history can often lead to a diagnosis as pattern recognition is the key to diagnosis in rheumatic diseases.

Pain

- *Where is it? Is it localized or generalized?* The pattern of joint involvement is a useful clue to the diagnosis (e.g. distal interphalangeal joints in osteoarthritis).
- *Is it arising from joints, the spine, muscles or bone?* Soft tissue lesions and inflamed joints are locally tender.
- *Could it be referred from another site?* Joint pain is localized but may radiate distally – shoulder to upper arm; hip to thigh and knee.
- *Is it constant, intermittent or episodic? How severe is it – aching or agonizing?* For example, the pain of gout, or of septic arthritis in a previously fit, non-immunocompromised patient is agonizing. Joint pain and swelling lasting a day or so may indicate palindromic rheumatism, whilst longer bouts of a few days are typical of untreated gout or pseudogout. Constant pain, especially pain at night, may be due to an underlying malignancy.
- *Are there aggravating or precipitating factors?* For example:
  - Mechanical problems are made worse by activity and eased by rest.
  - Inflammatory joint and spine pain are worse after rest and improve with activity.
  - Trauma is a common cause of musculoskeletal pain.
- *Are there any associated neurological features?* Numbness, pins and needles and/or loss of power suggest 'nerve' involvement. Examples include carpal tunnel syndrome (see p. 508) or a spinal problem such as disc prolapse (p. 510). Nerve root pain, such as that due to a disc prolapse, reflects the anatomical distribution of the affected root.

Stiffness

- *Is it generalized or localized*? Spine or joint stiffness is common after injury.
- *Does it affect the limb girdles or periphery?*
- *Is it worse in the morning and relieved by activity?* Joints that are stiff for more than 15 minutes each morning are usually inflamed – think of rheumatoid arthritis (RA) (see p. 525) or another cause of inflammatory arthritis. Spinal stiffness and pain which is much worse in the morning may indicate ankylosing spondylitis (p. 533), especially in patients in their twenties or thirties. Shoulder and pelvic girdle stiffness and pain, which are worse in the morning in a patient over 55 years, may be polymyalgia rheumatica (p. 549).

Swelling

- *Is it of one joint, or of several*? Look for symmetry or asymmetry, and/or a peripheral or proximal pattern; these are clues to the type of arthritis. Rheumatoid arthritis is typically polyarticular. An acute monoarthritis may be due to trauma, gout (in a middle-aged male) or sepsis (fever or immunosuppression).
- *Is it constant or episodic*?
- *Are episodes of swelling short-lived, or longer*?
- *Is there associated inflammation (redness and warmth)*?

Gender

Gout (see p. 536), reactive arthritis (p. 535) and ankylosing spondylitis (p. 532) are more common in men. Rheumatoid arthritis and other autoimmune rheumatic diseases are more common in women.

Age

- *Is the person young, middle-aged or older?* Injury is common in young people but can occur at any age.
- *How old was the patient when the problem first started?* Osteoarthritis (see p. 518) and polymyalgia rheumatica (p. 549) rarely affect the under-fifties. Rheumatoid arthritis starts most commonly in women aged 20–50 years.

General health

- *Is there any associated ill-health or other worrying feature, such as weight loss or fever?* Systemic illness is a common feature of many rheumatic diseases. If there is weight loss and/or fever, think of autoimmune rheumatic disease, sepsis (joint infection may be due to septicaemia and is a medical emergency) or malignancy.
- *Are there other associated medical conditions that may be relevant?* Psoriasis (see p. 1240) or inflammatory bowel disease is associated with asymmetrical arthritis. Charcot's joints (p. 555) are seen in diabetics.

Medication

Could a drug be a cause? Diuretics may precipitate gout in men and older women. Hormone replacement therapy or the oral contraceptive pill may precipitate systemic lupus erythematosus (SLE) (p. 541). Steroids can cause avascular necrosis. Some drugs cause a lupus-like syndrome (p. 542).

Race

Is this relevant? Sickle cell disease causes joint pain in young black Africans, but osteoporosis (see p. 560) is uncommon in older black Africans.

Past history

Have there been any similar episodes or is this the first? Are there any clues from previous medical conditions? Gout is

recurrent; the episodes settle without treatment in 7–10 days. Acute episodes of palindromic rheumatism may predate the onset of rheumatoid arthritis (see p. 525).

Family history

Does anyone in the family have a similar problem or another related disorder? Osteoarthritis may be familial. Seronegative spondyloarthropathies (see p. 532) are seen in families with a history of arthritis, psoriasis, ankylosing spondylitis, iritis or inflammatory bowel disease. Autoimmunity has a familial tendency.

Occupational history

What job does the patient do? This can be a factor in soft tissue problems and osteoarthritis (e.g. in heavy labourers and dancers). Work-related problems, particularly in those who use a keyboard, are becoming more common and are complained of more.

Psychosocial history

The biopsychosocial model of disease is highly relevant to many rheumatic disorders:

- *Has there been any recent major stress in family or working life? Could this be relevant?* Stress rarely causes rheumatic disease but may precipitate a flare-up of inflammatory arthritis. Stress also tends to reduce a person's ability to cope with pain or disability. Remember that the diagnosis of a chronic arthritis has a major influence on the lifestyle of patients and their families. The extent of disability should be noted.
- *Has there been an injury for which a legal case for compensation is pending*?

Extent of disability

The World Health Organization describes the impact of disease on an individual in terms of:

- *Impairment*: any loss or abnormality of psychological or anatomical structure or function.
- *Disability* (activity limitation): any restriction or lack of ability to perform an activity in the manner or within the range considered normal for a human being.
- *Handicap* (participation restriction): a disadvantage for an individual resulting from an impairment or disability that limits or prevents the fulfilment of a role that is normal for that individual.

The patient's own perception of limitation must be taken into account during assessment, as well as the impact of physical causes due to disease. Subjective and objective assessments must be made. Quality of life (QoL) involves physical and psychosocial factors. The aim of treatment is to reduce or cure physical and/or psychological disease and to reduce the impact of any impairment or disability on the individual. A variety of different standard questionnaires can be used to assess pain, disease impact and outcome (e.g. Health Assessment Questionnaire (HAQ), Arthritis Impact Measurement Scale (AIMS)).

EXAMINATION OF THE JOINTS

Always observe a patient, looking for disabilities, as he or she walks into the room and sits down. General and neurological examinations are often necessary. Guidelines for rapid examinations of the limbs and spine are shown in Practical box 10.1.

Examining an individual joint involves three stages – looking, feeling and moving:

- *Appearance*. Look at it for swelling, rash or erythema, muscle wasting, deformity such as a distal bone displaced laterally as in knock knees (genu valgus) or bowed legs (genu varus), fixed flexion or hyperextension, loss of normal range and lack of fluidity of movement, and any pain caused by movement.
- *Feel* it for tenderness, warmth (indicates inflammation) and swelling which may be due to fluid, soft tissue or bone. Common descriptors are 'fluctuant' (fluid), 'firm' or 'boggy' (swelling of the synovium), and 'hard' (bony).

Practical Box 10.1 Rapid examinations of the limb and spine

Rapid examination of the upper limbs

- *Raise arms sideways to the ears (abduction). Reach behind neck and back.* Difficulties with these movements indicate a shoulder or rotator cuff problem.
- *Hold the arms forward, with elbows straight and fingers apart, palm up and palm down.* Fixed flexion at the elbow indicates an elbow problem. Examine the hands for swelling, wasting and deformity.
- *Place the hands in the 'prayer' position with the elbows apart.* Flexion deformities of the fingers may be due to arthritis, flexor tenosynovitis or skin disease. Painful restriction of the wrist limits the person's ability to move the elbows out with the hands held together.
- *Make a tight fist.* Difficulty with this indicates a loss of flexion or grip. Grip strength can be measured.

Rapid examination of the lower limbs

- *Ask the patient to walk* a short distance away from and towards you, and to stand still. Look for abnormal posture or stance.
- *Ask the patient to stand on each leg.* Severe hip disease causes the pelvis on the non-weight-bearing side to sag (positive Trendelenburg test).
- *Watch the patient stand and sit,* looking for hip and/or knee problems.
- *Ask the patient to straighten and flex each knee.*
- *Ask the patient to place each foot in turn on the opposite knee with the hip externally rotated.* This tests for painful restriction of hip or knee. Abnormal hips or knees must be examined lying.
- *Move each ankle up and down.* Examine the ankle joint and tendons, medial arch and toes whilst standing.

Rapid examination of the spine

Stand behind the patient.

- *Ask the patient to (a) bend forwards to touch the toes with straight knees, (b) extend backwards, (c) flex sideways, and (d) look over each shoulder, flexing and extending and side-flexing the neck.* Observe abnormal spinal curves – scoliosis (lateral curve), kyphosis (forward bending) or lordosis (backward bending). A cervical and lumbar lordosis and a thoracic kyphosis are normal. Muscle spasm is worse whilst standing and bending. Leg length inequality leads to a scoliosis which decreases on sitting or lying (the lengths are measured lying).
- *Ask the patient to lie supine.* Examine any restriction of straight-leg raising (see disc prolapse, p. 511).
- *Ask the patient to lie prone.* Examine for anterior thigh pain during a femoral stretch test (flexing knee whilst prone), which indicates a high lumbar disc problem.
- *Palpate* the spine and buttocks for tender areas.

- *Movement*. Move it to assess the passive range of movement (e.g. flexion, extension, abduction, adduction and rotation), any instability, or the production of pain and crepitus (grating) seen with cartilage damage. The normal range varies between individuals. Comparing right with left and asking the patient about any change in range help to assess whether the endpoints are normal or not. A screening examination of the locomotor system, known by the acronym GALS (global assessment of the locomotor system) has been devised.

X-ray of the joint often forms an integral part of the examination.

INVESTIGATIONS

Investigations are unnecessary in many of the common regional musculoskeletal problems and osteoarthritis (OA); the diagnosis is clear from the history and examination findings. Tests help to exclude another condition and to reassure the patient or their primary care physician.

Useful blood screening tests

- **Full blood count**
  Haemoglobin. Normochromic, normocytic anaemia occurs in chronic inflammatory and autoimmune diseases. Hypochromic, microcytic anaemia indicates iron deficiency, sometimes due to non-steroidal anti-inflammatory drug (NSAID) induced gastrointestinal bleeding.
  White cell count. Neutrophilia is seen in bacterial infection (e.g. septic arthritis). It also occurs with corticosteroid treatment. Lymphopenia occurs with viral illnesses or active systemic lupus erythematosus (SLE). Neutropenia may reflect drug-induced bone marrow suppression. Eosinophilia is seen in the Churg–Strauss syndrome (p. 870).
  Platelets. Thrombocythaemia occurs with chronic inflammation. Thrombocytopenia is seen in drug-induced bone marrow suppression.
- **Erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP).** An increase of these reflects inflammation and acute phase reactants. Plasma viscosity is also raised in inflammatory disease and measured in some laboratories in place of the sedimentation rate.
- **Bone and liver biochemistry.** A raised serum alkaline phosphatase may indicate liver or bone disease. A rise in liver enzymes is seen with drug-induced toxicity. For other investigations of bone, see page 559.

Other blood and urine tests

- *Protein electrophoretic strip (and/or immunofixation) and urinary Bence Jones protein* – to exclude myeloma as a cause of a raised ESR.
- *Serum uric acid* – for gout.
- *Antistreptolysin-O titre* – in rheumatic fever.

Serum autoantibody studies

- **Rheumatoid factors (RFs)** (see also p. 524)**.** IgM rheumatoid factors were detected by agglutination tests using IgG-coated latex particles (the Rose–Waaler test) or sensitized sheep red cells (sheep cell agglutination test or SCAT) but increasingly the ELISA technique is being used. RFs are antibodies (usually IgM, but also IgG or IgA) against the Fc portion of IgG. They are detected in 70% of patients with rheumatoid arthritis (RA), but are not diagnostic. RFs are detected in many autoimmune rheumatic disorders (e.g. SLE), in chronic infections, and in asymptomatic older people (Table 10.1).
- **Cyclic citrullinated peptide antibodies (anti-CCP).** These are measured by an ELISA technique and are present in up to 80% of patients with RA. They have a high specificity for RA (90% with a sensitivity of 60%). They are helpful in early disease when the RF is negative to distinguish it from acute transient synovitis (see Box 10.6, p. 525). Both RF and anti-CCP are positive in established RA in more than 90% of patients.
- **Antinuclear antibodies (ANAs).** These are detected by indirect immunofluorescent staining of fresh-frozen sections of rat liver or kidney or Hep-2 cell lines. Different patterns reflect a variety of antigenic specificities that occur with different clinical pictures (see Box 10.11, p. 544). ANA is used as a screening test for SLE, but low titres occur in RA and chronic infections and in normal individuals, especially the elderly (Table 10.2).
- **Anti-double-stranded DNA (dsDNA) antibodies.** These are usually detected by a precipitation test (Farr assay), by ELISA, or by an immunofluorescent test using *Crithidia luciliae* (which contains double-stranded DNA). They are diagnostic of active SLE but may be negative in mild or inactive disease. High titres of IgG anti-dsDNA indicate a poor prognosis and are specific to SLE. They are probably pathogenic. Anti-single-stranded DNA antibodies are non-specific.
- **Anti-extractable nuclear antigen (ENA) antibodies** (see Box 10.11, p. 544)**.** These produce a speckled ANA fluorescent pattern, which can also be distinguished by ELISA.

There are seven types of antibodies to tRNA synthetase enzymes but only a few are available for routine analysis. *Anti*

Table 10.1 Conditions in which rheumatoid factor is found in the serum

Autoimmune rheumatic diseases	**RF (IgM) %**
Rheumatoid arthritis	70
Systemic lupus erythematosus	25
Sjögren's syndrome	90
Systemic sclerosis	30
Polymyositis/dermatomyositis	50
Juvenile idiopathic arthritis	Variable
Viral infections	**Hyperglobulinaemias**
Hepatitis	Chronic liver disease
Infectious mononucleosis	Sarcoidosis
Cryoglobulinaemia	
Chronic infections	**Normal population**
Tuberculosis	Elderly
Leprosy	Relatives of patients with RA
Infective endocarditis	
Syphilis	

Table 10.2	Conditions in which serum antinuclear antibodies (%) are found
Systemic lupus erythematosus	95
Systemic sclerosis	70
Sjögren's syndrome	80
Polymyositis and dermatomyositis	40
Rheumatoid arthritis	30
Juvenile idiopathic arthritis	Variable
Other diseases	
Autoimmune hepatitis	100
Drug-induced lupus	>95
Myasthenia gravis	50
Idiopathic pulmonary fibrosis	30
Diabetes mellitus	25
Infectious mononucleosis	5–10
Normal population	8

Jo-1 antibodies are found in polymyositis and dermatomyositis. *Anti-RNA polymerase antibodies I and III* are present in systemic sclerosis. *Anti-topoisomerase 1 (ScL-70) antibodies* are seen in systemic sclerosis.

- **Anti-neutrophil cytoplasmic antibodies (ANCAs)** (see p. 551)**.** These are predominantly IgG autoantibodies directed against the primary granules of neutrophil and macrophage lysosymes. Two major clinically relevant ANCA patterns are recognized on *immunofluorescence*:
  - proteinase 3 (PR3-ANCA), also called cytoplasmic or cANCA, producing a granular immunofluorescence
  - myeloperoxidase (MPO-ANCA), also called perinuclear or pANCA, producing a perinuclear stain.

Both are strongly associated with small-vessel vasculitides: PR3-ANCA in Wegener's granulomatosis, and MPO-ANCA in e.g. microscopic polyarteritis (polyangiitis) and Churg–Strauss syndrome. An MPO-ANCA is also found in inflammatory bowel disease.

- **Antiphospholipid antibodies** (see p. 599)**.** These are detected in the antiphospholipid syndrome (see p. 544).
- **Immune complexes.** Immune complexes are infrequently measured, largely because of variability between assays and difficulty in interpreting their meaning. Assays based on the polyethylene glycol precipitation method (PEG) or C1q binding are available commercially.
- **Complement.** Low complement levels indicate consumption and suggest an active disease process in SLE.

FURTHER READING

Male D, Brostoff J, Roth DB, Roitt I. *Immunology,* 7th edn. St Louis: Elsevier, 2006.

Joint aspiration (Practical box 10.2)

Examination of joint (or bursa) fluid is used mainly to diagnose septic, reactive or crystal arthritis. The nature of the fluid is an indicator of the level of inflammation. Clear fluid indicates little inflammation in the joint, whereas translucent or opaque fluid indicates increasing cellularity and underlying inflammation. Purulent fluid is seen in septic arthritis, but crystal arthritis and reactive arthritis may also produce a highly cellular effusion. The procedure is often undertaken in combination with injection of a corticosteroid. Aspiration alone is therapeutic in crystal arthritis.

Practical Box 10.2 Joint aspiration

This is a sterile procedure which should be carried out in a clean environment

Explain the procedure to the patient; obtain consent.

1. Decide on the site to insert the needle and mark it.
2. Clean the skin and your hands scrupulously; remove rings and wristwatch. Put on gloves.
3. Draw up local anaesthetic (and corticosteroid if it is being used) and then use a new needle.
4. Warn the patient, insert the needle, injecting local anaesthetic as it advances and, if a joint effusion is suspected, attempt to aspirate as you advance it.
5. If fluid is obtained, change syringes and aspirate fully.
6. Examine the fluid in the syringe and decide whether or not to proceed with a corticosteroid injection (if fluid clear or slightly cloudy).
7. Cover the injection site and advise the patient to rest the affected area for a few days. Warn the patient that the pain may increase initially but to report urgently if this persists beyond a few days, if the swelling worsens, or if they become febrile, since this might indicate an infected joint.

Examination of synovial fluid

Aspiration and analysis of synovial fluid are always indicated when an infected or crystal induced arthritis is suspected, particularly a monoarthritis. Normal fluid is straw coloured and contains <3000 WCC/mm^3. Inflammatory fluid is cloudy, viscous and contains >3000 WCC/mm^3. Septic fluid is opaque and less viscous (it contains hyaluronidase ++) and contains up to 75 000 WCC/m^3. There is much overlap.

Polarized light microscopy is performed for crystals.

- Gout – negatively birefringent, needle-shaped crystals of sodium urate.
- Pyrophosphate arthropathy (pseudogout) – rhomboidal, weakly positively birefringent crystals of calcium pyrophosphate.

Gram staining is essential if septic arthritis is suspected and may identify the organism immediately. Joint fluid should be cultured and antibiotic sensitivities requested.

Diagnostic imaging and visualization

- **X-rays** can be diagnostic in certain conditions (e.g. established rheumatoid arthritis) and are the first investigation in many cases of trauma. X-rays can detect joint space narrowing, erosions in rheumatoid arthritis, calcification in soft tissue, new bone formation, e.g. osteophytes and decreased bone density (osteopenia) or increased bone density (osteosclerosis):
  (a) In acute low back pain, X-rays are indicated only if the pain is persistent, recurrent, associated with neurological symptoms or signs, or worse at night. They should also be performed if the pain is associated with such symptoms as fever or weight

loss (red flags), which might indicate a more serious underlying pathology.

(b) Radiological changes are common in older people and may not indicate symptomatic osteoarthritis or spondylosis.

(c) X-rays are of little diagnostic value in early inflammatory arthritis but are useful as a baseline from which to judge later change.

- **Ultrasound (US)** is particularly useful for periarticular structures, soft tissue swellings and tendons. It is increasingly used to examine the shoulder and other structures during movement, e.g. shoulder impingement syndrome (see p. 507). US with Doppler measures blood flow and hence inflammation. US can be used to guide local injections. It is also used to assess bone density.
- **Magnetic resonance imaging (MRI)** shows bone changes and intra-articular structures in striking detail. Visualization of particular structures can be enhanced with different resonance sequences. T_1 weighted is used for anatomical detail, T_2 weighted for fluid detection and short tau inversion recovery (STIR) for the presence of bone marrow oedema. It is more sensitive than X-rays in the early detection of articular and peri-articular disease. It is the investigation of choice for most spinal disorders but is inappropriate in uncomplicated mechanical low back pain. Gadolinium injection enhances inflamed tissue. MRI can also detect muscle changes, e.g. myositis.
- **Computerized axial tomography (CT)** is useful for detecting changes in calcified structures.
- **Bone scintigraphy** utilizes radionuclides, usually ^{99m}Tc, and detects abnormal bone turnover and blood circulation and, although non-specific, helps in detecting areas of inflammation, infection or malignancy. It is best used in combination with other anatomical imaging techniques.
- **DXA scanning** uses very low doses of X-irradiation to measure bone density and is used in the screening and monitoring of osteoporosis.
- **Positron emission tomography (PET) scanning** uses radionuclides which decay by emission of positrons. ^{18}F-Fluorodeoxyglucose uptake indicates areas of increased glucose metabolism. It is used to locate tumours and demonstrate large vessel vasculitis, e.g. Takayasu's arteritis (see p. 809). PET scans are combined with CT to improve anatomical details.
- **Arthroscopy** is a direct means of visualizing a joint, particularly the knee or shoulder. Biopsies can be taken, surgery performed in certain conditions (e.g. repair or trimming of meniscal tears), and loose bodies removed.

COMMON REGIONAL MUSCULOSKELETAL PROBLEMS (Fig.10.2)

Analgesic and anti-inflammatory drugs used for treatment of musculoskeletal problems are discussed on page 517.

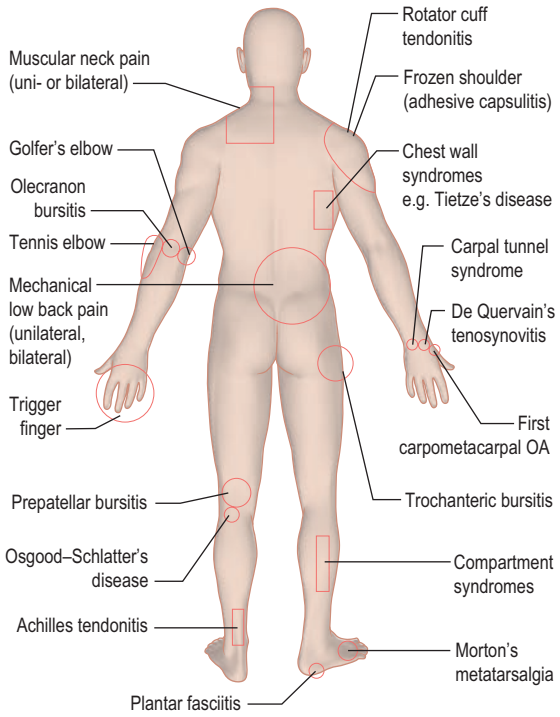


Fig. 10.2 **Common regional musculoskeletal problems.**

PAIN IN THE NECK AND SHOULDER (Table 10.3)

Mechanical or muscular neck pain (shoulder girdle pain)

Unilateral or bilateral muscular-pattern neck pain is common and usually self-limiting. It can follow injury, falling asleep in an awkward position, or prolonged keyboard working. Chronic burning neck pain occurs because of muscle tension from anxiety and stress.

Spondylosis (see p. 506) seen on X-ray increases after the age of 40 years, but it is not always causal. Spondylosis can, however, cause stiffness and increases the risk of mechanical or muscular neck pain. Muscle spasm can be palpable, is tender and may lead to abnormal neck posture

Table 10.3 Pain in the neck and shoulder

Trauma (for example, a fall)
Mechanical or muscular neck pain
Whiplash injury
Disc prolapse – nerve root entrapment (p. 510)
Ankylosing spondylitis
Shoulder lesions:
Rotator cuff tendonitis
Calcific tendonitis or bursitis
Impingement syndrome or rotator cuff tear
Adhesive capsulitis (true 'frozen' shoulder)
Inflammatory arthritis or osteoarthritis
Polymyalgia rheumatica
Fibromyalgia (chronic widespread pain)
Chronic (work-related) upper limb pain syndrome
Tumour

(e.g. acute torticollis). *Muscular-pattern neck pain* is not localized but affects the trapezius muscle, the C7 spinous process and the paracervical musculature (shoulder girdle pain). Pain often radiates upwards to the occiput but rarely laterally to the tip of the shoulder. It is commonly associated with unilateral or bilateral tension headaches; pain radiating over the head to the temple and eye, described as like a pressure or tight band. These features are also seen in chronic widespread pain (see p. 516).

Treatment

Patients are given short courses of analgesic therapy along with reassurance and explanation. Physiotherapists can help to relieve spasm and pain, teach exercises and relaxation techniques, and improve posture. An occupational therapist can advise about the ergonomics of the workplace if the problem is work-related (see p. 516).

Nerve root entrapment

This is caused by an acute cervical disc prolapse or pressure on the root from spondylotic osteophytes narrowing the root canal.

Acute cervical disc prolapse presents with unilateral pain in the neck, radiating to the interscapular and shoulder regions. This diffuse, aching dural pain is followed by sharp, electric shock-like pain down the arm, in a nerve root distribution, often with pins and needles, numbness, weakness and loss of reflexes (see Table 10.4).

Cervical spondylosis occurs in the older patient with postero-lateral osteophytes compressing the nerve root and causing root pain (see Fig. 21.46, p. 1176), commonly at C5/C6 or C6/C7; it is seen on oblique radiographs of the neck. An MRI scan shows facet joint OA, root canal narrowing and any associated disc prolapse clearly.

Treatment

A support collar, rest, analgesia and sedation are used as necessary. Patients should be advised not to carry heavy items. It usually recovers in 6–12 weeks. MRI is the investigation of choice if surgery is being considered or the diagnosis is uncertain (Fig. 10.3). A cervical root block administered under direct vision by an experienced pain specialist may relieve pain while the disc recovers. Neurosurgical referral is essential if the pain persists or if the neurological signs of weakness or numbness are severe or bilateral. Bilateral root pain is a neurosurgical emergency because a central disc prolapse may compress the cervical spinal cord. Posterior osteophytes may cause spinal claudication and cervical myelopathy (see p. 1177).

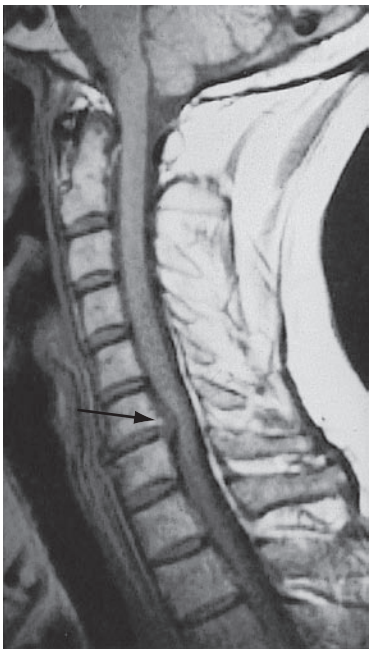


Fig. 10.3 **MRI of cervical spine**, showing a large central disc prolapse impinging on the spinal cord (arrow) at the C6/7 level.

Table 10.4 Cervical nerve root entrapment – symptoms and signs

Nerve root	Sensory changes	Reflex loss	Weakness
C5	Lateral arm	Biceps	Shoulder abduction Elbow flexion
C6	Lateral forearm	Biceps	Elbow flexion
	Thumb and index finger	Supinator	Wrist extension
C7	Middle finger	Triceps	Elbow extension
C8	Medial forearm	None	Finger flexion
	Little and ring fingers		
T1	Medial upper arm	None	Finger ab- and adduction

Whiplash injury

Whiplash injury results from acceleration–deceleration forces applied to the neck, usually in a road traffic accident when the car of a person wearing a seat belt is struck from behind. Delayed recovery depends in part upon the severity of the initial injury. A simple decision plan based on clinical criteria helps to distinguish those most at risk and who warrant radiography. There is a low probability of serious bony injury if there is no midline cervical tenderness, no focal neurological deficit, normal alertness, no intoxication and no other painful distracting injury. CT scans are reserved for those with bony injury. MRI scans occasionally show severe soft tissue injury. Whiplash injuries commonly lead to litigation.

Whiplash injury is a common cause of chronic neck pain, although most people recover within a few weeks or months. The pattern of chronic neck pain is often complex, involving pain in the neck, shoulder and arm. Headache, dizziness, and loss of memory and poor concentration sometimes accompany this. The subjective nature of these symptoms has led to controversy about their cause. The problem is more commonly seen in industrialized countries. It has often been suggested that the syndrome is caused in part by the prospect of financial compensation. This may not be wholly conscious and may be contributed to in part by the conflictive nature of the compensation process. There appears to be a direct relationship between poor prognosis and potential for compensation – elimination of compensation for pain and suffering due to this type of injury may lead to an improved prognosis.

Treatment is with reassurance (as the patient may be very anxious), analgesia, a short-term support collar and physiotherapy. Pain may take a few weeks or months to settle and the patient should be warned of this.

PAIN IN THE SHOULDER

The shoulder is a shallow joint with a large range of movement. The humeral head is held in place by the rotator cuff (Fig. 10.4) which is part of the joint capsule. It comprises the tendons of infraspinatus and teres minor posteriorly, supraspinatus superiorly and teres major and subscapularis anteriorly. The rotator cuff (particularly supraspinatus) prevents the humeral head blocking against the acromion during abduction; the deltoid pulls up and the supraspinatus pulls in to produce a turning moment. This permits the greater tuberosity to glide under the acromion without impingement.

Specific diagnoses are difficult to make clinically and are best made by *ultrasound investigation*. In the short term, however, this may not matter for pain management. Ultrasound should be reserved for patients whose pain persists.

Pain in the shoulder can sometimes be due to problems in the neck. The differential diagnosis of this is shown in Box 10.1. Although the term 'frozen shoulder' is commonly used for any painful stiff shoulder, true frozen shoulder (adhesive capsulitis) is uncommon – see below. A painful, stiff shoulder can result from rotator cuff lesions and is also seen following hemiplegia, chest or breast surgery or myocardial infarction. Painful shoulders may also be the initial presentation of RA, less commonly a seronegative spondyloarthropathy, and of polymyalgia rheumatica in the elderly.

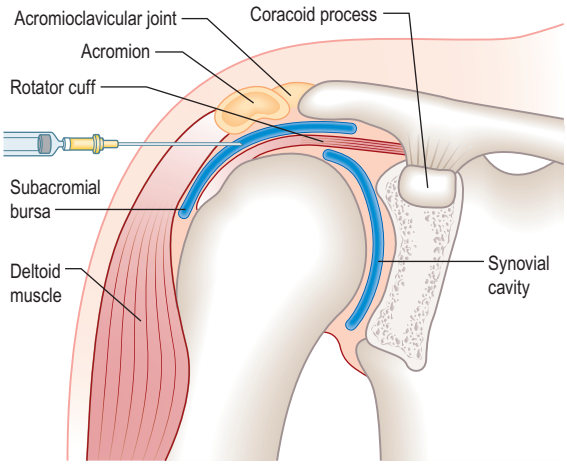


Fig. 10.4 **The shoulder region**, showing site of injection and subacromial space.

Box 10.1 Differential diagnosis of 'shoulder' pain

- Rotator cuff tendonitis pain is worse at night and radiates to the upper arm.
- Painful shoulders produce secondary muscular neck pain.
- Muscular neck pain (also known as shoulder girdle pain) does not radiate to the upper arm.
- Cervical nerve root pain is usually associated with pins and needles or neurological signs in the arm.

Management. Ultrasound examination can distinguish shoulder problems accurately and an ultrasound-guided corticosteroid injection can be given if required.

Rotator cuff (supraspinatus) tendonosis

This is a common cause of painful restriction of the shoulder at all ages. It follows trauma in 30% of cases and is bilateral in under 5%. The pain radiates to the upper arm and is made worse by arm abduction and elevation, which are often limited. The pain is often worse during the middle of the range of abduction, reducing as the arm is raised fully and the painful part of the tendon rotates through to the proximal side of the acromion – a so-called 'painful arc syndrome'. When examined from behind, the scapula rotates earlier than usual during elevation. Passive elevation reduces impingement and is less painful. Severe pain virtually immobilizes the joint, although some rotation is retained (cf. adhesive capsulitis, see below). There is also painful spasm of the trapezius. There may be an associated subacromial bursitis. Isolated subacromial bursitis occurs after direct trauma, falling on to the outstretched arm or elbow.

X-rays or ultrasound are necessary only when rotator cuff tendonosis is persistent or the diagnosis is uncertain.

Treatment

Analgesics or NSAIDs may suffice, but severe pain responds to an injection of corticosteroid into the subacromial bursa (Fig. 10.4). Patients should be warned that 10% will develop worse pain for 24–48 hours after injection. Seventy per cent improve over 5–20 days and mobilize the joint themselves. Physiotherapy helps persistent stiffness but therapeutic ultrasound is of unproven value. Further injections may be needed.

Torn rotator cuff

This is caused by trauma in the young but also occurs spontaneously in the elderly and in rheumatoid arthritis (RA). It prevents active abduction of the arm, but patients learn to initiate elevation using the unaffected arm. Once elevated, the arm can be held in place by the deltoid muscle. In younger people, the tear is repaired surgically but this is rarely possible in the elderly or in RA. Repeated trauma of the cuff between humerus and acromion/acromioclavicular joint causes osteophyte and cyst formation.

Shoulder impingement syndrome causes pain and crepitus on abduction and rotation. Some patients require arthroscopic surgery.

Calcific tendonosis and bursitis

Calcium pyrophosphate deposits in the tendon are visible on X-ray, but they are not always symptomatic. The pathogenesis is unclear, although ischaemia may play a part. The deposit is usually just proximal to the greater tuberosity. It may lead to acute or chronic recurrent shoulder pain and restriction of movement. A local corticosteroid injection may relieve the pain. The calcification may persist or resolve. Aspiration of the deposit under X-ray or ultrasound control may be required for persistent pain. Rarely, arthroscopic removal is necessary.

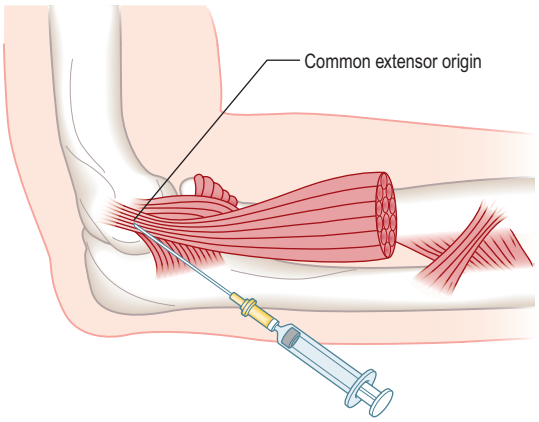


Fig. 10.5 **Injection for tennis elbow.**

Shedding of crystals into the subacromial bursa causes a bursitis with severe pain and shoulder restriction. The shoulder feels hot and is swollen, and an X-ray will show a diffuse opacity in the bursa. The differential diagnosis of calcific bursitis is gout, pseudogout or septic arthritis.

Aspiration and injection with corticosteroid can help.

Adhesive capsulitis (true 'frozen' shoulder)

This is uncommon. Severe shoulder pain is associated with complete loss of all shoulder movements, including rotation. High doses of NSAIDs and intra-articular injections of local anaesthetic and corticosteroids are helpful. Once the pain settles, arthroscopic release speeds functional recovery. When untreated, it recovers in 1–2 years.

PAIN IN THE ELBOW

Pain in the elbow can be due to epicondylitis, inflammatory arthritis or occasionally osteoarthritis.

Epicondylitis

Two common sites where the insertions of tendons into bone become inflamed (enthesitis) are the insertions of the wrist extensor tendon into the lateral epicondyle ('tennis elbow') and the wrist flexor tendon into the medial epicondyle ('golfer's elbow'). Both are usually unrelated to either sporting activity.

There is local tenderness. Pain radiates into the forearm on using the affected muscles – typically, gripping or holding a heavy bag in tennis elbow or carrying a tray in golfer's elbow. Pain at rest also occurs.

Treatment

Advise rest and arrange review by a physiotherapist. A local injection of corticosteroid at the point of maximum tenderness is helpful when the pain is severe (Fig. 10.5). Avoid the ulnar nerve when injecting golfer's elbow. Both conditions settle spontaneously eventually, but occasionally become disabling.

PAIN IN THE HAND AND WRIST (Table 10.5)

Hand pain is commonly caused by injury or repetitive work-related use. When associated with pins and needles or numbness it suggests a neurological cause arising at the wrist, elbow or neck. Pain and stiffness that are worse in the morning are due to tenosynovitis or inflammatory arthritis. The distribution of hand pain often indicates the diagnosis.

Table 10.5 Pain in the hand and wrist – causes

All ages	Older patients
Trauma/fractures	Nodal OA:
Tenosynovitis:	DIPs (Heberden's nodes)
Flexor with/without triggering	PIPs (Bouchard's nodes)
Dorsal	First carpometacarpal joint
De Quervain's	Trauma – scaphoid fracture
Carpal tunnel syndrome	Pseudogout
Ganglion	Gout:
Inflammatory arthritis	Acute
Raynaud's syndrome (p. 809)	Tophaceous
Chronic regional pain syndrome type I (p. 517)	

DIPs, PIPs, distal and proximal interphalangeal joints.

Tenosynovitis

The finger flexor tendons run through a series of synovial sheaths and under loops which hold them in place. Inflammation occurs with repeated or unaccustomed use, or in inflammatory arthritis when the thickened sheaths are often palpable.

Flexor tenosynovitis causes finger pain when gripping and stiffness of the fingers in the morning. Occasionally a tendon causes a trigger finger, when the finger remains flexed in the morning or after gripping and has to be pulled straight. A tendon nodule is palpable, usually in the distal palm.

Dorsal tenosynovitis is less common except in rheumatoid arthritis. The swelling is on the back of the hand and wrist.

De Quervain's tenosynovitis causes pain and swelling around the radial styloid where the abductor pollicis longus tendon is held in place by a retaining band. There is local tenderness, and the pain at the styloid is worsened by flexing the thumb into the palm (Finkelstein's sign).

Treatment

Resting, splinting and NSAIDs are used. Therapeutic ultrasound helps some people. Local corticosteroids are used and injected alongside the tendon under low pressure (not into the tendon itself). Occasionally surgery is needed if symptoms persist after six months.

Carpal tunnel syndrome

This is due to median nerve compression in the limited space of the carpal tunnel. Thickened ligaments, tendon sheaths or bone enlargement can cause it, but it is usually idiopathic. Causes are discussed on page 1172. The history is usually typical and diagnostic with the patient waking with numbness, tingling and pain in a median nerve distribution. The pain radiates to the forearm. The fingers feel swollen but usually are not. Wasting of the abductor pollicis brevis develops with sensory loss in the palm and radial three and a half fingers. Tinel's sign may be positive, i.e. reproduction of the pain on tapping the nerve in the carpal tunnel. Phalen's test is pain of the palm on maximal wrist flexion.

Treatment is with a splint to hold the wrist in dorsiflexion overnight. This relieves the symptoms and is diagnostic; used nightly for several weeks it may produce full recovery. If it does not, a corticosteroid injection into the carpal tunnel (N.B. not into the nerve!) helps in about 70% of cases, although it may recur. Persistent symptoms or nerve damage require nerve conduction studies showing prolonged latency across the carpal tunnel and surgical decompression of the carpal tunnel.

Other conditions causing pain

Inflammatory arthritis. This may present with pain, swelling and stiffness of the hands. In RA the wrists, proximal interphalangeal (PIP) joints and metacarpophalangeal (MCP) joints are affected symmetrically. In psoriatic arthritis and reactive arthritis a finger may be swollen (dactylitis) or the distal interphalangeal (DIP) joints and nails are affected asymmetrically.

Nodal osteoarthritis. This affects the DIP and less commonly PIP joints, which are initially swollen and red. The inflammation and pain settle but bony swellings remain. There is often a strong family history and it rarely presents before 50 years of age. Reassurance and local treatment are all that is needed.

First carpometacarpal osteoarthritis. This causes pain at the base of the thumb when gripping, or painless stiffness at the base of the thumb, often in persons with nodal osteoarthritis.

Scaphoid fractures. These cause pain in the anatomical snuffbox. They may not be seen immediately on X-ray. If untreated, scaphoid fractures can eventually cause pain because of failed union.

Ganglion. A ganglion is a jelly-filled, often painless swelling caused by a partial tear of the joint capsule. The wrist is a common site. Treatment is not essential as many resolve or cause little trouble. They rarely respond to injection, and surgical excision is the best option.

Dupuytren's contracture

This is a painless, palpable fibrosis of the palmar aponeurosis, with fibroblasts invading the dermis. It causes puckering of the skin and gradual flexion of the affected fingers, usually the ring and little fingers. It is more common in males, Caucasians, in diabetes mellitus and in those who abuse alcohol. A similar fibrosis occurs in the feet in a younger age group. It is often more aggressive. It is also associated with Peyronie's disease of the penis – a painful inflammatory disorder of the corpora cavernosa, leading eventually to painless fibrosis and angulation of the penis during erection. Intralesional steroid injections may help in early disease and some advocate transcutaneous needle aponeurotomy. Plastic surgical release of the contracture is restricted to those with severe deformity of the fingers.

PAIN IN THE LOWER BACK

Low back pain is a common symptom. It is often traumatic and work-related, although lifting apparatus and other mechanical devices and improved office seating are used to avoid it. Episodes are generally short-lived and self-limiting, and patients attend a physiotherapist or osteopath more often than a doctor. Chronic back pain is the cause of 14% of long-term disability in the UK. The causes are listed in Table 10.6, and the management of back pain is summarized in Box 10.2.

Table 10.6 Pain in the back (lumbar region) – causes

Mechanical
Trauma
Muscular and ligamentous pain
Fibrositic nodulosis
Postural back pain (sway back)
Lumbar spondylosis
Facet joint syndrome
Lumbar disc prolapse
Spinal and root canal stenosis
Spondylolisthesis
Disseminated idiopathic skeletal hyperostosis (DISH)
Fibromyalgia, chronic widespread pain (see p. 516)
Inflammatory
Infective lesions of the spine
Ankylosing spondylitis/sacroiliitis (see p. 532)
Metabolic
Osteoporotic spinal fractures (see p. 561)
Osteomalacia (see p. 567)
Paget's disease (see p. 565)
Neoplastic (see p. 555)
Metastases
Multiple myeloma
Primary tumours of bone
Referred pain

Box 10.2 Management of back pain

- Most back pain presenting to a primary care physician needs no investigation.
- Pain between the ages of 20 and 55 years is likely to be mechanical and is managed with analgesia, brief rest if necessary and physiotherapy.
- Patients should stay active within the limits of their pain.
- Early treatment of the acute episode, advice and exercise programmes reduce long-term problems and prevent chronic pain syndromes.
- Physical manipulation of uncomplicated back pain produces short-term relief and enjoys high patient satisfaction ratings.
- Psychological and social factors may influence the time of presentation.
- Appropriate early management reduces long-term disability.

Investigations

- **Spinal X-rays** are required only if the pain is associated with certain 'red flag' symptoms or signs, which indicate a high risk of more serious underlying problems:
  (a) starts before the age of 20 or after 50 years
  (b) is persistent and a serious cause is suspected
  (c) is worse at night or in the morning, when an inflammatory arthritis (e.g. ankylosing spondylitis), infection or a spinal tumour may be the cause

(d) is associated with a systemic illness, fever or weight loss

(e) is associated with neurological symptoms or signs.

- **MRI** is preferable to CT scanning when neurological signs and symptoms are present. CT scans demonstrate bony pathology better.
- **Bone scans** are useful in infective and malignant lesions but are also positive in degenerative lesions.
- **Full blood count, ESR and biochemical tests** are required only when the pain is likely to be due to malignancy, infection or a metabolic cause.

Mechanical low back pain

Mechanical low back pain starts suddenly, may be recurrent and is helped by rest. Mechanical back pain is often precipitated by an injury and may be unilateral or bilateral. It is usually short-lived.

Examination and management

The back is stiff and a scoliosis may be present when the patient is standing. Muscular spasm is visible and palpable and causes local pain and tenderness. It lessens when sitting or lying. Patients with spondylosis on X-ray probably have an increased risk of developing mechanical back pain but changes are often absent in the young and may be coincidental in the elderly. Pain relief and physiotherapy are helpful. Excessive rest should be avoided. The patient needs to be re-educated in lifting and shown exercises to prevent recurrent attacks of pain. Once a patient has presented to a general practitioner with low back pain, although the episode itself is usually self-limiting, there is a significantly increased risk of further back pain episodes. Risk factors for recurrent back pain include female sex, increasing age, pre-existing chronic widespread pain (fibromyalgia) and such psychosocial factors as high levels of psychological distress, poor self-rated health and dissatisfaction with employment. Chronic low back pain is a major cause of disability and time off work and is reduced by appropriate early management.

Spinal movement occurs at the disc and the posterior facet joints, and stability is normally achieved by a complex mechanism of spinal ligaments and muscles. Any of these structures may be a source of pain. An exact anatomical diagnosis is difficult, but some typical syndromes are recognized (see below). They are often associated with but not necessarily caused by radiological spondylosis (see p. 1176).

Postural back pain develops in individuals who sit in poorly designed, unsupportive chairs.

Lumbar spondylosis. The fundamental lesion in spondylosis occurs in an intervertebral disc, a fibrous joint whose tough capsule inserts into the rim of the adjacent vertebrae. This capsule encloses a fibrous outer zone and a gel-like inner zone. The disc allows rotation and bending.

Changes in the discs may start in teenage years or early twenties and increase with age. The gel changes chemically, breaks up, shrinks and loses its compliance. The surrounding fibrous zones develop circumferential or radial fissures. In the majority this is initially asymptomatic but visible on MRI as decreased hydration. Later the discs become thinner and less compliant. These changes cause circumferential bulging of the intervertebral ligaments.

Reactive changes develop in adjacent vertebrae; the bone becomes sclerotic and osteophytes form around the rim of the vertebra (Fig. 10.6). The most common sites of lumbar spondylosis are L5/S1 and L4/L5. In young people, disc prolapse through an adjacent vertebral endplate produces a Schmorl's node on X-ray. This is painless but may accelerate disc degeneration.

Spondylosis may be symptomless, but it can cause:

- episodic mechanical spinal pain
- progressive spinal stiffening
- facet joint pain
- acute disc prolapse, with or without nerve root irritation
- spinal stenosis
- spondylolisthesis.

Facet joint syndrome. Lumbar spondylosis also causes secondary osteoarthritis of the misaligned facet joints. Pain is typically worse on bending backwards and when straightening from flexion. It is lumbar in site, unilateral or bilateral and radiates to the buttock. The facet joints are well seen on MRI and may show osteoarthritis, an effusion or a ganglion cyst. Diagnostic local anaesthetic injections into the joints (under X-ray vision) can be followed by a corticosteroid injection. The long-term value of this is unclear but many patients find the procedure helpful. Physiotherapy to reduce hyperlordosis and reducing weight are helpful.

Fibrositic nodulosis. This causes unilateral or bilateral low back pain, radiating to the buttock and upper posterior thigh. There are tender nodules in the upper buttock and along the iliac crest. Such nodules are relevant only if they are tender and associated with pain. They are probably traumatic. Local, intralesional corticosteroid injections help.

Postural back pain and sway back of pregnancy. Low back pain is common in pregnancy and reflects altered spinal posture and increased ligamentous laxity. There is usually a hyperlordosis on examining the patient standing. Weight control and pre- and postnatal exercises are helpful, and the pain usually settles after delivery. Analgesics and NSAIDs are best avoided during pregnancy and breast-feeding. Epidurals during delivery are not associated with an increased incidence of subsequent back pain. Poor posture causes a similar syndrome in the non-pregnant, owing to obesity or muscular weakness. Poor sitting posture at work is a frequent cause of chronic low back pain.

Treatment of mechanical back pain

Adequate analgesia to allow normal mobility and avoid bed rest is best, combined with physical treatments. The evidence base for physiotherapy, back muscle training regimens and manipulation is not good. Manipulation produces more rapid pain relief in some patients. Acupuncture may help. A positive approach probably reduces the development of chronic pain. A comfortable sleeping position should be adopted using a mattress of medium (not hard) firmness.

Acute lumbar disc prolapse

The central disc gel may extrude into a fissure in the surrounding fibrous zone and cause acute pain and muscle spasm, which in turn leads to a forwards and sideways tilt when standing. These events are often self-limiting. A disc prolapse occurs when the extrusion extends beyond the

Ask the authors
www.studentconsult.com

limits of the fibrous zone (Fig. 10.6). The weakest point is posterolateral, where the disc may impinge on emerging spinal nerve roots in the root canal.

The episode starts dramatically during lifting, twisting or bending and produces a typical combination of low back pain and muscle spasm, and severe, lancinating pains, paraesthesia, numbness and neurological signs in one leg (rarely both). The back pain is diffuse, usually unilateral and radiates into the buttock. The muscle spasm leads to a scoliosis that reduces when lying down. The nerve root pain develops with, or soon after, the onset. The site of the pain and other symptoms is determined by the root affected (Table 10.7). A central high lumbar disc prolapse may cause spinal cord compression and long tract signs (i.e. upper motor neurone). Below L2/L3 it produces lower motor neurone lesions.

On examination, the back often shows a marked scoliosis and muscle spasm. The straight-leg-raising test, whilst lying, is positive in a lower lumbar disc prolapse – raising the straight leg beyond 30 degrees produces pain in the leg. Slight limitation or pain in the back limiting this movement is seen with mechanical back pain. Pain in the affected leg produced by a straight raise of the other leg suggests a large or central disc prolapse. Look for perianal sensory loss and urinary retention, which indicate a cauda equina lesion – a neurosurgical emergency. An upper lumbar disc prolapse produces a positive femoral stretch test – pain in the anterior thigh when the knee is flexed in the prone position.

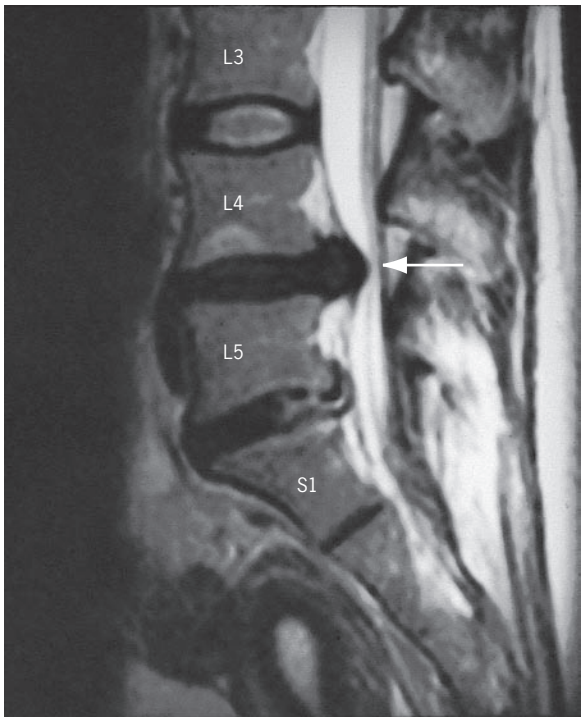


Fig. 10.6 **MRI of lumbar spine**, showing a central disc prolapse at the L4/L5 level (arrow). The signal from the L4/L5 and L5/S1 discs indicates dehydration, while the L3/L4 signal appearance is normal.

Treatment

Advise a short period (2–3 days) of bed rest – lying flat for a lower disc but semi-reclining for a high lumbar disc – and prescribe analgesia and muscle relaxants. Once the pain is tolerable, encourage the patient to mobilize and refer to a physiotherapist for exercises and preventative advice. An X-ray-guided epidural or nerve root canal injection by a pain specialist reduces pain rapidly, although the evidence that it speeds resolution or prevents surgery is unclear. Caudal epidural injections are less effective than lumbar ones. Resuscitation equipment must be available for these procedures. Referral to a surgeon for possible microdiscectomy or hemilaminectomy is necessary if the neurological signs are severe, if the pain persists and is severe for more than 6–10 weeks, or if the disc is central. If bladder or anal sphincter tone is affected it becomes a neurosurgical emergency.

Spinal and root canal stenosis

Progressive loss of disc height, OA of the facet joints, posterolateral osteophytes and buckling of the ligamentum flavum all contribute to root canal stenosis. This causes nerve root pain or spinal root claudication – pain and paraesthesiae in a root distribution brought on by walking and relieved slowly by rest. The associated sensory symptoms, slow recovery when the patient rests, and presence of normal foot pulses distinguish this from peripheral arterial claudication. Severe cervical spondylosis may also produce spinal claudication, often with arm symptoms and signs.

Spinal canal stenosis at more than one level is often associated with severe spondylosis and a congenitally narrow spinal canal. It causes buttock and bilateral leg pain, paraesthesiae and numbness when walking. Rest helps, as does bending forwards, a manoeuvre that opens the spinal canal. Specialist surgical advice is necessary.

Spondylolisthesis

This occurs in adolescents and young adults when bilateral congenital pars interarticularis defects cause instability and permit the vertebra to slip, with or without preceding injury.

Table 10.7 **Lumbar nerve root entrapment – symptoms and signs**

Nerve root	Sensory changes	Reflex loss	Weakness	Usual disc prolapse
L2	Front of thigh	None	Hip flexion/adduction	L2/3
L3	Inner thigh and knee	Knee	Knee extension	
L4	Inner calf	Knee	Knee extension	L3/4
L5	Outer calf	None	Inversion of foot	L4/5
	Upper, inner foot		Dorsiflexion of toes	
S1	Posterior calf	Ankle	Plantar flexion of foot	L5/S1
	Lateral border of foot			

Rarely a cauda equina syndrome with loss of bladder and anal sphincter control and saddle-distribution anaesthesia develops (p. 1177). It is diagnosed radiologically. Low back pain in adolescents warrants investigation, and spondylolisthesis requires orthopaedic assessment. It needs careful monitoring during the growth spurt.

A degenerative spondylolisthesis may also develop in older people with lumbar spondylosis and osteoarthritis of the facet joints.

Diffuse idiopathic skeletal hyperostosis (DISH)

DISH (Forrestier's disease) affects the spine and extraspinal locations. It causes bony overgrowths and ligamentous ossification and is characterized by flowing calcification over the anterolateral aspects of the vertebrae. The spine is stiff but not always painful, despite the dramatic X-ray changes. Ossification at muscle insertions around the pelvis produces radiological 'whiskering'. Similar changes occur at the patella and in the feet.

Treatment is with analgesics or NSAIDs for pain, and exercise to retain movement and muscle strength.

Osteoporotic crush fracture of the spine

Osteoporosis is asymptomatic but leads to an increased risk of fracture of peripheral bones, particularly neck of femur and wrist, and thoracic or lumbar vertebral crush fractures. Such vertebral fractures develop without trauma, after minimal trauma, or as part of a major accident. They may develop painlessly or cause agonizing localized pain that radiates around the ribs and abdomen. Multiple fractures lead to an increased thoracic kyphosis ('widow's stoop'). The diagnosis is confirmed by X-rays, showing loss of anterior vertebral body height and wedging, with sparing of the vertebral endplates and pedicles. Bone oedema on MRI indicates that a fracture is recent.

Treatment

Advise bed rest and analgesia until the severe pain subsides over a few weeks, then gradual mobilization. It may warrant hospitalization, and intravenous bisphosphonates or subcutaneous calcitonin are given to relieve pain. There may be some residual pain and deformity. Bone density measurement and preventative treatment of osteoporosis are essential (see p. 563).

Septic discitis

Septic discitis may cause severe pain and rapid adjacent vertebral destruction. It requires urgent neurosurgical referral.

Ankylosing spondylitis (see also p. 532)

Buttock pain and low back stiffness in a young adult suggests ankylosing spondylitis, especially if it is worse at night and in the morning.

FURTHER READING

Koes BW. Evidence-based management of acute low back pain. *Lancet* 2007; **370:** 1595–1596.

PAIN IN THE HIP (Table 10.8)

'Hip' refers to a wide area between the upper buttock, trochanter and groin. It is useful to ask the patient to point to the site of pain and its field of radiation. Pain arising from the hip joint itself is felt in the groin, lower buttock and anterior thigh, and may radiate to the knee. Occasionally and inexplicably, hip arthritis causes pain only in the knee.

Table 10.8 Pain in the hip – causes

Hip region problems	Main sites of pain
Osteoarthritis of hip	Groin, buttock, front of thigh to knee
Trochanteric bursitis (or gluteus medius tendonopathy)	Lateral thigh to knee
Meralgia paraesthetica	Anterolateral thigh to knee
Referred from back	Buttock
Facet joint pain	Buttock and posterior thigh
Fracture of neck of femur	Groin and buttock
Inflammatory arthritis	Groin, buttock, front of thigh to knee
Sacroiliitis (AS)	Buttock(s)
Avascular necrosis	Groin and buttocks
Polymyalgia rheumatica	Lumbar spine, buttocks and thighs

AS, ankylosing spondylitis.

Osteoarthritis (OA) (see also p. 518)

OA is the most common cause of hip joint pain in a person over the age of 50 years. It causes pain in the buttock and groin on standing and walking. Stiff hip movements cause difficulty in putting on a sock and may produce a limp.

Trochanteric bursitis and gluteus medius tendonopathy

This may be due to trauma or unaccustomed exercise, but sometimes has an unknown cause. It also occurs in inflammatory arthritis. The pain over the trochanter is worse going up stairs and when abducting the hip, and the trochanter is tender to lie on. A local corticosteroid injection onto the surface of the trochanter is helpful. A tear of the gluteus medius tendon at its insertion into the trochanter causes a similar syndrome, but does not respond to injection. MRI scans have demonstrated this new syndrome. Its best management is unclear.

Meralgia paraesthetica (see also p. 1172)

This causes numbness and burning dysaesthesia (increased sensitivity to light touch) over the anterolateral thigh and may be precipitated by a sudden increase in weight, an injury or during pelvic surgery. It is usually self-limiting but can be helped by amitriptyline 10 mg at night or gabapentin starting with 300 mg.

Fracture of the femoral neck

This usually occurs after a fall, occasionally spontaneously. There is pain in the groin and thigh, weight-bearing is painful or impossible, and the leg is shortened and externally rotated. Occasionally a fracture is not displaced and remains undetected. X-rays are diagnostic. Anyone with a hip fracture, especially after minimal trauma, should be reviewed for osteoporosis (see p. 562).

Table 10.9 Pain in the knee

- Trauma and overuse
  - Periarticular problems
    - Anterior knee pain or medial knee pain
  - Internal derangements – meniscal tears or cruciate ligament tears
- Osteoarthritis
- Inflammatory arthritis
  - Acute monoarthritis
    - Gout, pseudogout, Reiter's disease or septic arthritis
  - Pauciarticular (<4 joints)
    - Seronegative spondyloarthropathy or atypical rheumatoid arthritis
  - Polyarticular
    - Rheumatoid arthritis
- Popliteal (Baker's) cyst/ruptured cyst
- Osteochondritis dissecans
- Hypermobility syndrome
- Referred from hip joint

Avascular necrosis (osteonecrosis) of the femoral head

This is uncommon but occurs at any age. There is severe hip pain. X-rays are diagnostic after a few weeks, when a well-demarcated area of increased bone density is visible. In the femur this lies at the upper pole of the femoral head. The affected bone may collapse. Early, the X-ray is normal but bone scintigraphy or MRI demonstrates the lesion. Risk factors include treatment with corticosteroids or heparin, exposure to high barometric pressures (divers and tunnellers), excessive alcohol consumption, and sickle cell disease.

Inflammatory arthritis of the hip

This produces pain in the groin and stiffness, which are worse in the morning. Rheumatoid arthritis (RA) rarely presents with hip pain, although the hip is involved eventually in severe RA. Ankylosing spondylitis and other seronegative spondyloarthropathies cause inflammatory hip arthritis in younger people.

Polymyalgia rheumatica (see also p. 549)

Bilateral hip, buttock and thigh pain and stiffness that are worse in the morning in an elderly patient may be attributable to polymyalgia rheumatica. Neck and shoulder pain and stiffness are usually also present.

PAIN IN THE KNEE (Table 10.9)

The knee depends on ligaments and quadriceps muscle strength for stability. It is frequently injured, particularly during sports. Trauma or overuse of the knee leads to a variety of peri- and intra-articular problems. Some are self-limiting; others require physiotherapy, local corticosteroid injections or surgery.

The knee is also a common site of inflammatory arthritis and osteoarthritis. Minor radiographic changes of osteoarthritis (see Fig. 10.9, p. 519) are common in the over-fifties and often coincidental, the cause of the pain being periarticular. Knee pain should not be attributed to osteoarthritis until other causes have been excluded. Symptomatic osteoarthritis of the knee correlates poorly with the severity of the radiological changes.

Common periarticular knee lesions

Medial knee pain

There may be medial or lateral ligament strain, but the medial ligament is more commonly affected. There is pain at the ligament's insertion into the upper medial tibia, which is worsened by standing or stressing the affected ligament.

Anserine bursitis causes pain and localized tenderness 2–3 cm below the posteromedial joint line in the upper part of the tibia at the site of the bursa. It occurs in obese women, often with valgus deformities, and in breast-stroke swimmers.

Treatment is with physiotherapy and a local corticosteroid injection.

Anterior knee pain

Anterior knee pain is common in adolescence. In many cases no specific cause is found despite careful investigation. This is called 'anterior knee pain syndrome' and settles with time. Isometric quadriceps exercises and avoidance of high heels both help the condition. Patient and parents often need firm reassurance. Abnormal patellar tracking may be a cause and need surgical treatment. Hypermobility of joints causes joint pain, maltracking and rarely recurrent patellar dislocation (see also p. 554).

Pre- and infrapatellar bursitis are caused by unaccustomed kneeling ('housemaid's knee'). There is local pain, tenderness and fluctuant swelling. Avoidance of kneeling and a local corticosteroid injection are helpful. Septic bursitis can occur.

Chondromalacia patellae is diagnosed arthroscopically. The retropatellar cartilage is fibrillated. In most cases the pain settles eventually. When there is patellar misalignment it may need surgery, as does recurrent patellar dislocation in adolescent girls.

Osgood–Schlatter disease (p. 554) causes pain and swelling over the tibial tubercle. It is a traction apophysitis of the patellar tendon and occurs in enthusiastic teenage sports players.

Enthesitis may occur at the patellar end of the tendon (jumper's knee).

Common intra-articular traumatic lesions of the knee

Torn meniscus

The menisci are partially attached fibrocartilages that stabilize the rounded femoral condyles on the flat tibial plateaux. In the young they are resilient but this decreases with age. They can be torn by a twisting injury, commonly in sports that involve twisting and bending. The history is usually diagnostic. There is immediate medial or lateral knee pain and dramatic swelling within a few hours. The affected side is tender. If the tear is large the knee may lock flexed. The immediate treatment is to apply ice. MRI demonstrates the tear (Fig. 10.7). In most circumstances, especially in active sportsmen, early arthroscopic repair or trimming of the torn meniscus is essential. Surgical intervention reduces recurrent pain, swelling and locking but not the risk of secondary osteoarthritis. The long-term benefit of repairing tears is not

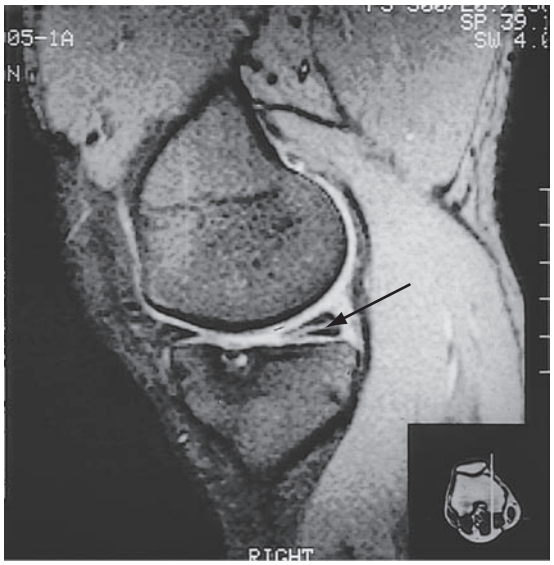


Fig. 10.7 **MRI of a knee**, showing a complete tear of the posterior horn of the medial meniscus, extending to its lower surface (arrow).

yet known. Post-surgical quadriceps exercises aid a return to sport and other activities.

Torn cruciate ligaments

Torn cruciate ligaments account for around 70% of knee haemarthroses in young people. They often coexist with a meniscal tear. Partial cruciate tears are difficult to diagnose clinically. On flexing the knee to 90 degrees, a torn anterior cruciate allows the tibia to be pulled forwards on the femur. MRI is the investigation of choice. Such injuries need urgent orthopaedic referral. There is a significant incidence of secondary OA.

Osteochondritis dissecans

This occasionally causes knee pain and swelling in adolescents and young adults, more commonly males. It is probably traumatic, possibly with hereditary predisposing factors. A fragment of bone and its attached cartilage detach by shearing, most commonly from the lateral aspect of the medial femoral condyle.

There is aching pain after activity and, if the fragment becomes loose, locking or 'giving way' occurs. The lesion is seen on a tunnel-view X-ray, but MRI is more sensitive, especially if the fragment is undisplaced. Undisplaced lesions are treated with rest, then isometric quadriceps exercises. Loose fragments can be fixed arthroscopically or removed. A similar lesion affecting the lateral femoral condyle occurs in older people.

Knee joint effusions

An effusion of the knee causes swelling, stiffness and pain. The pain is more severe with an acute onset and with increasing inflammation, because of stretching of the capsule that contains the pain receptors. A full clinical history and examination must include a past medical, family and drug history.

Inflammatory arthritis affects the knees and causes warmth and swelling. An acute inflammatory monoarthritis of the knee is a common presentation of a seronegative spondyloarthropathy and occasionally is the first sign of RA.

Monoarthritis of the knee, associated with severe pain and marked redness, may be due to septic arthritis, or gout in the middle-aged male, or to gout or pseudogout in an older male or female. A cool, clear, viscous effusion is seen in elderly patients with moderate or severe symptomatic OA (see p. 518).

Examination

A large and tense effusion is easily seen and felt on each side of the patella and in the suprapatellar pouch, and is fluctuant. The effusion delays the patella tapping against the femur when it is pressed firmly and quickly (the 'patellar tap' sign) with the knee held straight and relaxed. Small effusions also demonstrate the 'bulge' sign when the patient is lying with the quadriceps relaxed. For this, apply a gentle sweeping pressure, first to the medial side of the joint and then, watching the medial dimple, to the lateral side. Slightly delayed bulging of the medial dimple indicates fluid in the joint.

Investigations

These are (a) blood tests, and (b) aspiration and examination of the knee effusion. The basic technique of aspiration is described in Practical box 10.2 on page 504.

If there is a flexion deformity, externally rotate the leg or support the knee. Stand on the side opposite the joint and insert the needle between the patella, just proximal to its midpoint, and the medial femoral condyle. Angle the needle slightly backwards and inject small volumes of local anaesthetic, advancing until aspiration detects fluid. Then change the syringe for a larger one and aspirate as much fluid as possible. Examine the fluid and decide whether to inject corticosteroid or arrange microbiological tests.

Haemarthrosis of the knee

This is caused by:

- trauma – meniscal, cruciate or synovial lining tear
- clotting or bleeding disorders, such as haemophilia, sickle cell disease or von Willebrand's disease.

Popliteal cyst (Baker's cyst). In approximately 5% of patients with a knee effusion, a swollen, painful popliteal cyst develops. This is usually due to a bursa (usually the semimembranosus bursa), which in some individuals has a valve-like connection to the knee. This allows the effusion to flow into the bursa but not back. Occasionally there is a synovial herniation through the posterior joint capsule. The cyst is best seen and felt in the popliteal fossa with the patient standing.

Ruptured popliteal cyst. A popliteal cyst may rupture if the patient is mobile, particularly on standing up quickly or climbing stairs. Fluid escapes into the soft tissue of the popliteal fossa and upper calf, causing sudden and severe pain, swelling and tenderness of the upper calf. Dependent oedema of the ankle develops and the knee effusion reduces dramatically in size and may be undetectable.

A history of previous knee problems and the sudden onset of pain and tenderness high in the calf suggest a ruptured cyst rather than a deep vein thrombosis (DVT). However, the diagnosis is often missed and treated inappropriately

with anticoagulants. A diagnostic ultrasound examination distinguishes a ruptured cyst from a DVT (see p. 809). Analgesics or NSAIDs, rest with the leg elevated, and aspiration and injection with corticosteroids into the knee joint are required.

PAIN IN THE FOOT AND HEEL (Table 10.10)

The feet are subjected to extreme pressures by weight-bearing and inappropriate shoes. They are commonly painful. Broad, deep, thick-soled shoes are essential for sporting activities, prolonged walking or standing, and in people with congenitally flat or arthritic feet.

There are two common types of foot deformity:

- *Flat feet* – stress the ankle and throw the hindfoot into a valgus (everted) position. A flat foot is rigid and inflexible
- *High-arched feet* – place pressure on the lateral border and ball of the foot.

The foot is affected by a variety of inflammatory arthritic conditions. After the hand, the foot joints are the most commonly affected by rheumatoid arthritis. The diagnosis depends upon careful assessment of the distribution of the joints affected, the pattern of other joint problems, or by finding the associated condition (e.g. psoriasis, see p. 1240).

Hallux valgus

The great toe migrates laterally. In the congenital form the first metatarsal bone is displaced medially (metatarsus primus varus). The shape of modern shoes causes later onset of hallux valgus. It is a common complication of RA.

Hallux rigidus

Osteoarthritis of the first metatarsophalangeal (MTP) joint in a normally aligned or valgus joint causes hallux rigidus – a stiff, dorsiflexed and painful great toe. Careful choice of footwear and the help of a podiatrist suffice for most cases, but some require surgery.

Metatarsalgia

This is common, especially in women who wear high heels, after trauma and in those with hammer toes. The ball of the foot is painful to walk and stand on. Callosities and pressure-induced bursae develop under the metatarsal heads. Rheumatoid arthritis causes misalignment of the metatarsal bones and severe metatarsalgia.

Treatment is with podiatry and the wearing of appropriate shoes. Surgery is occasionally needed, particularly in the rheumatoid forefoot.

Morton's metatarsalgia is due to a neuroma, usually between the third and fourth metatarsal heads. It causes pain, burning and numbness in the adjacent surfaces of the affected toes when walking. It is helped by wearing wider, cushioned-soled shoes. Occasionally a steroid injection is necessary.

Table 10.10 Pain in the foot and heel – causes

Structural (flat (pronated) or high arched (supinated))	
Hallux valgus/rigidus (±OA)	
Metatarsalgia	
Morton's neuroma	
Stress fracture	
Inflammatory arthritis	
Acute, monoarticular – gout	
Chronic, polyarticular – RA	
Chronic, pauciarticular – seronegative spondyloarthropathy	
Tarsal tunnel syndrome	
Heel pain	
Plantar fasciitis	Below heel
Plantar spur	Below heel
Achilles tendonitis/bursitis	Behind heel
Sever's disease	Behind heel
Arthritis of ankle/subtaloid joints	

Stress (march) fractures

These cause sudden, severe, weight-bearing pain in the distal shaft of the fractured metatarsal bone. They occur after unaccustomed walking or with new shoes. There is local tenderness and swelling, but initially X-rays are normal and diagnosis delayed. A radioisotope bone scan or MRI reveals the fracture earlier than X-rays. Reduced weight-bearing for a few weeks usually suffices.

Tarsal tunnel syndrome

This is an entrapment neuropathy of the posterior tibial nerve as it rounds the medial malleolus. It produces burning, tingling and numbness of the toes, sole and medial arch. The nerve is tender below the malleolus and, when tapped, produces a shock-like pain (Tinel's sign). A local steroid injection under the retinaculum, between the medial malleolus and calcaneum, is helpful.

Pain under the heel

Plantar fasciitis is an enthesitis at the insertion of the tendon into the calcaneum. It produces localized pain when standing and walking, and tenderness in the midline. It occurs alone or in seronegative spondyloarthropathy.

Plantar spurs are traction lesions at the insertion of the plantar fascia in older people and are usually asymptomatic. They become painful after trauma.

Calcaneal bursitis is a pressure-induced (adventitious) bursa that produces diffuse pain and tenderness under the heel. Compression of the heel pad from the sides is painful, which distinguishes it from plantar fascia pain.

Whatever the cause, the pain is always worse in the morning as soon as weight is placed on the foot.

All of these lesions are treated with heel pads, and reduced walking; they are often self-limiting. A splint at night to hold the foot dorsiflexed and to stretch the plantar fascia is preferable to a local corticosteroid injection in plantar fasciitis. When an injection is necessary, a medial approach is used, rather than through the heel pad, often under ultrasound guidance.

Pain behind the heel and leg

Sever's disease is a traction apophysitis of the Achilles tendon in young people (cf. Osgood–Schlatter disease, p. 554).

Pain at the insertion of the Achilles tendon into the calcaneum is an enthesitis. This is traumatic or it can complicate seronegative spondyloarthropathy. Tendon damage or ruptive can occur with quinolone e.g. ciprofloxacin therapy. Raising the shoe heel reduces pain. Occasionally a low-pressure corticosteroid injection near the enthesis is necessary.

Achilles tendonosis causes a painful, tender swelling a few centimetres above the tendon's insertion. Advise against walking barefoot and jumping. Therapeutic ultrasound is helpful. (Caution – a local injection may cause the tendon to rupture.)

Achilles' bursitis lies clearly anterior to the tendon and can be safely injected with corticosteroid.

Compartment syndromes

The muscles of the lower leg are enclosed in fascial compartments, with little room for expansion to occur. Compartment syndromes can be acute and severe, such as following exercise.

In the *anterior tibial syndrome* there is severe pain in the front of the shin, occasionally with foot drop. Immediate surgical decompression to prevent muscle necrosis is sometimes required.

Chronic compartment syndrome produces pain in the lower leg that is aggravated by exercise and may therefore be mistaken for a vascular or neurological disorder.

PAIN IN THE CHEST

Musculoskeletal conditions are sometimes a cause of chest pain. An example is Tietze's disease. In this condition, pain arises from the costosternal junctions. It is usually unilateral and affects one, two or three ribs. There is local tenderness, which helps to make the diagnosis. The condition is benign and self-limiting. It often responds well to anti-inflammatory drugs. Other causes of chest wall pain include rib fractures due to trauma or osteoporosis or a malignant deposit. Costochondral pain occurs in ankylosing spondylitis (see p. 532). In patients with heart disease this may cause severe anxiety but this pain is not like angina and the patient should be reassured.

CHRONIC PAIN SYNDROMES

(see also p. 1193)

Chronic pain syndromes are difficult to manage. Psychological factors are at least as relevant as inflammation or damage in determining a patient's perception of pain. It is essential to be objective and non-judgmental when discussing physical, psychological and social factors without assuming which is primary. Chronic pain syndromes are difficult to explain scientifically and it is all too easy for a doctor to 'blame' the patient for this lack of explanation. Some chronic pain states may be exacerbated partly by the process of litigation that may follow an injury.

Any chronic painful condition can change the way a person copes. Some people with chronic diseases or chronic pain cope well, but others adopt coping strategies and patterns of behaviour which make things worse. They become anxious, depressed or socially isolated, and their quality of life is reduced. In chronic pain syndromes patients need help to lead a more normal life despite their pain, and are best referred to a specialist, multidisciplinary pain service.

Psychological states such as depression and anxiety produce physical symptoms, of which one is pain, while people with frank physical diseases are often understandably anxious and depressed. A biopsychosocial approach is best.

Chronic widespread pain (previously fibromyalgia)

Chronic widespread pain is defined as pain for more than three months both above and below the waist (p. 1193). It is a diagnosis of exclusion although it is still not universally accepted as a diagnosis. Patients value a name to explain symptoms previously dismissed or attributed simply to psychological or social problems. A feature for diagnosing fibromyalgia is tender trigger points although this is not specific. The patient is usually a middle-aged woman who struggles on with her work and/or housework despite the pain. Such individuals are difficult to live with and there is often family discord. Many patients have sleep disturbances, so they awake unrefreshed and have poor concentration. It can occur at any age and is increasingly recognized in teenagers.

The pain is a widespread, unremitting, aching discomfort. There are often other health problems, such as chronic fatigue syndrome (see below), irritable bowel syndrome, premenstrual syndrome, tension headache, anxiety and depression; doctors sometimes inappropriately label them 'heart sink' patients. The patient's frustration is compounded by the fact that most tests are normal, and they fear doctors believe it is 'all in their mind'. This is best regarded as a 'wind up' or pain amplification syndrome and is attributable to changes in the descending inhibitory pathways and in the spinal cord.

Treatment

A sympathetic approach is appropriate, with reassurance for the patient that their chronic widespread pain often improves and is not inevitably disabling. A graded aerobic exercise regimen over 3 months is safe and effective. When depression is present, it should be treated, but potentially addictive anxiolytic agents are best avoided. A behavioural psychologist may persuade the person to pace their life more effectively and to cope better, although patients often resist referral for psychological help.

Drugs

Analgesics or NSAIDs help in some cases but are best used intermittently. Gabapentin and pregabalin also help although drowsiness may be a problem.

Low doses of sedative antidepressant drugs, such as amitriptyline or dosulepin, help when taken a few hours before bedtime. They act by increasing the levels of serotonin (5HT) in the CNS and probably increasing descending sensory inhibition. It should be explained that these doses are analgesic and not antidepressant, and their side-effects should be outlined. Trigger-point injections with local anaesthetic, corticosteroids or acupuncture are sometimes helpful.

Oral corticosteroids are not helpful.

Chronic fatigue syndrome

Diffuse muscular pain and stiffness is common in this condition, which is described on page 1192.

Chronic (work-related) upper-limb pain syndrome

This name is preferred to 'repetitive strain injury' (RSI). The predominant symptoms are pain in all or part of one or both

arms. A specific lesion, such as tennis elbow or carpal tunnel syndrome, or muscular-pattern neck pain often develops first, and early recognition and treatment may prevent chronicity. After a variable period, the pain becomes more diffuse and no longer simply work-related, and there is often severe distress. It is seen in keyboard workers and others who perform the same task without breaks for prolonged periods, and also in musicians. When it arises at work, it is often at a time of changing work practices, shortage of staff or disharmony. Middle managers find it difficult to deal with and this compounds the stress.

It is seen throughout the developed world. It peaked in incidence in Australia in the 1970s and 1980s but has largely disappeared there, apparently because of changes in work practices, improvements in early medical management, changes in workers' compensation legislation, and reduced media discussion of the problem.

Treatment

If possible there should be a brief period off work and a gradual return to activity once the pain has settled. Use of analgesia and NSAIDs, with physiotherapy, is helpful during the initial phase to prevent a vicious circle developing. Amitriptyline or gabapentin is helpful for some patients.

A review of working practices and the positioning of screen, keyboard and chair are essential, as is support of the patient by their manager. Musicians are helped by expert advice on playing technique and should reduce playing times temporarily, but not stop completely.

Temporomandibular pain dysfunction syndrome

This is a disorder of the temporomandibular joint associated with nocturnal tooth grinding or abnormalities of bite. It particularly occurs in anxious people. It gives rise to pain in one or both temporomandibular joints.

Dental correction of the bite helps a few but when no dental cause is found, low-dose tricyclic antidepressant therapy is used. Many patients are exposed to much unnecessary dental treatment.

Chronic regional pain syndrome type 1 (previously called reflex sympathetic dystrophy or Sudek's atrophy)

This is defined as 'a complex disorder or group of disorders that may develop as a consequence of trauma affecting the limbs, with or without obvious nerve lesions'. It may also develop after central nervous system lesions (e.g. strokes) or without cause. It occurred in 1.5% of soldiers injured in Vietnam. Its features are pain and other sensory abnormalities, including hyperaesthesia, autonomic vasomotor dysfunction, leading to abnormal blood flow and sweating, and motor system abnormalities. This leads to structural changes of superficial and deep tissues (trophic changes). Not all components need be present. The sensory, motor and sympathetic nerve changes are not restricted to the distribution of a single nerve and may be remote from the site of injury. The early phase – with pain, swelling and increased skin temperature – is difficult to diagnose but potentially reversible.

After a period of weeks or months, a second, still painful, dystrophic phase develops, characterized by articular stiffness, cold skin and trophic changes, often with localized osteoporosis.

A late phase involves continued pain, skin and muscle atrophy, and muscle contractures, and is extremely disabling.

Diagnosis is initially clinical – a high index of suspicion and recognizing the unusual distribution of the pain. A three-phase bone scan shows diffuse or patchy increase in uptake in the affected limb in all three phases: early (a few seconds – arterial); middle (a few minutes – soft tissue); and late (several hours – mineral). The bone phase abnormalities appear early and well before demineralization is seen on X-ray. There is never loss of joint space, which distinguishes the appearances from the periarticular osteoporosis of inflammatory joint disease.

Treatment

Management is difficult and the problem often very disabling. The evidence base to treatment is poor, as in many rare disorders. Early diagnosis, effective pain relief and general care of the patient are essential. NSAIDs and corticosteroids are used in the early phase, together with active exercise of the limb encouraged by a physiotherapist. Calcitonin or intravenous disodium pamidronate may also help at this stage. If pain persists despite initial treatment, a phentolamine test is used to test for evidence of sympathetically maintained pain (i.v. infusion of up to 40 mg with careful cardiac monitoring). If this is positive, a stellate ganglion block is used for upper limb and a sympathetic chain block for lower limb involvement. Guanethidine (an alpha-blocking agent) or lidocaine administered to the limb under tourniquet is also useful. Referral to a pain management clinic is essential.

Chronic regional pain syndrome type II is discussed on page 1116.

ANALGESIC AND ANTI-INFLAMMATORY DRUGS FOR MUSCULOSKELETAL PROBLEMS

The key to using drugs, particularly in chronic disorders and the elderly, is to balance risk and benefit and constantly to review their appropriateness. Box 10.3 shows the main drugs available.

Simple and compound analgesic agents

Simple agents such as paracetamol, aspirin, or codeine compounds (or combination preparations), used when necessary or regularly, relieve pain and improve function. Paracetamol with dextropropoxyphene (co-proxamol) should not be prescribed and is being withdrawn. Sleep may also be improved. Side-effects are relatively infrequent, although drowsiness and constipation occur with codeine preparations, especially in the elderly.

Stronger analgesics, such as dihydrocodeine or morphine derivatives, should be used only with severe pain.

Non-steroidal anti-inflammatory drugs (NSAIDs)

NSAIDs have anti-inflammatory and centrally acting analgesic properties. They inhibit cyclo-oxygenase (COX), a key enzyme in the formation of prostaglandins, prostacyclins and thromboxanes (see Fig. 14.31). There are two specific cyclo-

Box 10.3 Analgesics and NSAIDs

Analgesics (in order of potency)

Advise that they be taken only if needed. Maximum doses are indicated here.

Paracetamol	500–1000 mg	6-hourly
Paracetamol with codeine	1–2 tablets	6-hourly
Paracetamol with dihydrocodeine	1–2 tablets	Every 6–8 hours
Dihydrocodeine	30–60 mg	Every 6–8 hours

Non-steroidal anti-inflammatory drugs (NSAIDs)

Always to be taken with food. Use slow-release preparations in inflammatory conditions or if more regular pain control is needed. Examples:

Ibuprofen	200–400 mg	Every 6–8 hours
Ibuprofen slow release	600–800 mg	12-hourly
Diclofenac	25–50 mg	8-hourly
Diclofenac slow release	75–100 mg	× 1–2 daily
Celecoxib*	100–200 mg	× 2 daily

*COX-2-specific NSAID (coxib).

FURTHER READING

Drenth JPH, Verheugt FWA. Do COX-2 inhibitors give enough gastrointestinal protection? *Lancet* 2007; **369:** 439–440.

oxygenase enzymes: COX-1, the constitutive form present in many normal tissues; and COX-2, the form mainly induced in response to pro-inflammatory cytokines and not found in most normal tissues. The latter is associated with oedema and the nociceptive and pyretic effects of inflammation. COX-2 appears to be constitutive in the kidney. Most of the older NSAIDs block both enzymes but with variable specificity; their therapeutic effect depends on blocking COX-2 and their side-effects mainly on blocking COX-1, for example, by the loss of gastric mucosal protection and a decrease in renal blood flow. COX-2 inhibitors (coxibs) were designed to reduce gastrointestinal toxicity by not blocking COX-1.

Uses

- *Short courses* of NSAIDs or coxibs are used occasionally in osteoarthritis and spondylosis, even when there is minimal inflammation. They are commonly used in musculoskeletal pain but simple analgesia is often more appropriate.
- *In crystal synovitis*, NSAIDs and coxibs have a true anti-inflammatory effect (see p. 538).
- *In chronic inflammatory synovitis*, NSAIDs and coxibs do not alter the chronic inflammatory process, or decrease the risk of joint damage, but they do reduce pain and stiffness.
- Slow-release preparations are useful for inflammatory arthritis and when more constant pain control is needed.
- Consider the patient's gastrointestinal and cardiac risks before prescribing NSAIDs or coxibs.
- NSAID gels have some value in chronic arthritis.

Side-effects

The most common with NSAIDs are indigestion or skin rashes. Gastric erosions and peptic ulceration with perforation and bleeding also occur. Ibuprofen should be avoided in combination with low dose aspirin as it significantly increases the risk of severe gastrointestinal bleeding. Proton pump inhibitors are probably the best drugs to protect those at high risk from serious gastrointestinal events and also reduce dyspepsia. H_2 blockers are less effective as gastroprotective agents. The value of prostaglandin E_2 analogues is limited by their tendency to cause nausea and diarrhoea. In the elderly, NSAIDs may cause gastric mucosal damage and gastrointestinal bleeding without warning symptoms, thereby causing significant morbidity and mortality. They may also reduce renal function, especially in the elderly (see Box 11.4). They should only be prescribed with a proton pump inhibitor in this group.

Coxibs produce fewer gastrointestinal side-effects but renal complications and fluid retention still occur. Rofecoxib was withdrawn because of an increase in cardiovascular events with prolonged use. It now appears that this may be a 'class' phenomenon of all NSAIDs and coxibs. Different agents may carry a lower risk in both groups but the studies remain unclear. Coxibs should be used in patients with a high risk of gastrointestinal disease and with no cardiovascular risk. Patients with a high risk of both may be better off taking an NSAID (diclofenac) and a proton pump inhibitor.

OSTEOARTHRITIS (OA)

Osteoarthritis is the most common type of arthritis. It occurs in synovial joints and is characterized by cartilage loss with an accompanying periarticular bone response. It is probably not a single disease entity but a multifactorial process in which mechanical factors have a central role. The whole joint structure including cartilage, subchondral bone, ligaments, menisci, synovium and capsule is involved. Pathologically, there is significant inflammation of articular and periarticular structures and alteration in cartilage structure. Enthesophytes and chrondrophytes seen on MRI precede radiologically apparent osteophytes and joint space narrowing. Clinically some patients complain of pain and disability. There are several patterns of involvement (p. 520).

Epidemiology

The prevalence of OA increases with age, and most people over 60 years will have some radiological evidence of it, although only a quarter of these will be symptomatic. It occurs world-wide, but with a variable distribution, e.g. in Asians, hip OA is less common and knee OA is more common than in Europeans. Women over 55 years are affected more commonly than are men of a similar age. There is a familial pattern of inheritance with distal interphalangeal joint involvement in nodal OA and also in primary generalized OA. OA has a variable distribution (Fig. 10.8). The resulting disabilities have major socio-economic resource implications, particularly in the developed world.

Aetiology (see Box 10.4)

Genes that encode collagen type II have been proposed as candidate genes for familial OA but the genes responsible have not yet been identified. Abnormal local mechanical factors play a role in most types of OA, possibly excluding the nodal form. Inflammation starting at periarticular entheses is seen during the inflammatory phase on MRI in nodal OA. Osteoarthritis is the result of active, sometimes inflammatory but potentially reparative processes rather than the inevitable result of trauma and ageing. Focal destruction of the articular cartilage is the common pathological feature. The spectrum of OA ranges from atrophic disease in which

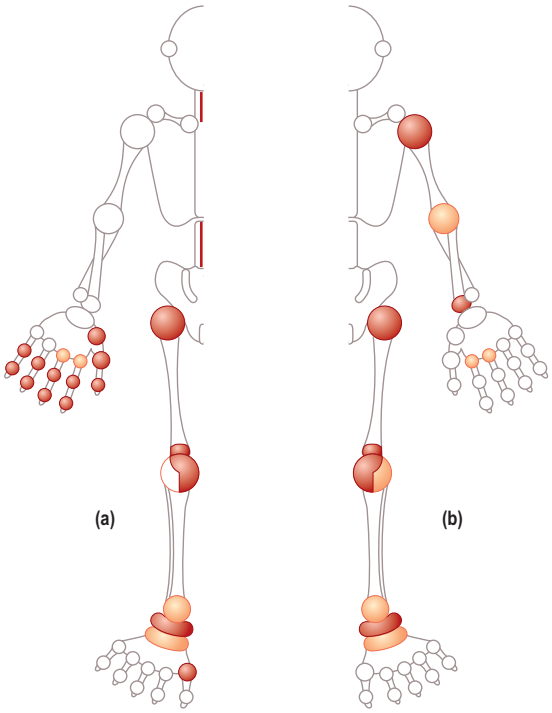


Fig. 10.8 **Typical distribution of affected joints in (a) primary generalized OA and (b) pyrophosphate arthropathy** (see p. 000). ●, more commonly affected; ●, less commonly affected.

Box 10.4 Factors predisposing to osteoarthritis

- *Obesity* – Predicts later risk of radiological and symptomatic OA in population studies.
- *Heredity* – Familial tendency to develop nodal and generalized OA.
- *Gender* – Polyarticular OA is more common in women; a higher prevalence after the menopause suggests a role for sex hormones.
- *Hypermobility* (see p. 554) – Increased range of joint motion and reduced stability lead to OA.
- *Osteoporosis* – There is reduced risk of OA.
- *Other diseases* – See Table 10.11.
- *Trauma* – A fracture through any joint. Meniscal and cruciate ligament tears cause OA of the knee.
- *Congenital joint dysplasia* – Alters joint biomechanics and leads to OA. Mild acetabular dysplasia is common and leads to earlier onset of hip OA.
- *Joint congruity* – Congenital dislocation of the hip or a slipped femoral epiphysis or Perthes' disease; osteonecrosis of the femoral head (see p. 565) in children and adolescents causes early-onset OA.
- *Occupation* – Miners develop OA of the hip, knee and shoulder, cotton workers OA of the hand, and farmers OA of the hip.
- *Sport* – Repetitive use and injury in some sports causes a high incidence of lower-limb OA.

cartilage destruction occurs without any subchondral bone response, to hypertrophic disease in which there is massive new bone formation at the joint margins.

Cartilage is a matrix of collagen fibres (mainly type II, see p. 568), which enclose a mixture of proteoglycans and water. Proteoglycans are present mainly as large molecular aggrecans, which consist of a protein core with polysaccharide side-chains of chondroitin sulphate and keratan sulphate.

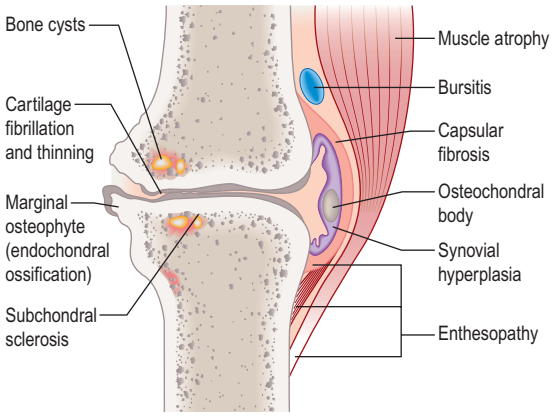


Fig. 10.9 **Early osteoarthritis of a knee** (diagram and X-ray). There is a medial compartment narrowing owing to cartilage thinning with subarticular sclerosis and marginal osteophyte formation (arrows).

The gene for human aggrecan has been cloned, and polymorphisms of the gene have been correlated with OA of the hand in older men.

Cartilage is smooth-surfaced and shock-absorbing. Under normal circumstances there is a dynamic balance between cartilage degradation by wear and its production by chondrocytes. Early in the development of OA this balance is lost and, despite increased synthesis of extracellular matrix, the cartilage becomes oedematous. Focal erosion of cartilage develops. Chondrocytes die and, although repair is attempted from adjacent cartilage, the process is disordered. Eventually the synthesis of extracellular matrix fails and the surface becomes fibrillated and fissured. Cartilage ulceration exposes underlying bone to increased stress, producing microfractures and cysts. The bone attempts repair but produces abnormal sclerotic subchondral bone and overgrowths at the joint margins, called osteophytes (Fig. 10.9). There is some secondary inflammation.

Pathogenesis

Several mechanisms have been suggested for the pathogenesis:

- Matrix loss is caused by the action of matrix metalloproteinases such as collagenase (MMP-1 and

MMP-13) which cleave collagen, and stromelysin (MMP-3) which is active against fibronectin and laminin in the extracellular matrix. MMPs are secreted by chrondrocytes in an inactive form. Extracellular activation then leads to the degradation of both collagen and proteoglycans.

- *Tissue inhibitors of metalloproteinases* (TIMPs) regulate the MMPs. Disturbance of this regulation may lead to increased cartilage degradation and contribute to the development of OA. TIMPs have not yet proven to be of therapeutic value.
- There is *synovial inflammation* in OA, and CRP may be raised. Interleukin-1 (IL-1) and tumour necrosis factor (TNF-α) release stimulates metalloproteinase production and IL-1 inhibits type II collagen production. IL-6 and IL-8 may also be involved. Anticytokine therapy has not yet been tested in OA. The production of cytokines by macrophages and of MMPs by chondrocytes in OA are dependent on the transcription factor NFκB. Inhibition of NFκB may have a therapeutic role in OA.
- *Growth factors*, including insulin-like growth factor (IGF-1) and transforming growth factor (TGF-β), are involved in collagen synthesis, and their deficiency may play a role in impairing matrix repair.
- *Vascular endothelial growth factor* (VEGF) from macrophages is a potent stimulator of angiogenesis and may contribute to inflammation and neovascularization in OA. Innervation can accompany vascularization of the articular cartilage where pressure and anoxia may stimulate these new nerves and lead to pain even when the inflammation has subsided.
- *Mutations* in the gene for type II collagen (*COL2A1*) have been associated with early polyarticular OA.
- *Twin studies* suggest a strong hereditary element underlying OA, and further studies may reveal genetic markers for the disease. The influence of genetic factors is estimated at 35–65%.
- In the *Caucasian* population there is an inverse relationship between the risk of developing OA and osteoporosis.
- A large population study has suggested that a high intake of *vitamin C* and other antioxidants may reduce the risk of OA. The lack of antioxidants is thought to contribute to many ageing processes.
- *Gender*. In *women*, weight-bearing sports produce a two- to threefold increase in risk of OA of the hip and knee. In *men*, there is an association between hip OA and certain occupations – farming and labouring.
- *Obesity* is a risk factor for developing OA in later life.

The term primary OA is sometimes used when there is no obvious known predisposing factor.

Box 10.4 shows some of the predisposing factors for the development of OA, whilst Table 10.11 shows other conditions that sometimes cause secondary arthritis.

Table 10.11 Causes of osteoarthritis

Primary OA	No known cause
Secondary OA	Pre-existing joint damage: Rheumatoid arthritis Gout Seronegative spondyloarthropathy Septic arthritis Paget's disease Avascular necrosis, e.g. corticosteroid therapy Metabolic disease: Chondrocalcinosis Hereditary haemochromatosis Acromegaly Systemic diseases: Haemophilia – recurrent haemarthrosis Haemoglobinopathies, e.g. sickle cell disease Neuropathies

Table 10.12 Features of nodal OA

Familial
Has a higher incidence in women
Typical pattern of polyarticular involvement of the hand joints
Develops in late middle age
Has a generally good long-term functional outcome
Associated with OA of the knee, hip and spine

Clinical features

Osteoarthritis affects many joints, with diverse clinical patterns. Hip and knee OA is the major cause of disability. Early OA is rarely symptomatic unless accompanied by a joint effusion, whilst advanced radiological and pathological OA is not always symptomatic.

Some flare-ups are due to inflammation and there may be a slight rise in ESR or CRP. Focal synovitis is caused by fragments of shed bone or cartilage. Radiological OA is usually, but not inevitably, progressive. This progression may be stepwise or continual. Radiological improvement is uncommon but has been observed, suggesting that repair is possible.

Symptoms

- Joint pain
- Joint gelling (stiffening and pain after immobility)
- Joint instability
- Loss of function.

Signs

- Joint tenderness
- Crepitus on movement
- Limitation of range of movement
- Joint instability
- Joint effusion and variable levels of inflammation
- Bony swelling
- Wasting of muscles.

Clinical subsets

Localized OA

Nodal OA (Table 10.12)

Joints of the hand are usually affected one at a time over several years, with the distal interphalangeal joints (DIPs) being more often involved than the proximal interphalangeal joints (PIPs). The onset may be painful and associated with tenderness, swelling and inflammation and impairment of hand function. At this stage enthesitis can be seen on MRI.

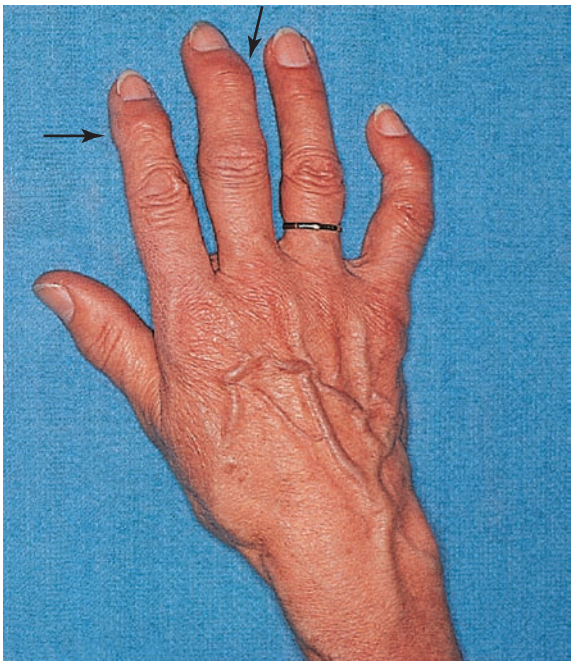


Fig. 10.10 **Severe nodal osteoarthritis.** The DIP joints demonstrate Heberden's nodes (arrows). The middle finger DIP joint is deformed and unstable. The thumb is adducted and the bony swelling of the first carpometacarpal joint is clearly shown – 'the squared hand of nodal OA'.

The inflammation often occurs around the female menopause. PIP-predominant nodal OA has a superficial similarity to early rheumatoid arthritis. Even if a weakly positive rheumatoid factor is found, it is of no significance. The inflammatory phase settles after some months or years, leaving painless bony swellings posterolaterally – Heberden's nodes (DIPs) and Bouchard's nodes (PIPs) – along with stiffness and deformity (Fig. 10.10). Functional impairment is slight for most, although PIP osteoarthritis restricts gripping more than DIP involvement. On X-ray, the nodes are marginal osteophytes and there is joint space loss.

Carpometacarpal and metacarpophalangeal OA of the thumb coexist with nodal OA and cause pain, which decreases as the joint stiffens. The 'squared' hand in OA is caused by bony swelling of the carpometacarpal joint and fixed adduction of the thumb. Function is rarely severely compromised.

Polyarticular hand OA is associated with a slightly increased frequency of OA at other sites.

Hip OA (see p. 512)

Hip OA affects 7–25% of white adult Caucasians but is significantly less common in black African and Asian populations. There are two major subgroups defined by the radiological appearance. The most common is *superior-pole hip OA*, where joint space narrowing and sclerosis predominantly affect the weight-bearing upper surface of the femoral head and adjacent acetabulum. This is most common in men and unilateral at presentation, although both hips may become involved because the disease is progressive. Early onset of hip OA is associated with acetabular dysplasia. Less commonly, *medial cartilage loss* occurs. This is most common in women and associated with hand involvement (nodal generalized OA – NGOA), and is usually bilateral. It is more rapidly disabling.

Knee OA

The prevalence of symptomatic knee OA is 40% in individuals aged over 75 years. It is commoner in women than in men. There is a strong relationship with obesity. The disease is generally bilateral and strongly associated with polyarticular OA of the hand in elderly women, sometimes alone, sometimes as part of generalized OA. The medial compartment is most commonly affected and leads to a varus (bow-legged) deformity. There is often also retropatellar OA. Previous trauma, meniscal and cruciate ligament tears are risk factors for developing knee OA.

Primary generalized OA

This is less common than nodal OA of the hands but is usually seen in combination. It is also called 'nodal generalized OA' (NGOA). The other joints affected are the knees, first MTP, hip, and intervertebral (spondylosis). There is a female preponderance and a strong familial tendency. NGOA is associated with immune complex deposition and may have an autoimmune cause. Its onset is often sudden and severe.

Erosive OA

This is rare. The DIPs and PIPs are inflamed and equally affected. In contrast to nodal OA, the functional outcome is poor. Radiologically, there are marked subchondral cysts. Erosive OA may develop into RA and may not be a true subset of OA.

Crystal-associated OA

This is most commonly seen with calcium pyrophosphate deposition in the cartilage (chondrocalcinosis). *Chondrocalcinosis* increases in frequency with age, but is usually asymptomatic. The joints most commonly affected are the knees (hyaline cartilage and fibrocartilage) and wrists (triangular fibrocartilage, see Fig. 10.8). There is patchy linear calcification on X-ray (Fig. 10.11).

A chronic arthropathy (pseudo-OA) occurs, predominantly in elderly women with severe chondrocalcinosis. There is a florid inflammatory component and marked osteophyte and cyst formation visible on X-rays. The joints affected differ from NGOA, being predominantly the knees, then wrists and shoulders, but also elbows, ankles and hips. Chondrocalcinosis is associated with pseudogout, an acute crystal-induced arthritis (see p. 539).

A rare, rapidly destructive arthritis in elderly women, affecting shoulders, hips and knees, is associated with finding crystals of calcium apatite in a bloody joint effusion. The outlook is poor and joints require early surgical replacement.

Investigations in OA

- **Blood tests**. There is no specific test; the ESR is normal although high sensitivity CRP may be slightly raised. Rheumatoid factor and antinuclear antibodies are negative.
- **X-rays** are abnormal only when the damage is advanced. They are useful to assess severity for operative intervention. For knees, a standing X-ray (stressed) is used to assess cartilage loss and 'skyline' views in flexion for patello-femoral OA.
- **MRI** demonstrates meniscal tears, early cartilage injury and subchondral bone changes.

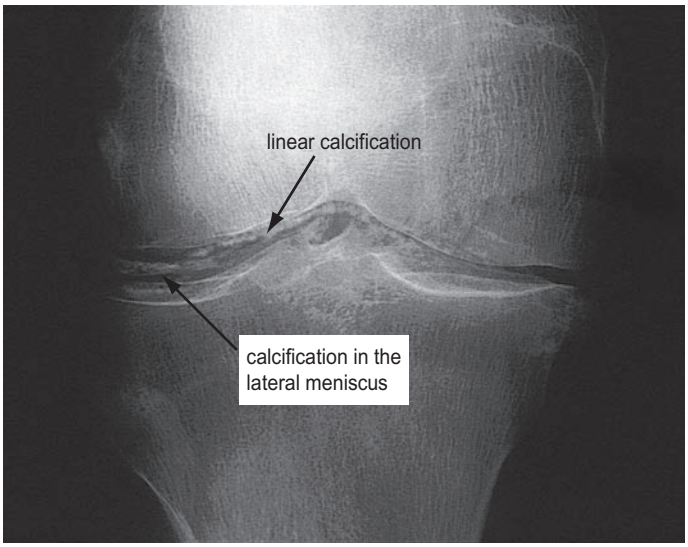


Fig. 10.11 **Chondrocalcinosis of the knee.** Note the linear calcification in the hyaline cartilage and calcification of the lateral meniscus (plus mild secondary OA).

- **Arthroscopy** reveals early fissuring and surface erosion of the cartilage.
- **Aspiration of synovial fluid** (if there is a painful effusion) shows a viscous fluid with few leucocytes and occasionally calcium pyrophosphate crystals (p. 536).

Management

The guiding principle is to treat the symptoms and disability, not the radiological appearances; depression and poor quadriceps strength are better predictors of pain than is radiological severity in OA of the knee. Education of the individual about the disease and its effects reduces pain, distress and disability and increases compliance with treatment. Psychological or social factors alter the impact of the disease.

Physical measures

Weight loss and exercises for strength and stability are useful. Hydrotherapy helps, especially in lower-limb OA. Local heat, ice packs, massage and rubifacients or local NSAID gels are all used. Insoles for shoes and a walking stick held on the contralateral side to the affected lower limb joint are useful.

There is increasing evidence that acupuncture helps knee OA. Other forms of complementary medicine are commonly used and, despite lack of scientific evidence, little is lost in trying it since a number of patients do seem to be helped.

Medication

Balance the potential benefit against potential side-effects. Patients should be prescribed short courses of simple analgesics before NSAIDs (see Box 10.3). NSAIDs or coxibs should be used intermittently.

Intra-articular corticosteroid injections produce short-term improvement when there is a painful joint effusion. Frequent injections into the same joint should be avoided. The role of intra-articular hyaluronan preparations is unproven.

The benefit of chondroitin sulphate and glucosamine compounds (sold as food supplements) is unproven. They appear to do no harm and some patients benefit.

There are no proven agents which halt or reverse OA although they are greatly needed. The role of bisphosphonates in reducing bone changes is unclear. The role of drugs which block tissue metalloproteinases or cytokines (see pathogenesis) is also unclear.

Surgery

Arthroscopic surgery for the knee is not beneficial. However, replacement arthroplasty has transformed the management of severe OA. The safety of hip and knee replacements is now equal, with a complication rate of about 1%; loosening, and late blood-borne infection are the most serious. Resurfacing hip surgery and unicompartmental knee replacement are less major procedures and appropriate for some patients. These slight but definite risks make it essential that the patient is certain that surgery is wanted, when all else has been tried. For the vast majority, a total hip or knee replacement reduces pain and stiffness and greatly increases function.

Other surgical procedures include realignment osteotomy of the knee or hip, excision arthroplasty of the first MTP and base of the thumb, and fusion of a first MTP joint.

FURTHER READING

Brandt KD, Doherty M, Lohmander LS (eds). *Osteoarthritis,* 2nd edn. Oxford: Oxford University Press 2003.

Kirkley A, Birmingham TB, Litchfield RB et al. A randomized trial of arthroscopic surgery for osteoarthritis of the knee. *New England Journal of Medicine* 2008; **359**: 1097–1107.

INFLAMMATORY ARTHRITIS

Inflammatory arthritis includes a large number of arthritic conditions in which the predominant feature is synovial inflammation. This disparate group includes postviral arthritis, rheumatoid arthritis, seronegative spondyloarthropathy, crystal arthritis and Lyme arthritis. The diagnosis of these conditions is helped by the pattern of joint involvement (Table 10.13), along with any nonarticular disease; a past and family history is helpful. The distribution of the affected joints (symmetrical or asymmetrical; large or small) as well as the periodicity of the arthritis (single acute, relapsing, chronic and progressive) also help in the diagnosis.

Certain non-articular diseases – for example, psoriasis, iritis, inflammatory bowel disease, non-specific urethritis or recent dysentery – suggest a seronegative spondyloarthropathy. There may be evidence of recent viral illness (rubella, hepatitis B or erythro (parvo) virus), of rheumatic fever, or of a tick bite and skin rash (Lyme disease). In early arthritis it

Table 10.13 Pattern of joint involvement in inflammatory arthritis

Diseases presenting as an inflammatory monoarthritis
Crystal arthritis, e.g. gout, pseudogout
Septic arthritis
Palindromic rheumatism
Traumatic ± haemarthrosis
Arthritis due to juxta-articular bone tumour
Occasionally, psoriatic, reactive, rheumatoid may present as monoarthritis
Diseases presenting as an inflammatory polyarthritis
Rheumatoid arthritis
Reactive arthritis
Seronegative arthritis associated with psoriasis or ankylosing spondyloarthropathy
Postviral arthritis
Lyme arthritis
Enteropathic arthritis
Arthritis associated with erythema nodosum

may not be possible to make a specific diagnosis until the disease has evolved.

There is a distinct genetic separation of rheumatoid-pattern synovitis and the seronegative group; RA (see below) is associated with a genetic marker in the class II major histocompatibility genes, whilst seronegative spondyloarthropathy shares certain alleles in the B locus of class I MHC genes, usually B27 (see p. 60).

In general the pain and stiffness of inflammatory arthritis are worse in the morning and after rest. This early-morning exacerbation may last several hours, in contrast to the much shorter post-rest gelling of OA. Inflammatory markers (ESR and CRP) are often raised in inflammatory arthritis, and there is often a normochromic, normocytic anaemia. Specific types of arthritis are discussed below.

RHEUMATOID ARTHRITIS (RA)

Rheumatoid arthritis is an autoimmune disease causing a chronic symmetrical polyarthritis with systemic involvement. Its course is extremely variable and it is associated with nonarticular features.

Epidemiology

RA has a world-wide distribution and affects 0.5–1% (with a female preponderance) of the population. The prevalence is high in the Pima Indian population and low in black Africans and Chinese people. The incidence, however, appears to be falling. RA is a significant cause of disability and mortality and carries a high socio-economic cost. It presents from early childhood (when it is rare) to late old age. The most common age of onset is between 30 and 50 years.

Aetiology and pathogenesis

The cause is multifactorial and genetic and environmental factors play a part.

- *Gender*. Women before the menopause are affected three times more often than men. After the menopause the frequency of onset is similar between the sexes, suggesting an aetiological role for sex hormones. A meta-analysis of the use of the oral contraceptive pill has shown no affect on RA overall, as previously thought, but it may delay the onset of disease.
- *Familial*. The disease is familial with an increased incidence in first degree relatives and a high concordance amongst monozygotic twins (up to 15%) and dizygotic twins (3.5%). In occasional families it affects several generations.
- *Genetic* factors are estimated to account for up to 60% of disease susceptibility. There is a strong association between susceptibility to RA and certain HLA haplotypes. HLA-DR4, which occurs in 50–75% of patients, correlates with a poor prognosis, as does HLA-DRB1* 0404/0401. The possession of a specific pentapeptide (Q-K-RAA) in HLA-DRB1 increases susceptibility; combined with a positive rheumatoid factor it identifies individuals with a 13 times greater risk for developing bone erosions in early disease. In a genome-wide association study in anti-CCP positive RA, an association was found with loci near *HLA-DRBI* and *PTPN22* in people of European descent. A variant of *PAD14* has also been shown to be a risk factor in Asian populations. These genes affect the presentation of autoantigens (for *HLA-DRBI*), T cell receptor signal transduction (for *PTPN22*) and targets of anti-CCP antibodies (citrullination of proteins) for *PAD14*. Other variants have also been identified which modify changes at a molecular level, e.g. tumour necrosis factor receptor associated factor (*TRAF1*), variant C5, STAT4.

Immunology

Many factors have been implicated but the chronic synovial inflammation is caused by ongoing T cell activation.

- The presence of *activated T cells* and macrophages and local production of rheumatoid factor autoantibodies in the joint in RA suggests that immune dysregulation plays a fundamental role in pathogenesis.
- *Anti-CCP (anti-citrullinated cyclic peptide)* (p. 503) antibodies react with proteins where arginine has been replaced by citrulline. CCPs predate the clinical disease by several years. They help distinguish early RA from transient polyarthritis.
- Alternatively the inflammation may be maintained by the *local production of rheumatoid factors* and continuous stimulation of macrophages via IgG Fc receptors. Considering the extent of synovial inflammation and lymphocytic infiltration, there are only minimal amounts of factors normally produced by T cells (interferon and IL-2 and -4). Conversely, the cytokines (IL-1, IL-8, TNF-α, granulocyte–macrophage colony-stimulating factor) and chemokines produced by macrophages (macrophage inflammatory protein (MIP) and monocyte chemoattractant protein (MCP)) and fibroblasts (producing IL-6) are abundant.
- *Activated mast cells* which release histamine and TNF-α may also play a role.
- *CD4-specific antibodies*, when they were used therapeutically, produced a specific helper T cell lymphopenia but did not significantly alter the disease, raising the possibility that T cells play a lesser role. They are no longer used.

- *Temporary CD20 positive B cell ablation* (a technique used for treating B cell lymphomas) induces remission, reinforcing the central place of rheumatoid factor production in maintaining the chronic inflammation of RA. As the B cells return, the CRP rises and the disease flares again.
- The *TNF superfamily of cytokines* produced mainly by activated macrophages and T cells play a role in the development of joint inflammation. Antibodies to TNF-α, IL-1 or specific blocking agents produce marked short-term improvement in synovitis, indicating the pivotal role of these cytokines in the chronic synovitis (see p. 54). They also reduce the malaise and tiredness felt in active RA.
- *Synovial fibroblasts* have high levels of the adhesion molecule, vascular cell adhesion molecule (VCAM-1 – a molecule which supports B lymphocyte survival and differentiation), decay accelerating factor (DAF – a factor that prevents complement-induced cell lysis) and cadherin-II (which mediates cell to cell interactions). These molecules may facilitate the formation of ectopic lymphoid tissue in synovium. Recent studies have shown that mice deficient in cadherin-II are resistant to a form of inflammatory arthritis. High-affinity antibodies are not a feature of RA, unlike other autoimmune diseases.

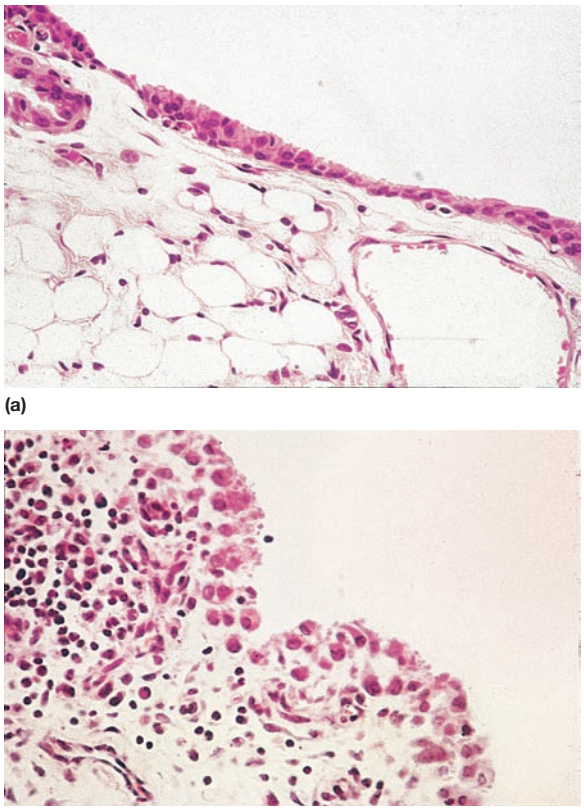
The *triggering antigen* remains unclear, although it is suggested that the glycosylation pattern of immunoglobulins may be abnormal in RA and lead to their becoming potentially antigenic. There is little evidence that collagen type II is the triggering antigen, although it is a cause of arthritis in animal models of RA.

Bacterial or slow virus infections have been implicated but are unproven. It has been suggested that an immune response to any pathogen is to produce autoantibodies by B cell clonal expansion. In susceptible individuals such clones may persist.

Pathology

Rheumatoid arthritis is typified by widespread persisting synovitis (inflammation of the synovial lining of joints, tendon sheaths or bursae). The cause of this is unclear, but the production of rheumatoid factors (RFs, see below) by plasma cells in the synovium and the local formation of immune complexes play a part. The normal synovium is thin and comprises a lining layer a few cells thick containing fibroblast-like synoviocytes and macrophages overlying loose connective tissue. The synoviocytes play a central role in synovial inflammation. In RA the synovium becomes greatly thickened to the extent that it is palpable as a 'boggy' swelling around the joints and tendons. There is proliferation of the synovium into folds and fronds, and it is infiltrated by a variety of inflammatory cells, including polymorphs, which transit through the tissue into the joint fluid, and lymphocytes and plasma cells. There are disorganized lymphoid follicles that are responsive to exogenous antigens. The normally sparse surface layer of lining cells becomes hyperplastic and thickened (Fig. 10.12). There is marked vascular proliferation. Increased permeability of blood vessels and the synovial lining layer leads to joint effusions that contain lymphocytes and dying polymorphs.

The hyperplastic synovium spreads from the joint margins on to the cartilage surface. This 'pannus' of inflamed synovium damages the underlying cartilage by blocking its normal route for nutrition and by the direct effects of cytokines on the chondrocytes. The cartilage becomes thinned and the underlying bone exposed. Local cytokine production and joint disuse combine to cause juxta-articular osteoporosis during active synovitis.



(a)

(b)

Fig. 10.12 **Histological appearance of RA synovium.** **(a)** Normal synovium. **(b)** Synovial appearances in established RA, showing marked hypertrophy of the tissues with infiltration by lymphocytes and plasma cells. From Shipley M *Colour Atlas of Rheumatology*, 3rd edn. London: Mosby-Wolfe, 1993, with permission.

Fibroblasts from the proliferating synovium also grow along the course of blood vessels between the synovial margins and the epiphyseal bone cavity and damage the bone. This is shown by MRI to occur in the first 3–6 months following onset of the arthritis, and before the diagnostic, ill-defined juxta-articular bony 'erosions' appear on X-ray (Fig. 10.13). This early damage justifies the use of DMARDs (see p. 530) within 3–6 months of onset of the arthritis. Low-dose steroids delay and anti-TNF-α agents halt or even reverse erosion formation. Erosions lead to a variety of deformities and contribute to long-term disability.

Rheumatoid factors (RFs) and anti-CCP antibodies (see also p. 503)

RFs are circulating autoantibodies that have the Fc portion of IgG as their antigen. The nature of the antigen means that they self-aggregate into immune complexes and thus activate complement and stimulate inflammation, causing chronic synovitis. Transient production of RFs is an essential part of the body's normal mechanism for removing immune complexes, but in RA they show a much higher affinity and their production is persistent and occurs in the joints. They

Fig. 10.13 **X-ray of early RA**, showing typical erosions at the thumb (black and white arrows) and middle MCP joints (black arrow) and at the ulnar styloid (white arrow).

Box 10.5 Differential diagnosis of early rheumatoid arthritis

- Postviral arthritis – rubella, hepatitis B or parvovirus
- Seronegative spondyloarthropathies
- Polymyalgia rheumatica
- Acute nodal osteoarthritis (PIPs and DIPs involved)

Box 10.6 Presentations of rheumatoid arthritis

- **Palindromic** – Monoarticular attacks lasting 24–48 hours; 50% progress to other types of RA.
- **Transient** – A self-limiting disease, lasting less than 12 months and leaving no permanent joint damage. Usually seronegative for IgM rheumatoid factor and CCP. Some of these may be undetected postviral arthritis.
- **Remitting** – There is a period of several years during which the arthritis is active but then remits, leaving minimal damage.
- **Chronic, persistent** – The most typical form, it may be seropositive or seronegative for IgM rheumatoid factor. The disease follows a relapsing and remitting course over many years. Seropositive (plus anti-CCP) patients tend to develop greater joint damage and long-term disability. They warrant earlier and more aggressive treatment with disease-modifying agents.
- **Rapidly progressive** – The disease progresses remorselessly over a few years and leads rapidly to severe joint damage and disability. It is usually seropositive (plus anti-CCP), has a high incidence of systemic complications and is difficult to treat.

are of any immunoglobulin class (IgM, IgG or IgA), but the most common tests employed clinically detect IgM rheumatoid factor. Around 70% of patients with polyarticular RA have IgM rheumatoid factor in the serum. Positive titres can predate the onset of RA.

The term seronegative RA is used for patients in whom the standard tests for IgM rheumatoid factor are persistently negative. They tend to have a more limited pattern of synovitis.

IgM rheumatoid factor is not diagnostic of RA and its absence does not rule the disease out; however, it is a useful predictor of prognosis. A persistently high titre in early disease implies more persistently active synovitis, more joint damage and greater disability eventually, and justifies earlier use of DMARDs.

RF and the antibody to CCP (see p. 503) together are more specific. RF indicates persistence but anti-CCP predicts the development of erosions and a worse prognosis. Measurement of CCP may take over from RF in day to day management because of its greater predictive value of persistence in early inflammatory arthritis.

CLINICAL FEATURES OF RA

Typical presentation

The most typical presentation of rheumatoid arthritis (approximately 70% of cases) begins as a slowly progressive, symmetrical, peripheral polyarthritis, evolving over a period of a few weeks or months. The patient is usually in her thirties to fifties, but the disease can occur at any age. Less commonly (15%) a rapid onset can occur over a few days (or explosively overnight) with a severe symmetrical polyarticular involvement. These patients often have a better prognosis. A worse than average prognosis (with a predictive accuracy of about 80%) is indicated by being female, a gradual onset over a few months, and a positive IgM rheumatoid factor, and/or anaemia within 3 months of onset. The differential diagnosis of early RA is shown in Box 10.5.

Symptoms and signs

The majority of patients complain of pain and stiffness of the small joints of the hands (metacarpophalangeal, MCP; proximal and distal interphalangeal, PIP, DIP) and feet (metatarsophalangeal, MTP). The wrists, elbows, shoulders, knees and ankles are also affected. In most cases many joints are involved, but 10% present with a monoarthritis of the knee or shoulder or with a carpal tunnel syndrome.

The patient feels tired and unwell and the pain and stiffness are significantly worse in the morning and may improve with gentle activity. Sleep is disturbed.

The joints are usually warm and tender with some joint swelling. There is limitation of movement and muscle wasting. Deformities develop as the disease progresses. Non-articular features develop (see below).

Other presentations

The presentation and progression of RA is variable. Presentations are shown in Box 10.6. Relapses and remissions occur either spontaneously or in response to drug therapy. In some patients the disease remains active, producing progressive joint damage. Rarely the process may cease ('burnt-out RA').

A *seronegative, limited synovitis* initially affects the wrists more often than the fingers and has a less symmetrical joint involvement. It has a better long-term prognosis, but some cases progress to severe disability. This form can be confused with psoriatic arthropathy, which has a similar distribution (p. 1240).

Table 10.14 Complications of rheumatoid arthritis
Complications of the condition
Ruptured tendons
Ruptured joints (Baker's cysts)
Joint infection
Spinal cord compression (atlantoaxial or upper cervical spine)
Amyloidosis (rare)
Side-effects of therapy
See Table 10.15

Palindromic rheumatism is unusual (5%) and consists of short-lived (24–72 h) episodes of acute monoarthritis. The joint becomes acutely painful, swollen and red, but resolves completely. Further attacks occur in the same or other joints. About 50% go on to develop typical chronic rheumatoid synovitis after a delay of months or years. The rest remit or continue to have acute episodic arthritis. The detection of IgM rheumatoid factor/anti-CCP predicts conversion to chronic, destructive synovitis.

Complications (Table 10.14)

Septic arthritis

This is a serious complication with significant morbidity and mortality. Affected joints are hot and inflamed with accompanying fever and a neutrophil leucocytosis in the blood. However, these signs are often absent, and any effusion, particularly of sudden onset, should be aspirated. *Staphylococcus aureus* is the most common organism. Blood cultures are often positive. Treatment is with systemic antibiotics (see p. 539) and drainage.

Amyloidosis (see p. 1072)

Amyloidosis is found in a very small number of people with severe rheumatoid arthritis. RA is the most common cause of secondary AA amyloidosis. AL amyloidosis causes a polyarthritis that resembles RA in distribution and is also often associated with carpal tunnel syndrome and subcutaneous nodules.

Joint involvement in RA

Hands and wrists

The impact of RA on the hands is severe. In early disease the fingers are swollen, painful and stiff. Inflamed flexor tendon sheaths increase functional impairment and may cause a carpal tunnel syndrome. Joint damage causes a variety of typical deformities. Most typical is a combination of ulnar drift and palmar subluxation of the MCPs (Fig. 10.14). This leads to unsightly deformity, but function may be remarkably good once the patient has learned to adapt, and pain is controlled. Fixed flexion (buttonhole or boutonnière deformity) or fixed hyperextension (swan-neck deformity) of the PIP joints impairs hand function.

Swelling and dorsal subluxation of the ulnar styloid leads to wrist pain and may cause rupture of the finger extensor tendons, leading in turn to a sudden onset of finger drop of the little and ring fingers predominantly, which needs urgent surgical repair.

Shoulders

RA commonly affects the shoulders. Initially the symptoms mimic rotator cuff tendonosis (see p. 507) with a painful arc syndrome and pain in the upper arms at night. As the joints become more damaged, global stiffening occurs. Late in the disease rotator cuff tears are common (see p. 507) and interfere with dressing, feeding and personal toilet.

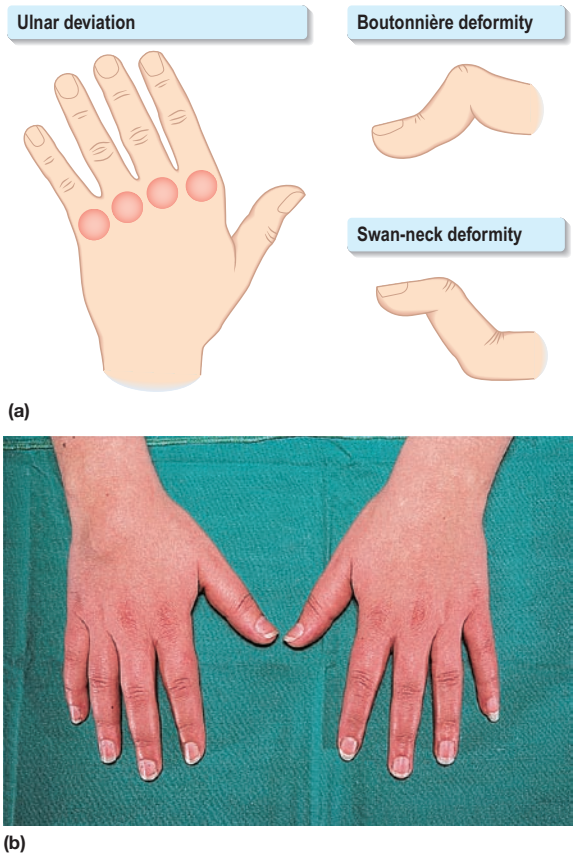


Fig. 10.14 **Rheumatoid arthritis (a) Characteristic hand deformities in RA. (b) Early rheumatoid arthritis** – dorsal tenosynovitis of the right wrist and small joints of both hands with spindling of the fingers.

Elbows

Synovitis of the elbows causes swelling and a painful fixed flexion deformity. In late disease flexion may be lost and severe difficulties with feeding result, especially combined with shoulder, hand and wrist deformities.

Feet

One of the earliest manifestations of RA is painful swelling of the MTP joints. The foot becomes broader and a hammer-toe deformity develops. Exposure of the metatarsal heads to pressure by the forward migration of the protective fibrofatty pad (Fig. 10.15) causes pain. Ulcers or callouses may develop under the metatarsal heads and the dorsum of the toes. Mid- and hindfoot RA causes a flat medial arch and loss of flexibility of the foot. The ankle often assumes a valgus position. Appropriate broad, deep, cushioned shoes are essential but rarely wholly adequate, and walking is often painful and limited. Podiatry helps and surgery may be required.

Knees

Massive synovitis and knee effusions occur, but respond well to aspiration and steroid injection (see p. 504). A persistent effusion increases the risk of popliteal cyst formation and

www.studentconsult.com
Ask the authors

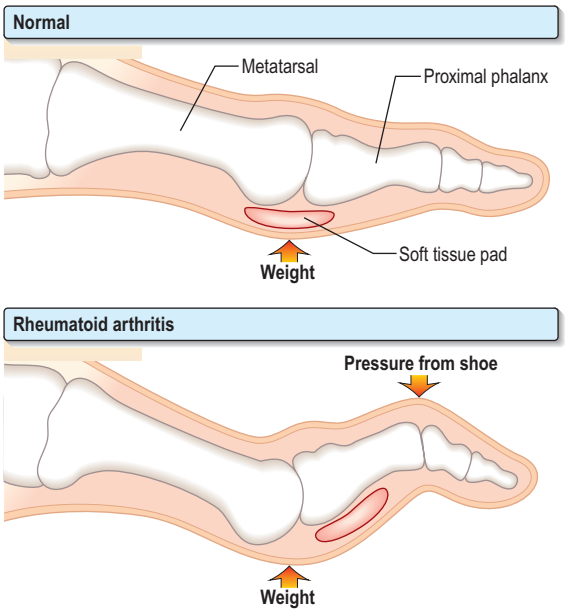


Fig. 10.15 **The toes in RA**, showing exposure of the metatarsal heads with forward migration of the soft tissue pad.

rupture (see p. 514). In later disease, erosion of cartilage and bone causes loss of joint space on X-ray and damage to the medial and/or lateral and/or retropatellar compartments of the knees. Depending on the pattern of involvement, the knees may develop a varus or valgus deformity. Secondary OA follows. Total knee replacement is often the only way to restore mobility and relieve pain.

Hips

The hips are rarely affected in early RA and are less commonly affected than the knees at all stages of the disease. Pain and stiffness are accompanied by radiological loss of joint space and juxta-articular osteoporosis. The latter may permit medial migration of the acetabulum (protrusio acetabulae). Later, secondary OA develops. Hip replacement is usually necessary.

Cervical spine

Painful stiffness of the neck in RA is often muscular, but it may be due to rheumatoid synovitis affecting the synovial joints of the upper cervical spine and the bursae which separate the odontoid peg from the anterior arch of the atlas and from its retaining ligaments. This synovitis leads to bone destruction, damages the ligaments and causes atlanto-axial or upper cervical instability. Subluxation and local synovial swelling may damage the spinal cord, producing pyramidal and sensory signs. MRI is the best way of visualizing this, but lateral flexed and extended neck X-rays can demonstrate instability. In late RA, difficulty walking which cannot be explained by articular disease, weakness of the legs or loss of control of bowel or bladder may be due to spinal cord compression and is a neurosurgical emergency. Image the cervical spine in flexion and extension in patients with RA before surgery or upper gastrointestinal endoscopy to check for instability and reduce the risk of cord injury during intubation. Severe neck involvement is less common than previously although the reason for this is not clear.

Other joints

The temporomandibular, acromioclavicular, sternoclavicular, cricoarytenoid and any other synovial joint can be affected.

Non-articular manifestations (Fig. 10.16)

Soft tissue surrounding joints

Subcutaneous nodules are firm, intradermal and generally occur over pressure points, typically the elbows, the finger joints and the Achilles tendon. Histologically there is a necrotic centre surrounded by rows of activated macrophages, which resembles synovitis without a synovial space.

The nodules may ulcerate and become infected, but usually resolve when the disease comes under control. The nodules can be removed surgically or injected with corticosteroids if causing a problem. They tend to recur.

The olecranon and other bursae may be swollen (bursitis).

Tenosynovitis of affected flexor tendons in the hand can cause stiffness and occasionally a trigger finger. Swelling of the extensor tendon sheath over the dorsum of the wrist is common.

Muscle wasting around joints is common. Muscle enzyme concentrations are normal; myositis is extremely rare. Corticosteroid-induced myopathy may occur.

Lungs (see also p. 871)

Peripheral, intrapulmonary nodules are usually asymptomatic but may cavitate. When pneumoconiosis is present (Caplan's syndrome), large cavitating lung nodules develop.

Other manifestations are:

- serositis causing pleural effusion
- pleural nodules
- fibrosing alveolitis (lung fibrosis)
- obstructive bronchiolitis
- infective lesions, e.g. TB in patients on biological DMARDs.

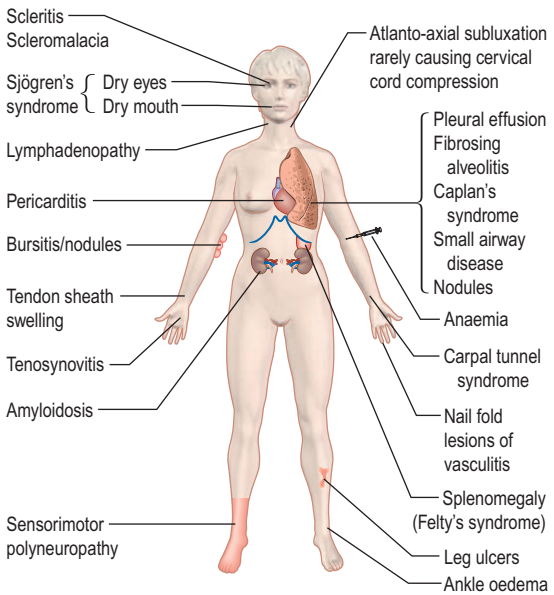


Fig. 10.16 **Non-articular manifestations of RA.**

Vasculitis

Vasculitis (see p. 548) is caused by immune complex deposition in arterial walls. It is uncommon. Smoking is a risk factor. Other manifestations are:

- nail fold infarcts due to cutaneous vasculitis
- widespread cutaneous vasculitis with necrosis of the skin (seen in patients with very active, strongly seropositive disease)
- mononeuritis multiplex (p. 1172)
- bowel infarction due to necrotizing arteritis of the mesenteric vessels (this may be indistinguishable from polyarteritis nodosa).

The heart and peripheral vessels

Poorly controlled RA with a persistently raised CRP is a risk factor for premature coronary artery and cerebrovascular atherosclerosis independent of traditional risk factors (i.e. raised lipids and hypertension). Clinical pericarditis is rare. In strongly seropositive RA, echocardiogram or post-mortem studies, however, show that 30–40% of patients have pericardial involvement.

Endocarditis and myocardial disease are rarely seen clinically, but are found at post-mortem in approximately 20% of cases. These are secondary to vasculitis.

Raynaud's syndrome occurs (see p. 809).

The nervous system

Neuropathies, either mononeuritis multiplex or a sensory loss in a glove and stocking pattern, are due to vasculitis of the vasa nervorum. Compression neuropathies such as carpal or tarsal tunnel syndrome are due to local synovial hypertrophy. Atlanto-axial subluxation can cause serious neurological abnormalities in advanced disease (see above).

The eyes

Scleritis and episcleritis occur in severe, seropositive disease and produce painful red lesions in the eye. Scleritis may lead to perforation of the eye (scleromalacia perforans) and requires active treatment with local and systemic corticosteroids.

Sicca syndrome causes dry mouth and eyes (see Sjögren's syndrome, p. 548).

The kidneys

Amyloidosis causes the nephrotic syndrome and renal failure. Presentation is with proteinuria. It occurs rarely in severe, long-standing rheumatoid disease and is due to the deposition of highly stable serum amyloid A protein (SAP) in the intercellular matrix of a variety of organs. SAP is an acute-phase reactant, produced normally in the liver. Amyloidosis is rare, and proteinuria in RA is more commonly due to DMARDs.

The spleen, lymph nodes and blood

Felty's syndrome is splenomegaly and neutropenia in a patient with RA. Leg ulcers or sepsis are complications. HLA-DR4 is found in 95% of patients, compared with 50–75% of patients with RA alone.

The lymph nodes may be palpable, usually in the distribution of affected joints. There may be peripheral lymphoedema of the arm or leg.

Anaemia is almost universal and is usually the normochromic, normocytic anaemia of chronic disease. It may be iron-deficient owing to gastrointestinal blood loss from NSAID ingestion, or rarely haemolytic (Coombs' positive). There may be a pancytopenia due to hypersplenism in Felty's syndrome or as a complication of DMARD treatment. A high platelet count occurs with active disease.

DIAGNOSIS AND INVESTIGATIONS

The diagnosis relies on the clinical features described above. The American College of Rheumatology (ACR) criteria are shown in Box 10.7 and are useful for epidemiological and investigative studies but unhelpful in early disease.

Initial investigations include:

- **Blood count**. Anaemia may be present. The ESR and/or CRP are raised in proportion to the activity of the inflammatory process and are useful in monitoring treatment.
- **Serology**. Anti-CCP (see p. 503) is positive earlier in the disease, and in early inflammatory arthritis indicates the likelihood of progressing to RA. Rheumatoid factor is present in approximately 70% of cases and ANA at low titre in 30%.
- **X-rays of the affected joint(s)** to establish a baseline. Only soft tissue swelling is seen in early disease. MRI demonstrates early erosions but is rarely warranted.
- **Aspiration of the joint** if an effusion is present. The aspirate looks cloudy owing to white cells. In a suddenly painful joint septic arthritis should be suspected (see p. 539).

Other investigations will depend on the clinical picture as outlined above. In severe disease extensive imaging of joints may be required. MRI is the technique of choice, especially for the knee and cervical spine.

MANAGEMENT OF RA (Box 10.8)

The diagnosis of RA inevitably causes concern and fear in the patient and requires a lot of explanation and reassurance. Early diagnosis with early referral to a rheumatologist is vital. The doctor and therapist should retain a positive approach and remind the patient that with the help of drugs most will continue to lead a more or less normal life despite their arthritis; 25% will recover completely. The earliest years are often the most difficult and people should be helped and encouraged to stay at work during this phase if possible. Uncertainty about when the disease will remit and flare, when

Box 10.7 Criteria for the diagnosis of rheumatoid arthritis (American College of Rheumatology, 1988 revision)

- Morning stiffness > 1 hour
- Arthritis of three or more joints
- Arthritis of hand joints and wrists

} For 6 weeks or more

- Symmetrical arthritis
- Subcutaneous nodules
- A positive serum rheumatoid factor
- Typical radiological changes (erosions and/or periarticular osteopenia)

Four or more criteria are necessary for diagnosis.

Box 10.8 Management of rheumatoid arthritis

- Establish the diagnosis clinically.
- Use NSAIDs and analgesics to control symptoms.
- Try to induce remission with i.m. depot methylprednisolone 80–120 mg if synovitis persists beyond 6 weeks.
- If synovitis recurs, refer to a rheumatologist to start sulfasalazine or methotrexate. Give a second dose of i.m. depot methylprednisolone.
- Refer for physiotherapy and general advice through a specialist team.
- If there is no significant improvement in 6–12 weeks as measured by less pain, less morning stiffness and reduced acute-phase response, use a combination of methotrexate and sulfasalazine.
- If no better, use an alternative agent, such as gold, D-penicillamine, leflunomide or anti-TNF-α therapy

Box 10.9 Problems associated with the use of corticosteroids

- Patients are increasingly anxious about the use of corticosteroids because of adverse publicity about their potential side-effects. This must be discussed frankly and the risks of not using corticosteroids in treatment should be described and balanced against the risks of the drugs themselves.
- Patients must be warned to avoid sugars and saturated fats and to eat less because of the risk of weight gain.
- The skin becomes thin and easily damaged.
- Monitor for diabetes and hypertension.
- Cataract formation may be accelerated.
- Osteoporosis develops within 6 months on doses above 7.5 mg daily. Monitor with DXA scan and treat with calcium and vitamin D and bisphosphonate (see p. 564).

and if drugs will work, and whether they produce side-effects, makes planning from day to day difficult. People learn to adjust remarkably but this takes time and support. A rheumatology unit will have a team, including doctors, specialist nurses and physiotherapists, to help the patient learn to cope. Leaflets and websites give helpful advice, as do local patient groups. Patients from socially deprived backgrounds and smokers have a worse prognosis. Statins have been shown to be of benefit in reducing cardiovascular risk and possibly inflammation whatever the cholesterol level; more studies are required.

Drug therapy

There is no curative agent available for RA but drugs are now available that prevent disease deterioration. Symptoms are controlled with analgesia and NSAIDs. Data now support the use of DMARDs early in the disease to prevent the long-term irreversible damaging effects of inflammation of the joints, and drugs that block TNF-α and IL-1 and the use of B cell ablation with rituximab are revolutionizing the management of RA.

Non-steroidal anti-inflammatory drugs (NSAIDs) and coxibs

Most patients with RA are unable to cope without an NSAID to relieve night pain and morning stiffness. NSAIDs do not reduce the underlying inflammatory process. They all act on the cyclo-oxygenase (COX) pathway (see Fig. 14.32). The individual response to NSAIDs varies greatly. It is desirable therefore to try several different drugs for a particular patient in order to find the best (see Box 10.8). Each compound should be given for at least a week. Start with an inexpensive NSAID with few side-effects and with which you are familiar. Regular doses are needed to be effective. The major side-effects of NSAIDs and the use of coxibs are discussed on page 517. If gastrointestinal side-effects are prominent, or the patient is over 65 years, add a proton pump inhibitor. Slow-release preparations (e.g. slow-release diclofenac, 75 mg, taken after supper), or a suppository at bedtime, usually work well and can be given in addition to daytime therapy if necessary. For additional relief a simple analgesic is taken as required (e.g. paracetamol or a combination of codeine or dihydrocodeine and paracetamol). Many patients need night sedation.

Corticosteroids

There is evidence to suggest that the early use of corticosteroids slows down the course of the disease.

Intra-articular injections with semicrystalline steroid preparations have a powerful but sometimes only short-lived effect.

Intramuscular depot injections (40–120 mg depot methylprednisolone) help to control severe disease flares, or can be used before a holiday or other life event, but should be used infrequently.

The use of oral corticosteroids has a number of problems (see Boxes 10.9 and 18.9). They are powerful disease-controlling drugs, but are avoided in the long term because side-effects are inevitable. Early intensive short-term regimens are used in some centres. Others use doses of 5–7.5 mg as maintenance therapy. There is some evidence that increased physical activity because of better symptom control reduces the risk of osteoporosis. Corticosteroids are also invaluable to patients with severe disease with extra-articular manifestations such as vasculitis.

Disease-modifying anti-rheumatic drugs

(Table 10.15)

- Traditional DMARDs, which mainly act through cytokine inhibition, reduce inflammation, with a reduction of joint swelling, a fall in the plasma acute-phase reactants and slowing of the development of joint erosions and irreversible damage. Their beneficial effect is not immediate (hence 'slow-acting agents') and may be partial or transient.
- DMARDs often only have a partial effect, achieving between 20 and 50% improvement by ACR criteria for disease remission (morning stiffness <15 min; no fatigue, joint pain, joint tenderness or soft tissue swelling; an ESR of <30 in women and <20 in men). TNF-α blocking agents and rituximab achieve almost 70% improvement in around 20% of patients and they also act more rapidly.
- Sulfasalazine, methotrexate, leflunomide, TNF-α blockers, ciclosporin and rituximab have all been shown to reduce the rate of progressive joint damage in early and late disease.

Generally DMARDs are used after symptomatic treatment. However, patients positive for RF and anti-CCP with persis-

Table 10.15 Disease-modifying anti-rheumatic drugs (DMARDs)

Drug	Dose	Side-effects	Monitoring to detect side-effects
Sulfasalazine (enteric coated)	500 mg daily after food, increasing to 2–3 g daily	Nausea Skin rashes and mouth ulcers	
		Neutropenia and/or thrombocytopenia	Initial, 2 weeks, then 4-monthly
		Abnormal liver biochemistry	Initial, 2 weeks, then 4-monthly
Methotrexate	2.5 mg increasing to 25 mg weekly, orally or s.c.	Nausea, mouth ulcers and diarrhoea	
		Abnormal liver biochemistry	Initial, 2 weeks, then monthly
		Neutropenia and/or thrombocytopenia	Initial, weekly, then monthly
		Renal impairment	Initial, then every 3–6 months
		Pulmonary fibrosis (rare)	Baseline chest X-ray
Leflunomide	100 mg daily for 1–3 days, then 20 (or 10) mg daily or 10–20 mg daily	Diarrhoea	
		Neutropenia and/or thrombocytopenia	Initial, then 2 weekly; monthly at 6 months
		Abnormal liver biochemistry Alopecia Hypertension	Initial, then 2 weekly; monthly at 6 months
TNF-α blockers		**For all**	
Etanercept (alone or with methotrexate)	s.c. 25 mg × 2 weekly or 50 mg weekly	Injection site reactions Infections, e.g. TB and septicaemia Hypersensitivity reactions Heart failure Rare – demyelination and autoimmune syndromes Reversible lupus-like syndrome	See British Society for Rheumatology Guidelines www.rheumatology.org.uk/guidelines/clinicalguidelines Stop if no response after 6 months
Adalimumab (with methotrexate)	s.c. 40 mg alternate weeks		
Infliximab (with methotrexate)	i.v. 3–10 mg/kg every 4–8 weeks		
Other biological agents (used with methotrexate)			
Anakinra	1 mg/kg s.c. daily	Injection site reaction Serious reactions – rare	Used after failure of anti-TNF agents
Rituximab	i.v. 500–1000 mg	Hypo/hypertension, skin rash, nausea, pruritis, back pain Rare – toxic epidermal necrolysis	
Abatacept	i.v. 10 mg/kg on days 1, 15, 30 and then monthly	Nausea, vomiting Headache Hypersensitivity – rare	
Tocilizumab	i.v. 8 mg/kg infusion	Headache, skin eruption, stomatitis, fever, anaphylactic reactions	Phase III studies

tent joint swelling have a poor prognosis and they should be treated early with DMARDs – preferably before the appearance of erosions on X-rays of hands and feet. Studies of early RA suggest that intervention with DMARDs at 6 weeks to 6 months improves the outcome. The use of combinations of three or four drugs (steroids, methotrexate, sulfasalazine and hydroxychloroquine) in early RA, reducing the number of agents once remission has been achieved, is more controversial and its long-term efficacy is yet to be proven. Toxicity appears to be reduced with such combinations, possibly because inflammation is controlled effectively. Effective treatment with DMARDs reduces the increased cardiovascular risk seen in patients with RA.

You must not, however, overtreat people who are not going to develop erosions. Most prognostic assessments in early disease are only 80% accurate for risk of erosions and subsequent disability, and clinical judgement retains a central place in managing RA. DMARDs are usually prescribed by a rheumatologist and early referral is essential.

Sulfasalazine

This is a combination of sulfapyridine and 5-amino-salicylic acid. Sulfapyridine is probably the active component. It is well tolerated and for many is the first-choice DMARD, especially in younger patients and women who are planning a family. It produces a response in about half the patients in the first 3–6 months. Serious side-effects are rare, being mainly leucopenia and thrombocytopenia.

Methotrexate

This is considered by many to be the drug of choice. It should not be used in pregnancy. Conception should be delayed for 3–6 months off the drug for either partner. It is given at an initial weekly dose of 2.5–7.5 mg orally, increased up to 15–25 mg if necessary. It is well tolerated and this therapy can be introduced early in the disease. Nausea or poor absorption may limit its efficacy. It can be self-administered by subcutaneous injection. Oral folic acid should be given in addition to reduce side-effects, although it may marginally

reduce efficacy. Full blood counts and liver biochemistry should be monitored. It usually works within 1–2 months. More patients remain on this agent than on most other DMARDs, indicating that it is effective and has relatively few side-effects.

Leflunomide

This DMARD exerts an immunomodulatory effect by preventing pyrimidine production in proliferating lymphocytes through blockade of the enzyme dihydroorotate dehydrogenase. Most cells are able to bypass this blockade but T cells cannot – thus it has a specific effect to block clonal expansion of T cells by slowing progression through the maturation phases G to S1. It is well absorbed when taken orally. It has a long half-life of 4–28 days. A loading dose of 100 mg daily is often used for 1–3 days (but is not essential), then 20 mg daily (10 mg if diarrhoea is a problem). The main side-effects are diarrhoea, nausea, alopecia and rash. Diarrhoea diminishes with time. Blood monitoring is obligatory. The onset of action is 4 weeks compared with 6 weeks for methotrexate. The initial response is similar to sulfasalazine but improvement continues and is better sustained at 2 years. Leflunomide works in some patients who have failed to respond to methotrexate, or it can be administered with methotrexate to enhance the response. It needs a 'washout' of 2 years before conception (3 months in men) so is best avoided in women planning a family.

Biological DMARDs

TNF-α blockers. The availability of agents that block TNF-α has significantly changed the traditional use of DMARDs. Because of their cost they are used after at least two DMARDs (usually sulfasalazine and methotrexate) have been tried. They represent a major therapeutic advance.

- *Etanercept* is a fully humanized p75 TNF-α receptor IgG_1 fusion protein given by self-administered subcutaneous injection. Around 65% of patients respond well. Some develop an injection reaction.
- *Adalimumab* is a fully human monoclonal antibody against TNF-α given along with methotrexate.
- *Infliximab* is a monoclonal antibody against TNF-α, given intravenously and co-prescribed with methotrexate to prevent loss of efficacy because of antibody formation.

These products slow or halt erosion formation in up to 70% of patients with RA and produce healing in a few. Patients often comment that their malaise and tiredness improve in a manner that is not seen with other DMARDs. There is a 50% incidence of secondary failure with infliximab in the first year, less with etanercept and adalimumab. Changing to another anti-TNF agent is justified and often regains control of the disease. Studies suggest that anti-TNF drugs given for 12 months in early disease may induce remission although this is not yet common practice. Cost effectiveness studies are underway.

Side-effects. The evidence to date of increased tumour development is controversial. A few people become ANA positive and develop a reversible lupus-like syndrome, leucocytoclastic vasculitis, some extracutaneous involvement or interstitial lung disease. Reactivation of old TB may occur (a pre-treatment chest X-ray is recommended, with or without skin testing) and any TB should be treated before using these agents. There is an increased risk of infections which requires close monitoring.

These agents are extremely expensive when compared with traditional DMARDs but they may save costs in the longer term by reducing disability and the need for hospitalization. Their use should be restricted to specialist centres. To date there is no evidence of an adverse effect on pregnancy outcome but care is essential.

Other biological agents

Anakinra is a human recombinant IL-1 receptor antagonist which is used in combination with methotrexate. It is used after other anti-TNF agents have been unsuccessful, mainly in clinical trials.

Rituximab is a genetically engineered chimeric monoclonal antibody (p. 76) that causes lysis of CD-20 positive B cells. CD20 is a pan-B cell surface antigenic phosphoprotein. Its expression is restricted to pre-B and mature B cells but it is not present on stem cells and is lost before differentiation into plasma cells. Rituximab produces significant improvement in RF-factor positive RA for 8 months to several years when used alone or in combination with corticosteroids and/or methotrexate. This is associated with a 6–9-month B cell lymphopenia with little change in circulating immunoglobulins. A re-flare is often accompanied by a return of peripheral lymphocytes and a rise in CRP. Rituximab can be reused as the disease flares. Repeated courses over up to 5 years are acceptable and well tolerated and around 80% of RF-positive patients respond with 50–60% showing persistent disease control. It is worth trying in patients who have failed to respond to anti-TNF agents. There may be an increased risk of chest infections, and immunoglobulin levels may fall progressively and need to be monitored.

Abatacept is a recombinant fusion protein of CTLA4 and the Fc portion of IgG_1, which selectively modulates T cell activation. It has FDA approval for use in the USA for RA.

Tocilizumab, a humanized anti-IL-6 receptor, is currently in clinical trials.

Drugs used less commonly

Gold (sodium aurothiomalate) is given by deep intramuscular injection. A test dose of 10 mg is followed by weekly doses of 50 mg until response occurs, usually in about 3 months. Occasionally remission is induced. If there is no improvement after a total dose of 1 g, treatment should be stopped. If a response is obtained, change to 4-weekly injections, continued for up to 5 years. Side-effects occur in a third of patients on i.m. gold. Rare side-effects include pulmonary fibrosis, colitis, polyneuropathy and cholestatic jaundice. Gold-induced glomerulonephritis can occur, particularly in patients who are HLA-DR3 positive, and routine urinalysis for proteinuria is performed.

Hydroxychloroquine, an antimalarial, 200–400 g daily, is well tolerated. It is used alone in mild disease or as an adjunct to other DMARDs. Retinopathy is the most serious side-effect, but this is rare before 6 years of treatment. Some rheumatologists arrange an initial check of macular function with an Amsler chart, and further reviews as retinopathy is irreversible.

Penicillamine, 125 g daily for 1 month, then 500–750 g daily, is given before food and for at least 3 months before improvement occurs. If proteinuria exceeds 2 g/24 h the drug must be stopped. Loss of taste is reversible. Other rare side-effects include a lupus erythematosus-like syndrome and a myasthenia gravis-like syndrome.

Azathioprine at a maximum dose of 2.5 mg/kg and *cyclophosphamide* 1–2 mg/kg have been used, usually when other DMARDs have been ineffective. They are often used when extra-articular features are severe, particularly with vasculitis. There is a high risk of neutropenia and possibly liver toxicity from azathioprine in patients who are genetically deficient for the enzyme thiopurine methyltransferase (TPMT) which metabolizes azathioprine into its active metabolites. Pretesting is not evidence based but a wise precaution.

Ciclosporin 2.5–4 mg/kg is used for active rheumatoid arthritis when conventional therapy has been ineffective. Side-effects include a rise in creatinine level and hypertension.

Physical measures

Patients with RA need constant advice and support from physiotherapists and nurse specialists, especially while they are learning to adjust. A combination of rest for active arthritis and exercises to maintain joint range and muscle power is essential. Exercise in a hydrotherapy pool is popular and effective. Advice about managing activities of daily living despite the arthritis, and about gadgets, seating or structural changes in the home or at work are helpful. Family and friends should be involved.

Surgery

Surgery has a useful role in the long-term approach to patient management. Its main objectives are prophylactic, to prevent joint destruction and deformity, and reconstructive, to restore function.

Single-joint disease can be treated by surgical synovectomy to reduce the bulk of inflamed tissue and prevent damage. Excision arthroplasty of the ulnar styloid reduces pain and the risk of extensor tendon damage. Excision arthroplasties of the metatarsal heads reduce metatarsal pain and relieve pressure points. The major surgical advance has been the development of total replacement arthroplasty of the hip, knee, finger joints, elbow and shoulder. Such procedures need careful planning and preparation, and the expected outcomes and risks should be explained to the patient.

PROGNOSIS

A poor prognosis is indicated by:

- A *clinical picture* of an insidious rather than an explosive onset of RA, female sex, increasing number of peripheral joints involved and the level of disability at the onset.
- *Blood tests* showing a high CRP/ESR, normochromic normocytic anaemia and high titres of anti-CCP antibodies and of rheumatoid factor.
- *X-rays* with early erosive damage (N.B. ultrasound and MRI can show cartilage and bone damage prior to conventional X-rays).

Prognosis can be altered dramatically with early DMARD therapy under expert supervision.

FURTHER READING

Kremer JM. COMET's path and the new biologicals in rheumatoid arthritis. *Lancet* 2008; **372**: 347–348.

Smolen JS, Aletaha D, Koeller M et al. New therapies for treatment of rheumatoid arthritis. *Lancet* 2007; **370**: 1861–1874.

Table 10.16 Seronegative spondyloarthropathies

Ankylosing spondylitis (AS)
Psoriatic arthritis
Reactive arthritis (sexually acquired, Reiter's disease)
Post-dysenteric reactive arthritis
Enteropathic arthritis (ulcerative colitis/Crohn's disease)

SERONEGATIVE SPONDYLOARTHROPATHIES

This awkward title describes a group of conditions affecting the spine and peripheral joints which cluster in families and are linked to certain type 1 HLA antigens (Table 10.16).

The joint involvement is usually more limited than that seen in RA and its distribution is different. There are associated extra-articular and genetic features. These diseases occasionally present in childhood.

Histologically the synovitis itself is difficult to distinguish from that of RA, but there is no production of rheumatoid factors – hence 'seronegative'. Anti-CCP is usually negative. Inflammation of the enthesis (junction of ligament or tendon and bone) and joint ankylosis develop more commonly than in RA. All are associated with an increased frequency of sacroiliitis and an increased frequency of HLA-B27.

Aetiology

The common aetiological thread of these disorders is their striking association with HLA-B27, particularly ankylosing spondylitis (AS). HLA type B27 is present in >90% of Caucasians with AS but only 8% of controls. HLA-B27 exhibits a number of unusual characteristics including a high tendency to mis-fold but its aetiological relevance remains unclear. The role of class I HLA antigens in pathogenesis is supported by the fact that HLA-B27 transgenic mice spontaneously develop arthritis, skin, gut and genitourinary lesions.

There are clues that infections play a role, possibly by molecular mimicry, with parts of the organism which are structurally similar to the HLA molecule triggering cross-reactive antibody formation. This is unproven. AIDS is increasing the prevalence of reactive arthritis and spondylitis in sub-Saharan Africa even in the absence of HLA-B27. The explanation for this changing epidemiology is unclear.

The types of arthritis that follow a precipitating infection are called reactive arthritis (p. 535).

The specialized immune systems of the gut and genitourinary mucous membranes may also play a causal role, perhaps reacting to local infections or to antigens which cross the damaged mucosa.

ANKYLOSING SPONDYLITIS (AS)

This is an inflammatory disorder of the spine affecting mainly young adults. It occurs world-wide and affects 1% of men and 0.5% of women in Caucasian populations. The frequency of AS in different populations is roughly paralleled by the incidence of HLA-B27; Africans and Japanese have a low incidence of both HLA-B27 and ankylosing spondylitis, while the North American Haida Indians have a high incidence of both. There are at least 24 subtypes of HLA-B27 (B*2701–B*2724). Some appear to increase risk; others have a protec-

tive role. Twin studies indicate a much higher disease concordance in HLA-B27-positive monozygotic (up to 70%) twins than in dizygotic twins (about 20–25%). HLA-B60 and HLA-DR1 are also associated but are of less significance. There are also other genes lying within the major histocompatibility complex (interleukin-1 gene cluster and the gene *CYP2D6*) which also influence susceptibility to AS but the disease is polygenic with loci on other chromosomes likely to harbour additional susceptibility genes. The disease occurs in men and women (2.5:1). Spinal involvement is milder in women, so men are more likely to present with symptoms of the disease (in a ratio of 4:1). Spinal and sacroiliac disease have a high degree of heritability (>90%). Both sexes may develop iritis.

Epidemiology and pathogenesis

Environmental factors may also be involved but although Gram-negative organisms, e.g. Yersinia, Klebsiella, Salmonella, Shigella, can cause a reactive arthropathy, there is no conclusive evidence for their involvement in the pathogenesis of AS.

There is lymphocyte and plasma cell infiltration and local erosion of bone at the attachments of the intervertebral and other ligaments (enthesitis). This heals with new bone (syndesmophyte) formation.

Clinical features

Episodic inflammation of the sacroiliac joints in the late teenage years or early twenties is the first manifestation of AS. Pain in one or both buttocks and low back pain and stiffness are typically worse in the morning and relieved by exercise. Initially the diagnosis is often missed because the patient is asymptomatic between episodes and radiological abnormalities are absent. Retention of the lumbar lordosis during spinal flexion is an early sign. Later, paraspinal muscle wasting develops.

Classification criteria with a high sensitivity and specificity have been established. The presence of three of the four following indices in adults <50 years with chronic back pain indicates AS:

- morning stiffness >30 min
- improvement of back pain with exercise but not rest
- awakening because of back pain during second half of the night only
- alternating buttock pain.

Spinal stiffness can be measured by Schoeber's test – a tape measure is placed in the midline 10 cm above the dimples of Venus. Any movement of a marker at 15 cm during flexion is recorded. A reading of less than 5 cm implies spinal stiffness. Individuals may be able to touch the floor with a stiff back if they have good hip movements but serial measurement of the finger tip to floor distance highlights any change.

Non-spinal complications (uveitis or costochondritis) suggest the diagnosis of seronegative spondyloarthropathies (Box 10.10). Costochondral junction inflammation causes anterior chest pain. Measurable reduction of chest expansion is due to costovertebral joint involvement.

Peripheral joint involvement is asymmetrical and affects a few, predominantly large joints. Hip involvement leads to fixed flexion deformities of the hips and further deterioration of the posture. Young teenage boys occasionally present with a lower-limb monoarthritis (see p. 553) which later develops into AS.

Acute anterior uveitis is strongly associated with HLA-B27 in AS and related diseases and is occasionally the presenting complaint. Severe eye pain, photophobia and blurred vision are an emergency (see p. 1092). Overall clinical assessment is based on pain, tenderness, stiffness and fatigue using, for example, the Bath Ankylosing Spondylitis Disease Activity Index.

Box 10.10 Non-articular problems in seronegative spondyloarthropathies

- Uveitis, in all types
- Cutaneous lesions in reactive arthritis (keratoderma blenorrhagica), histologically identical to pustular psoriasis
- Nail dystrophy, in psoriasis and reactive arthritis
- Aortitis, occasionally in AS and reactive arthritis

Investigations

- **Blood**. The ESR and CRP are usually raised.
- **HLA testing** is rarely of value because of the high frequency of HLA-B27 in the population, but may give supporting evidence in a difficult case.
- **X-rays**. The medial and lateral cortical margins of both sacroiliac joints lose definition owing to erosions and eventually become sclerotic (Fig. 10.17). The earliest radiological appearances in the spine are blurring of the upper or lower vertebral rims at the thoracolumbar junction (best seen on a lateral X-ray) caused by an enthesitis at the insertion of the intervertebral ligaments. These changes may eventually affect the whole spine.

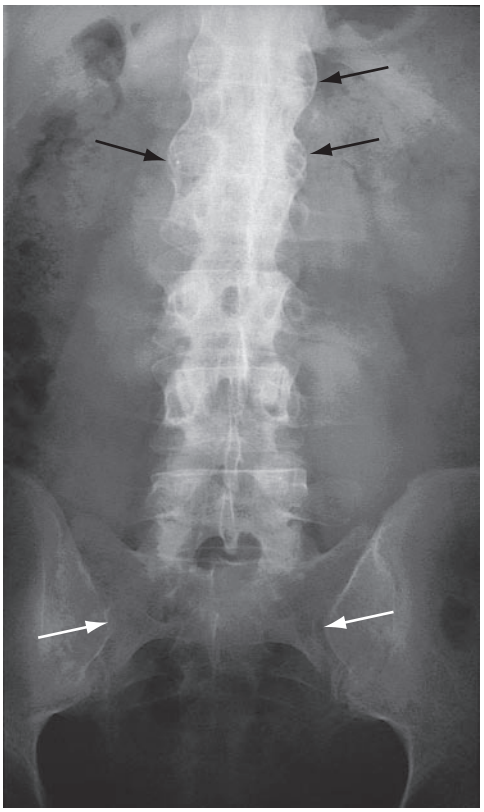


Fig. 10.17 **X-ray of ankylosing spondylitis.** The sacroiliac joints are eroded and show marginal sclerosis (white arrows). There is bridging syndesmophyte formation at the thoracolumbar junction (black arrows).

Persistent inflammatory enthesitis causes bony spurs (syndesmophytes). Syndesmophytes are more vertically oriented than the beak-like osteophytes of spondylosis and the disc is preserved, unlike in spondylosis (see p. 510). Syndesmophytes cause bony ankylosis and permanent stiffening. The sacroiliac joints eventually fuse, as may the costovertebral joints, reducing chest expansion. Calcification of the intervertebral ligaments and fusion of the spinal facet joints and syndesmophytes leads to what is often called a 'bamboo' spine (Fig. 10.18).

- **MRI** demonstrates sacroiliitis before it is seen on X-rays.

FURTHER READING

Braun J, Sieper J. Ankylosing spondylitis. *Lancet* 2007; **369**: 1379–1390.

Treatment

- The key to effective management of AS is early diagnosis so that a regimen of preventative exercises is started before syndesmophytes have formed. Morning exercises aim to maintain spinal mobility, posture and chest expansion.
- Failure to control pain and to encourage regular spinal and chest exercises leads to an irreversible dorsal kyphosis and wasted paraspinal muscles. This, along with stiffening of the cervical spine, makes forward vision difficult.
- When the inflammation is active, the morning pain and stiffness are too severe to permit effective exercise, an evening dose of a long-acting or slow-release NSAID or an NSAID suppository improves sleep, pain control and exercise compliance. Peripheral arthritis and enthesitis are managed with NSAIDs or local steroid injections.
- Methotrexate is effective for peripheral arthritis but not for spinal disease.
- In patients with persistent, active inflammation TNF-α blocking drugs (see Table 10.15) produce rapid, dramatic and sustained reduction of symptoms and of spinal and peripheral joint inflammation. Around half are able to stop NSAIDs. Relapse occurs on stopping therapy but may be delayed by several months making intermittent treatment feasible. Rituximab does not help seronegative arthritis.

Fig. 10.18 **X-ray of bamboo spine in ankylosing spondylitis.** In advanced disease there is calcification of the interspinous ligaments and fusion of the facet joints as well as syndesmophytes at all levels. The sacroiliac joints fuse.

Prognosis

With exercise and pain relief, the prognosis is excellent and over 80% of patients are fully employed. Anti-TNF therapies are likely to reduce the morbidity of severe disease, reducing the risk of permanent spinal stiffness and progressive peripheral joint disease.

Patients should be made aware that they risk passing the HLA-B27 gene to 50% of their children. HLA-B27 positive offspring then have a 30% risk of developing AS.

PSORIATIC ARTHRITIS (see also p. 1240)

The term 'psoriatic arthritis' describes a variety of different patterns of arthritis and enthesitis seen in people with psoriasis or with a family history of psoriasis. Five to eight per cent of individuals with psoriasis develop one of several different patterns of arthritis for which there is no serological marker. The aetiology and pathogenesis is described on page 1240.

Clinical features

The arthritis is typically more limited in distribution and less severe than in RA. The skin disease can be mild and may develop after the arthritis.

- A **distal interphalangeal arthritis** is the most typical pattern of joint involvement in psoriasis. It is unsightly, but rarely disabling, and there is often adjacent nail dystrophy (see p. 1266) reflecting inflammation in the enthesis extending into the nail root. Cutaneous lesions, interphalangeal joint synovitis and tenosynovitis causing a 'sausage' finger or toe (dactylitis) are seen in pauciarticular psoriatic arthritis (and in reactive arthritis).
- A **seronegative symmetrical polyarthritis** similar to rheumatoid arthritis also occurs.
- **Arthritis mutilans** affects about 5% of patients with psoriatic arthritis and causes marked periarticular osteolysis and bone shortening ('telescopic' fingers) (Fig. 10.19), in which, despite the deformity, pain may be mild and function often surprisingly good.

Fig. 10.19 **Hand showing psoriatic arthritis mutilans.** All the fingers are shortened and the joints unstable, owing to underlying osteolysis.

- Radiologically, **psoriatic arthritis** is erosive but the erosions are central in the joint, not juxta-articular, and produce a 'pencil in cup' appearance (Fig. 10.20).
- A **unilateral or bilateral sacroiliitis** and spondylitis develops in 15% of patients with early involvement of the cervical spine; only 50% are HLA-B27 positive.

Treatment and prognosis

NSAIDs and/or analgesics help the pain but they can occasionally worsen the skin lesions. Local synovitis responds to intra-articular corticosteroid injections.

In milder, polyarticular cases, sulfasalazine or methotrexate slows the development of joint damage.

When the disease is severe, methotrexate or ciclosporin is given because they control both the skin lesions and the arthritis. Anti-TNF-α agents, e.g. etanercept (see p. 530), are highly effective and safe for severe skin and joint disease. Corticosteroids orally may destabilize the skin disease and are best avoided but are valuable when injected into a single inflamed joint. Rituximab has no role in treating psoriatic arthritis.

The prognosis for the joint involvement is generally better than in RA.

REACTIVE ARTHRITIS

Reactive arthritis is a sterile synovitis, which occurs following an infection (see also post-streptococcal arthritis, p. 554).

Seronegative spondyloarthropathy develops in 1–2% of patients after an acute attack of dysentery, or a sexually acquired infection – non-specific urethritis (NSU) in the male, non-specific cervicitis in the female. In male patients who are HLA-B27 positive the relative risk is 30–50. Being HLA-B27 positive is not obligatory, however. Women are less commonly affected.

Aetiology

A variety of organisms can be the trigger, including some strains of *Salmonella* or *Shigella* spp. in bacillary dysentery. *Yersinia enterocolitica* causes diarrhoea and a reactive arthritis. In NSU the organisms are *Chlamydia trachomatis* or *Ureaplasma urealyticum*.

Patients with reactive arthritis are not more susceptible to infection but appear to respond differently with bacterial persistence and altered intracellular replication when HLA-B27 is present. Bacterial antigens or bacterial DNA have been found in the inflamed synovium of affected joints, suggesting that this persistent antigenic material is driving the inflammatory process. Other environmental factors may explain why, even in susceptible individuals, repeated infections do not necessarily produce a reactive arthritis. The methods by which HLA-B27 increases susceptibility to reactive arthritis may include:

- T cell receptor repertoire selection
- molecular mimicry causing autoimmunity against HLA-B27 and/or other self antigens
- mode of presentation of bacteria-derived peptides to T lymphocytes.

These are not mutually exclusive.

There are other organisms that also trigger reactive arthritis but have a different genetic basis; see post-streptococcal arthritis (p. 554), gonococcal arthritis (p. 540) and brucellosis (p. 540). In these, the borderline between reactive arthritis and septic arthritis is more indistinct and they can cause both.

Clinical features (Fig. 10.21)

The arthritis is typically an acute, asymmetrical, lower-limb arthritis, occurring a few days to a couple of weeks after the infection. The arthritis may be the presenting complaint if the infection is mild or asymptomatic. Enthesitis is common, causing plantar fasciitis or Achilles tendon enthesitis (see p. 515). Seventy per cent recover fully within 6 months but many have a relapse.

In susceptible individuals with reactive arthritis, sacroiliitis and spondylitis may also develop. Sterile conjunctivitis is common. Acute anterior uveitis may complicate more severe

Fig. 10.20 **X-ray of psoriatic arthritis.** There is osteolysis of the metatarsal heads and central erosion of the proximal phalanges to produce the 'pencil in cup' appearance (circle). All the lesser toes are subluxed.

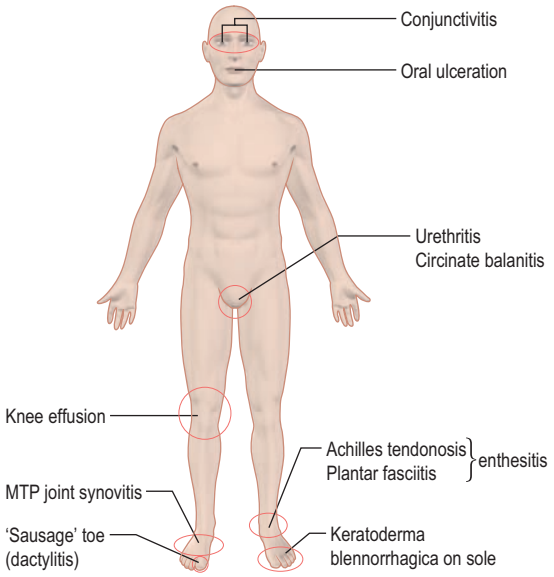


Fig. 10.21 **Clinical features of reactive arthritis.**

or relapsing disease but is not synchronous with the arthritis.

The skin lesions resemble psoriasis:

- Circinate balanitis in the uncircumcised male causes painless superficial ulceration of the glans penis. In the circumcised male the lesion is raised, red and scaly. Both heal without scarring.
- Keratoderma blenorrhagica – the skin of the feet and hands develops painless, red and often confluent raised plaques and pustules histologically similar to pustular psoriasis.
- Nail dystrophy may occur.

Other features

These include:

- bilateral conjunctivitis in 30%
- the classically described triad of *Reiter's disease* – urethritis, arthritis and conjunctivitis.

Treatment

There is some evidence that treating persisting infection with antibiotics will alter the course of the arthritis, once it has developed. Cultures should be taken and any infection treated. Sexual partners may require specialist advice about, and treatment for, sexually acquired infections.

Pain responds well to NSAIDs and local corticosteroid injections. The majority of individuals with reactive arthritis have a single attack which settles, but a few develop a disabling relapsing and remitting arthritis. Relapsing cases are sometimes treated with sulfasalazine or methotrexate (see Table 10.15). TNF-α blocking agents are being increasingly used in severe and persistent disease.

ENTEROPATHIC ARTHRITIS ASSOCIATED WITH INFLAMMATORY BOWEL DISEASE

Enteropathic synovitis occurs in approximately 10–15% of patients with ulcerative colitis and Crohn's disease see p. 288. The link between the bowel disease and the inflammatory arthritis is not clear. Selective mucosal leakiness may expose the individual to antigens that trigger synovitis.

The arthritis is asymmetrical and predominantly affects lower-limb joints. An HLA-B27-associated sacroiliitis or spondylitis is seen in 5% of patients with inflammatory bowel disease and is independent of disease activity. The joint symptoms may predate the development of bowel disease and lead to its diagnosis.

Remission of ulcerative colitis or total colectomy usually leads to remission of the joint disease, but arthritis may persist even in well-controlled Crohn's disease.

Treatment

The inflammatory bowel disease should be treated (see p. 285). In all cases of enteropathic arthritis, the joint disease should be managed symptomatically with NSAIDs, although they may make diarrhoea worse. A monoarthritis is best treated by intra-articular corticosteroids. Sulfasalazine is more frequently prescribed than mesalazine as it may help both bowel and joint disease. The TNF-α blocking drug infliximab is used in inflammatory bowel disease (p. 289) and can help the arthritis.

CRYSTAL ARTHRITIS

Aetiology

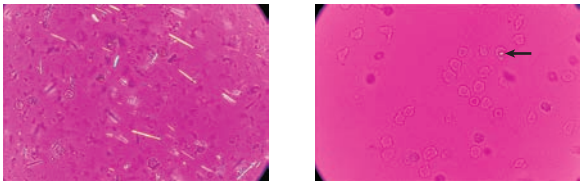
Two main types of crystal account for the majority of crystal-induced arthritis. They are sodium urate and calcium pyrophosphate and are distinguished by their different shapes and refringence properties under polarized light with a red filter (Fig. 10.22). Rarely, crystals of calcium apatite (see p. 521) or cholesterol cause acute synovitis.

GOUT AND HYPERURICAEMIA

Gout is an inflammatory arthritis associated with hyperuricaemia and intra-articular sodium urate crystals.

Epidemiology

The prevalence of gout is increasing mainly in developed countries, and in Europe and the USA it is approximately 0.2%, although hyperuricaemia in this population occurs in about 5%. Gout develops in men more than women (10:1) and rarely occurs before young adulthood (when it suggests a specific enzyme defect), and seldom in premenopausal females. The prevalence in older females is increasing with increased diuretic use. Hyperuricaemia is common in certain ethnic groups (e.g. Maoris).



(a) (b)

Fig. 10.22 **Crystal arthritis (a) Needle-shaped urate crystals. (b) A small intracellular pyrophosphate crystal.** Both viewed under polarized light with a red filter.

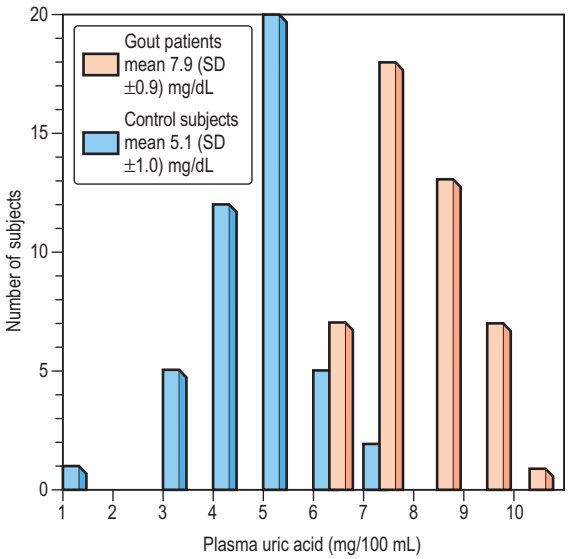


Fig. 10.23 **Serum uric acid levels in normals and in patients with gout.** 7.9 mg/dL is equivalent to 474 µmol/L, 5.1 mg/dL is equivalent to 306 µmol/L. From Snaith ML, Scott JT Uric acid clearance in patients with gout and normal subjects. (1971) *Annals of Rheumatic Disease*, with permission from the BMJ Publishing Group.

Table 10.17 Causes of hyperuricaemia

Impaired excretion of uric acid
Chronic renal disease (clinical gout unusual)
Drug therapy, e.g. thiazide diuretics, low-dose aspirin
Hypertension
Lead toxicity
Primary hyperparathyroidism
Hypothyroidism
Increased lactic acid production from alcohol, exercise, starvation
Glucose-6-phosphatase deficiency (interferes with renal excretion)
Increased production of uric acid
Increased purine synthesis de novo due to:
Hypoxanthine-guanine-phosphoribosyl transferase (HGPRT) reduction (an X-linked inborn error causing the Lesch–Nyhan syndrome)
Phosphoribosyl-pyrophosphate synthase overactivity
Glucose-6-phosphatase deficiency with glycogen storage disease type 1 (patients who survive develop hyperuricaemia due to increased production as well as decreased excretion)
Increased turnover of purines due to:
Myeloproliferative disorders, e.g. polycythaemia vera
Lymphoproliferative disorders, e.g. leukaemia
Others, e.g. carcinoma, severe psoriasis

Uric acid levels start to rise after puberty and are higher in men than in women until the female menopause. There is a normal distribution of serum uric acid in the population with a skewed distribution at the upper end of the range. Hyperuricaemia is defined as a serum uric acid level greater than two standard deviations from the mean (420 µmol/L in males, 360 µmol/L in females). This is close to the limit of solubility.

Most people with hyperuricaemia are asymptomatic even though sodium urate crystals are found in the joints. However, osteoarthritic joints are more prone to gouty attacks. The range for gouty individuals is higher than for normals, but the curves overlap (Fig. 10.23). Serum uric acid levels increase with age, obesity, a high-protein diet, a high alcohol consumption (particularly beer drinkers), a high fructose intake from sweetened drinks, combined hyperlipidaemia, diabetes mellitus, ischaemic heart disease and hypertension (metabolic syndrome, p. 799). There is often a family history of gout.

Pathogenesis

Causes of hyperuricaemia are shown in Table 10.17. In many patients with gout there is no obvious cause. Hyperuricaemia results from inadequate renal excretion of uric acid relative to its production and is the major determinant for developing gout. Only around 30% is derived from the diet.

Uric acid is the final product of purine metabolism in humans and levels in the blood depend on the balance between purine synthesis and the ingestion of dietary purines, and the elimination of urate by the kidney (66%) and intestine (33%). The body pool is about 1000 mg and 60% is turned over daily.

Acute gout is precipitated by the ingestion by polymorpholeucocytes of sodium urate crystals causing the release of proinflammatory cytokines from phagosomes and complement activation. Colchicine works by inhibition of microtubule formation necessary for this to occur.

Uric acid synthesis. The last two steps of purine metabolism in humans are the conversion of hypoxanthine to xanthine and of xanthine to uric acid, catalysed by the enzyme xanthine oxidase. Humans lack the enzyme uricase.

Uric acid excretion (Fig. 10.24). Uric acid excretion by the kidney is complex. It is completely filtered by the glomerulus; 98–100% is then reabsorbed in the proximal tubule via the urate transporter-1 (URAT1) expressed at the apical membrane and 50% is secreted by the distal tubule. Some postsecretory reabsorption also takes place. Low-dose aspirin blocks uric acid secretion. High-dose aspirin also blocks its reabsorption, leading to increased net excretion. Insulin resistance enhances uric acid resorption.

Ninety per cent of patients with gout have impaired excretion of uric acid, 10% have increased production also and in less than 1% an inborn error of metabolism leads to purine overproduction. One-third of uric acid is eliminated in the faeces.

Clinical features

Hyperuricaemia, whether in the context of the metabolic syndrome (see p. 799) or not, causes four clinical syndromes:

- acute sodium urate synovitis – gout
- chronic polyarticular gout
- chronic tophaceous gout
- urate renal stone formation (p. 610).

Acute gout presents typically in a middle-aged male with sudden onset of agonizing pain, swelling and redness of the

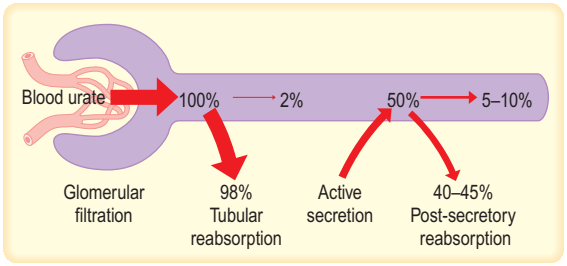


Fig. 10.24 **Urate renal transport.** The net result is that about 5–10% of the glomerular load is excreted in the urine under normal circumstances.

first MTP joint. The attack occurs at any time, but may be precipitated by too much food or alcohol, by dehydration or by starting a diuretic. Untreated attacks last about 7 days. Recovery is typically associated with desquamation of the overlying skin. In 25% of attacks, a joint other than the great toe is affected.

In severe attacks, overlying crystal cellulitis makes gout difficult to distinguish clinically from infective cellulitis. A family or personal history of gout and the finding of a raised serum urate suggest the diagnosis but, if in doubt, blood and joint fluid cultures should be taken.

Chronic polyarticular gout is unusual, except in elderly people on long-standing diuretic treatment, in renal failure, or occasionally in men who have been started on treatment with allopurinol too soon after an acute attack.

Chronic tophaceous gout (see below).

Investigations

The clinical picture is often diagnostic, as is the rapid response to NSAIDs or colchicine.

- **Joint fluid microscopy** is the most specific and diagnostic test but is technically difficult.
- **Serum uric acid** is usually raised (>600 µmol/L). If it is not, recheck it several weeks after the attack, as the level falls immediately after an acute attack. Acute gout never occurs with a serum uric acid in the lower half of the normal range below the saturation point of 360 µmol/L.
- **Serum urea and creatinine** are monitored for signs of renal impairment.

Treatment

The use of NSAIDs or coxibs in high doses rapidly reduces the pain and swelling. Initial doses, taken with food, are:

- naproxen: 750 mg immediately, then 500 mg every 8–12 hours
- diclofenac: 75–100 mg immediately, then 50 mg every 6–8 hours
- lumiracoxib: 100 mg once daily
- indometacin: 75 mg immediately, then 50 mg every 6–8 hours. Although regarded as the 'gold standard' treatment by some, the frequency of side-effects is unacceptably high with indometacin.

After 24–48 hours, reduced doses are given for a further week. *Caution*: NSAIDs may cause renal impairment. In individuals with renal impairment or a history of peptic ulceration, alternative treatments include:

- colchicine: 1000 µg immediately, then 500 µg every 6–12 hours, but this causes diarrhoea
- corticosteroids: intramuscular or intra-articular depot methylprednisolone.

Dietary advice

The first attacks may be separated by many months or years and are managed symptomatically. Individuals should be advised to reduce their alcohol intake, especially beer, which is high in purines and fructose. A diet which reduces total calorie and cholesterol intake and avoids such foods as offal, some fish and shellfish and spinach, all of which are rich sources of purines, is advised. This can reduce serum urate by 15% and delay the need for drugs that reduce serum urate levels. Dietary advice is readily available on the internet.

Treatment with agents that reduce serum uric acid levels

Allopurinol should only be used when the attacks are frequent and severe (despite dietary changes) or associated with renal impairment or tophi, or when the patient finds NSAIDs or colchicine difficult to tolerate. It should never be started within a month of an acute attack and always be started under cover of a course of NSAID or colchicine for the first 2–4 weeks before and 4 weeks after starting allopurinol, as it may induce acute gout.

Allopurinol (300–600 mg) blocks the enzyme xanthine oxidase, which converts xanthine into uric acid (see Fig. 15.11). It reduces serum uric acid levels rapidly and is relatively non-toxic but should be used at low doses (50–100 mg) in renal impairment. Alternatively the dose can be increased gradually from 100 mg every few weeks until the uric acid level is below the 360 µmol/L level. Skin rashes and gastrointestinal intolerance are the most common side-effects. A hypersensitivity reaction is the most serious adverse event, with rash, fever, eosinophilia, hepatitis and progressive renal failure. This is rare, as is bone marrow suppression.

Febuxostat (80–120 mg) is a new non-purine analogue inhibitor of xanthine oxidase and available in some countries. It appears to be well tolerated but longer term studies are needed. Uricosuric agents lower the serum uric acid. Benzbromarone is available in some European countries and probenecid in the USA. The angiotensin I-receptor antagonist, losartan, is uricosuric in hypertensive patients on diuretics.

Chronic tophaceous gout

Individuals with very high levels of uric acid can present with different clinical pictures. In chronic tophaceous gout, sodium urate forms smooth white deposits (tophi) in skin and around joints. They occur on the ear, the fingers (Fig. 10.25) or the

Fig. 10.25 **Gout with tophi.** Courtesy of Dr Varkey Alexander, Al-Sabah Hospital, Kuwait.

Achilles tendon. Large deposits are unsightly and ulcerate. There is chronic joint pain and sometimes superimposed acute gouty attacks.

Periarticular deposits lead to a halo of radio-opacity and clearly defined ('punched out') bone cysts on X-ray.

Tophaceous gout is often associated with renal impairment and/or the long-term use of diuretics. There may be acute or chronic urate nephropathy or renal stone formation. Whenever possible, stop the diuretics or change to less urate-retaining ones, such as bumetamide. Treat with allopurinol and/or uricosuric agents (see above).

PSEUDOGOUT (PYROPHOSPHATE ARTHROPATHY)

Calcium pyrophosphate deposits in hyaline and fibrocartilage produce the radiological appearance of chondrocalcinosis (see p. 522). Shedding of crystals into a joint precipitates acute synovitis which resembles gout, except that it is more common in elderly women and usually affects the knee or wrist. The attacks are often very painful. In young people it may be associated with haemochromatosis, hyperparathyroidism, Wilson's disease or alkaptonuria.

Diagnosis

The diagnosis is made by detecting rhomboidal, weakly positively birefringent crystals in joint fluid, or deduced from the presence of chondrocalcinosis on X-ray. The joint fluid looks purulent. Septic arthritis must be excluded and joint fluid should be sent for culture.

The attacks may be associated with fever and a raised white blood cell count.

Treatment

Aspiration of the joint reduces the pain dramatically but it is usually necessary to use an NSAID or colchicine, as for gout. If infection can be excluded, an intra-articular injection of a corticosteroid helps.

INFECTIONS OF JOINTS AND BONES

Joints become infected by direct injury or by blood-borne infection from an infected skin lesion or other site.

Chronically inflamed joints (e.g. in rheumatoid arthritis) are more prone to infection than are normal joints. Individuals who are immunosuppressed, by AIDS or by immunosuppressive agents, are particularly at risk, as are infants, the elderly and those who abuse alcohol. Artificial joints are also potential sites for infection.

SEPTIC ARTHRITIS

The organism that most commonly causes septic arthritis is *Staphylococcus aureus*. Other organisms include streptococci, other species of staphylococcus, *Neisseria gonorrhoeae, Haemophilus influenzae* in children, and these and other Gram-negative organisms in the elderly or complicating RA.

Clinical features

Suspected septic arthritis is a medical emergency. In young and previously fit people, the joint is hot, red, swollen, and agonizingly painful and held immobile by muscle spasm. In the elderly and immunosuppressed and in RA the clinical picture is less dramatic, so a high index of suspicion is needed to avoid missing treatable but potentially severely destructive and occasionally fatal septic arthritis. In 20% the sepsis affects more than one joint. Chronic destructive arthritis due to tuberculosis is rare.

Investigations

- **Aspirate** the joint and send the fluid for urgent Gram staining and culture. The fluid is usually frankly purulent. The culture techniques should include those for gonococci and anaerobes.
- **Blood cultures** are often positive.
- **Leucocytosis** is usual, unless the person is severely immunosuppressed.
- **X-rays** are of no value in diagnosis in acute septic arthritis.
- **Skin wound swabs, sputum and throat swab or urine** may be positive and indicate the source of infection.

Treatment

This should be started immediately on diagnosis because joint destruction occurs in days. The joint should be immobilized initially and then physiotherapy started early to prevent stiffness and muscle wasting. Intravenous antibiotics should be given for 1–2 weeks. It is usual to give two antibiotics to which the organism is sensitive for 6 weeks, then one for a further 6 weeks orally. Monitor clinically and with the ESR and CRP.

Empirical treatment in septic arthritis

This is started before the results of culture are obtained. Discuss the case with a microbiologist. Intravenous flucloxacillin 1–2 g is given 6-hourly, plus fusidic acid 500 mg orally 8-hourly. If the patient is allergic to penicillin, replace flucloxacillin with erythromycin 1 g i.v. 6-hourly or clindamycin 600 mg i.v. 8-hourly. In immunosuppressed patients, flucloxacillin 1–2 g i.v. 6-hourly plus gentamicin (to cover Gram-negative organisms) should be used. Change the antibiotics if the organism is not sensitive. Drainage of the joint and arthroscopic joint washouts are helpful in relieving pain.

Management of infected prostheses

If chronically infected, the prosthesis is removed and the joint space filled with an antibiotic-impregnated spacer for 3–6 weeks before a new prosthesis is inserted. The whole process is covered by antibiotics.

Prognosis

Surgical drainage may be required if there is joint destruction and osteomyelitis. Patients can start weight-bearing as soon as the inflammation subsides. Resolution of the septic arthritis with complete recovery can occur in a few days or weeks. It may lead to secondary osteoarthritis (see p. 520).

FURTHER READING

Moreland LW. Treatment for hyperuricaemia and gout. *New England Journal of Medicine* 2005; **353**: 2505–2506.

SPECIFIC TYPES OF BACTERIAL ARTHRITIS

Gonococcal arthritis

This is the most common cause of a septic arthritis in previously fit young adults, more commonly affecting women and homosexual men.

Initially the patient becomes febrile and develops characteristic pustules on the distal limbs. Polyarthralgia and tenosynovitis are common at this stage and about 40% have a gonococcaemia. This phase settles and blood cultures usually become negative. Later, large-joint mono- or pauciarticular arthritis may follow. Culture is usually positive from the genital tract, although the joint fluid may be sterile. It is not clear whether this is simply a septic arthritis – although it responds rapidly to antibiotics – or whether there is also a reactive element to bacterial lipopolysaccharide.

Treatment consists of oral penicillin, ciprofloxacin or doxycycline for 2 weeks, and joint rest.

Tuberculous arthritis

Around 1% of patients with tuberculosis develop joint and/or bone involvement. It occurs as the primary disease in children. In adults, it is usually due to haematogenous spread from secondary pulmonary or renal lesions. The onset is insidious and diagnosis often delayed.

The organism invades the synovium or intervertebral disc. There are caseating granulomas and rapid destruction of cartilage and adjacent bone. Some patients develop a reactive polyarthritis (Poncet's disease).

The hip or knee (30%) is quite commonly affected, but around 50% develop spinal disease. The patient is febrile, has night sweats, is anorexic and loses weight. The usual risk factors for tuberculosis apply – debility, alcohol abuse or immunosuppression. HIV-positive/AIDS patients are at particular risk.

Investigations should include culture of fluid, and culture and biopsy of the synovium. *M. tuberculosis* is the usual organism, but atypical mycobacteria are occasionally implicated. A chest X-ray should be performed. Initially joint or spinal X-rays may be normal but joint-space reduction and bone destruction develop rapidly if treatment is delayed. MRI shows the abnormality earlier in the spine and CT-guided biopsy from the affected disc is necessary to obtain cultures.

Treatment is as for tuberculosis with therapy for 9 months (see p. 866). The joint should be rested and the spine immobilized in the acute phase.

Meningococcal arthritis

This may complicate a meningococcal septicaemia and presents as a migratory polyarthritis. Organisms can only rarely be cultured from the joint and most cases are due to immune complex deposition. Treatment is with penicillin.

Infective endocarditis

This may present with arthralgia, polymyalgia rheumatica-like symptoms or an infective arthritis. It is discussed on page 769.

Lyme arthritis

A person with Lyme disease (see p. 139) develops a fever and headache, and an expanding, erythematous rash called erythema chronicum migrans. About 25% of cases develop an acute pauciarticular arthritis. This usually resolves but 10% of untreated cases go on to develop a chronic arthritis. There are no positive markers in these patients of an ongoing infection (p. 139).

Diagnosis is by the detection of IgM antibodies against the spirochaete *Borrelia burgdorferi*.

Treatment with antibiotics (amoxicillin or doxycycline) is highly effective in early disease. The response of chronic arthritis to antibiotic treatment may be delayed for months.

Brucellosis

Brucellosis (see p. 138) has a world-wide distribution. The most common cause of chronic brucellosis and of arthritis is *Brucella melitensis*. There is usually a peripheral mono- or oligoarticular arthritis, which may be septic or reactive. Arthritis is more common in chronic infections of more than 6 months.

Syphilitic arthritis

Congenital syphilis (see p. 178) can cause an acute painful epiphysitis or osteochondritis sometimes associated with para-articular swelling in the first few weeks of life. Later, at age 8–16 years, painless effusion of the knees may occur (Clutton's joints).

In acquired syphilis, arthralgia and arthritis occur in the secondary stage. Charcot's (neuropathic) joints usually involve the knees in tabes dorsalis (see p. 555).

Leprosy (p. 140)

Symmetrical polyarthritis, tenosynovitis and thickened nerves with or without cutaneous manifestations are seen in leprosy.

Actinomycetes infection

Actinomycetes (see p. 149) can affect the mandible or vertebrae.

ARTHRITIS IN VIRAL DISEASE

A transient polyarthritis or arthralgia can occur before, during or after many viral illnesses. These include infectious mononucleosis, chickenpox, mumps, adenovirus, rubella, erythro(parvo)virus B19, hepatitis B and C, arboviral infections and HIV. In most of these it is due to a direct toxic effect or immune complex deposition.

In *rubella* (see p. 113) the virus can occasionally be isolated from the joint. This arthritis is uncommon in countries where rubella vaccination is routine. It occurs most commonly in up to 50% of young adult females a few days after rubella infection (6% of men). It is a symmetrical polyarthritis involving the MCP or PIP joints most commonly, but many joints can be affected. It closely resembles rheumatoid arthritis. IgM rubella antibodies are present. It resolves within a few weeks in most cases. A mild arthritis occurs rarely 2–4 weeks after rubella vaccination.

Erythro(parvo)virus B19 (p. 109) causes an acute, self-limiting arthritis and is associated with erythema infectiosum ('slapped cheek disease').

In *hepatitis B* infection (see p. 336) a sudden symmetrical polyarticular arthritis of the small joints of the hands occurs in approximately one-third of patients, often in the prodromal

phase and mostly resolving before the onset of jaundice. Hepatitis C causes type II mixed cryoglobulinaemia (see p. 340).

Arbovirus infections (see p. 113) which are endemic in many parts of the world give rise to an arthralgia and/or arthritis. For example, the Ross River virus has caused an epidemic polyarthritis in Australia and the South Pacific; it involves the small joints of the hands and clears in 2–4 weeks. Other viral infections causing epidemic arthritis include chikungunya (p. 114) and O'nyong-nyong (p. 113).

Musculoskeletal aspects of infection with human immunodeficiency virus (HIV) and acquired immune deficiency syndrome (AIDS)

The clinical features seen in these patients are due to a number of causes such as opportunistic infections and drug therapy and are not usually caused directly by HIV. Infective arthritis seen in these immunosuppressed patients often has minimal symptoms and signs. Some of the antiviral agents cause an acute arthritis, possibly because of crystallization in the joint.

Arthralgia is common in AIDS. There is a *seronegative, predominantly lower-limb arthritis*, similar to psoriasis or Reiter's disease. *Spondylitis* also occurs but is not HLA-B27 associated. *Avascular necrosis*, possibly associated with corticosteroids or alcohol, is seen.

Non-articular diseases such as Sjögren's- and lupus-like syndromes, systemic vasculitis of the necrotizing and hypersensitivity types and myositis also occur.

FUNGAL INFECTION

Fungal infections of joints occur rarely. Bone abscesses may be seen. Destructive joint lesions can also occur with blastomycosis. A benign polyarthritis accompanied by erythema nodosum occasionally occurs in coccidioidomycosis and histoplasmosis. Culture of purulent synovial fluid and skin tests for fungi may help the diagnosis.

BONE INFECTIONS

Acute and chronic osteomyelitis

Osteomyelitis can be due either to metastatic haematogenous spread (e.g. from a boil) or to local infection. Malnutrition, debilitating disease and decreased immunity may play a part in the pathogenesis.

Staphylococcus is the organism responsible for 90% of cases of acute osteomyelitis. Other organisms include *Haemophilus influenzae* and salmonella; infection with the latter may occur as a complication of sickle cell anaemia. The classic presentation is with fever and localized pain with overlying erythema. Diagnosis and treatment within a few days carries a good prognosis. Delayed treatment leads to chronic osteomyelitis. In chronic osteomyelitis sinus formation is usual. Subacute osteomyelitis is associated with a chronic abscess within the bone (Brodie's abscess). Symptoms may be limited to local pain.

Treatment of osteomyelitis is with immobilization and antibiotic therapy with flucloxacillin and fusidic acid. Surgical drainage and removal of dead bone (sequestrum) may be possible but recurrence is common.

Tuberculous osteomyelitis

This is usually due to haematogenous spread from a reactivated primary focus in the lungs or gastrointestinal tract. The disease starts in intra-articular bone. The spine is commonly involved (Pott's disease), with damage to the bodies of two neighbouring vertebrae leading to vertebral collapse and acute angulation of the spine (gibbus). Later an abscess forms ('cold abscess'). Pus can track along tissue planes and discharge at a point far from the affected vertebrae. Symptoms consist of local pain and later swelling if pus has collected. Systemic symptoms of malaise, fever and night sweats occur.

Treatment is as for pulmonary tuberculosis but extended to 9 months (see p. 866) together with initial immobilization.

AUTOIMMUNE RHEUMATIC DISEASES

AUTOIMMUNITY AND AUTOANTIBODIES

Autoimmune diseases are conditions in which the immune system attacks tissues of the body. The antigens can be present in multiple organs so the clinical manifestations are systemic and diverse. In some diseases, such as Graves' disease, Hashimoto's thyroiditis, and insulin-dependent diabetes mellitus only a single organ is affected. The term 'autoimmune rheumatic disease' (ARD) is preferable to the older term 'connective tissue disease' because the clinical effects of ARD are not limited to connective tissues. Each individual ARD has a characteristic pattern of symptoms and signs, which are used to make the diagnosis. In some ARD there are also characteristic autoantibodies (i.e. antibodies that recognize antigens which are normal constituents of the body, such as DNA and phospholipids). Positive blood tests for autoantibodies are useful but not essential in the diagnosis of ARD (see Table 10.2 and p. 504).

SYSTEMIC LUPUS ERYTHEMATOSUS (SLE)

SLE is an inflammatory, multisystem disorder with arthralgia and rashes as the most common clinical features, and cerebral and renal disease as the most serious problems.

Epidemiology

SLE occurs world-wide but the prevalence varies from country to country, with the most common prevalence of 1:250 being in African American women. In other populations the prevalence varies between 1:1000 and 1:10 000. It is about nine times as common in women as in men, with a peak age of onset between 20 and 40 years.

Aetiology

The cause is unknown but there are several predisposing factors:

- *Heredity*. There is a higher concordance rate in monozygotic twins (up to 25%) compared to dizygotic twins (3%). First-degree relatives have a 3% chance of developing the disease, but approximately 20% have autoantibodies.

Ask the authors
www.studentconsult.com

- *Genetics*. Eight different chromosomal areas have been identified containing genes linked to development of SLE. These include some HLA genes, especially A1, B8 and DR3. Homozygous deficiencies of the complement genes C1q, C2 or C4 convey a high risk of developing SLE.
- *Sex hormone status*. Premenopausal women are most frequently affected. In addition, SLE has been seen in males with Klinefelter's syndrome (XXY) (see p. 40). Treatment with hormone replacement therapy increases the risk of a flare of SLE.
- *Drugs* such as hydralazine, isoniazid, procainamide and penicillamine can induce a form of SLE which is usually mild in that kidneys and the CNS are not affected.
- *Ultraviolet light* can trigger flares of SLE, especially in the skin.
- *Exposure to Epstein–Barr virus* has been suggested as a trigger for SLE.

Pathogenesis

When cells die by apoptosis the cellular remnants appear on the cell surface as small blebs which carry self-antigens. These antigens include nuclear constituents (e.g. DNA and histones) which are normally hidden from the immune system. In patients with SLE, removal of these blebs by phagocytes is inefficient so that they are transferred to lymphoid tissues where they can be taken up by antigen-presenting cells. The self-antigens from these blebs can then be presented to T cells which in turn stimulate B cells to produce autoantibodies directed against these antigens (p. 74). It has been shown that in some patients the autoantibodies are present in stored blood samples that were taken years before the patient developed clinical features of SLE. The combination of availability of self-antigens and failure of the immune system to inactivate B cells and T cells which recognize these self-antigens (i.e. a breakdown of *tolerance* – see Ch. 3, p. 74) leads to the following immunological consequences.

- Development of autoantibodies that either form circulating complexes or deposits by binding directly to tissues.
- This leads to activation of complement and influx of neutrophils causing inflammation in those tissues.
- Abnormal cytokine production – increased blood levels of IL-10 and alpha-interferon are particularly closely linked to high activity of inflammation in SLE.

Pathology

SLE of the skin and kidneys is characterized by deposition of complement and IgG antibodies and influx of neutrophils and lymphocytes. Biopsies of other tissues are carried out less frequently but can show vasculitis affecting capillaries, arterioles and venules. The synovium of joints can be oedematous and may contain immune complexes. Haematoxylin bodies (rounded blue homogeneous haematoxylin-stained deposits) are seen in inflammatory infiltrates and are thought to result from the interaction of antinuclear antibodies and cell nuclei.

The pathology of lesions in other organs is described in the appropriate chapters.

Clinical features

The manifestations of SLE vary greatly between patients. Most patients suffer fatigue, arthralgia and/or skin problems. Involvement of major organs is less common but more serious (Fig. 10.26).

General features

Fever is common in exacerbations. Patients complain of marked malaise and tiredness and these symptoms do not correlate with disease activity or severity of organ-based complications.

The joints and muscles

Joint involvement is the most common clinical feature (≥90%). Patients often present with symptoms resembling RA with symmetrical small joint arthralgia. Joints are painful but characteristically appear clinically normal, although sometimes there is slight soft-tissue swelling surrounding the joint. Deformity because of joint capsule and tendon contraction is rare, as are bony erosions. Rarely, major joint deformity resembling RA (known as Jaccoud's arthropathy) is seen. Aseptic necrosis affecting the hip or knee is a rare complication of the disease or of treatment with corticosteroids.

Myalgia is present in up to 50% of patients but a true myositis is seen only in <5%. If myositis is prominent, the patient may well have an overlap ARD with both polymyositis and SLE (see section on 'Overlap syndromes', p. 548).

The skin (see p. 1252)

This is affected in 85% of cases. Erythema, in a 'butterfly' distribution on the cheeks of the face and across the bridge of the nose (see Fig. 23.26), is characteristic. Vasculitic

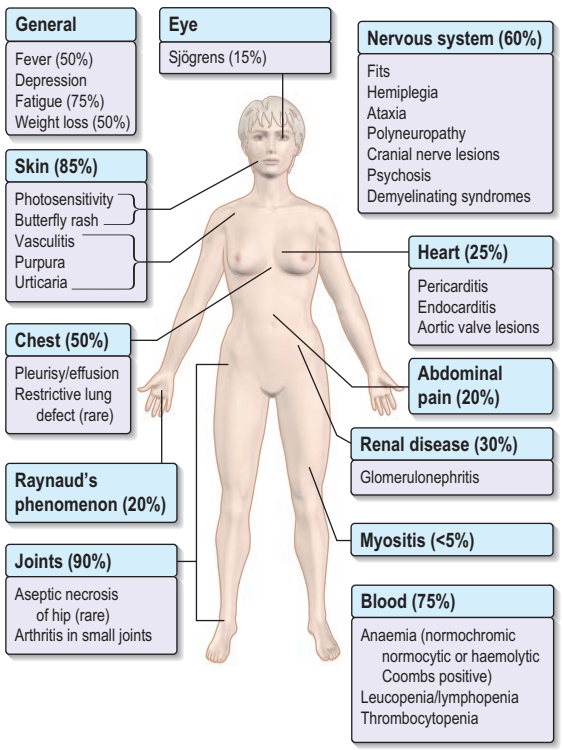


Fig. 10.26 Clinical features of systemic lupus erythematosus (SLE).

lesions on the finger tips and around the nail folds, purpura and urticaria occur. In 40–50% of cases there is photosensitivity (especially in patients positive for anti-Ro antibodies). Prolonged exposure to sunlight can lead to exacerbations of the disease. Livedo reticularis, palmar and plantar rashes, pigmentation and alopecia are seen. Scarring alopecia can lead to irreversible bald patches which are especially upsetting for women – who form the majority of patients with SLE. Raynaud's phenomenon (see p. 809) is common and may precede the development of other clinical problems by years.

Discoid lupus is a benign variant of lupus in which only the skin is involved. The rash is characteristic and appears on the face as well-defined erythematous plaques that progress to scarring and pigmentation (see p. 1252). Subacute cutaneous lupus erythematosus, a rare variant, is described on page 1252.

The lungs (see p. 871)

Up to 50% of patients will have lung involvement sometime during the course of the disease. Recurrent pleurisy and pleural effusions (exudates) are the most common manifestations and are often bilateral. Pneumonitis and atelectasis may be seen; eventually a restrictive lung defect develops with loss of lung volumes and raised hemidiaphragms. This 'shrinking lung syndrome' is poorly understood but may have a neuromuscular basis. Rarely, pulmonary fibrosis occurs, more commonly in overlap syndromes. Intrapulmonary haemorrhage associated with vasculitis is a rare but potentially life-threatening complication.

The heart and cardiovascular system

The heart is involved in 25% of cases. Pericarditis, with small pericardial effusions detected by echocardiography, is common. A mild myocarditis also occurs, giving rise to arrhythmias. Aortic valve lesions and a cardiomyopathy can rarely be present. A non-infective endocarditis involving the mitral valve (Libman–Sacks syndrome) is very rare. Raynaud's, vasculitis, arterial and venous thromboses can occur, especially in association with the antiphospholipid syndrome (see below). There is an increased frequency of ischaemic heart disease and stroke in patients with SLE. This is partly due to altered levels of common risk factors such as hypertension and lipid levels but the presence of chronic inflammation over many years may also play a role. It is not known whether intensive treatment of cardiovascular risk factors in SLE will alter the risk of developing coronary disease or stroke. The benefit of statin therapy in the absence of significant hypercholesterolaemia remains to be proved.

The kidneys

A classification of types of nephritis is on page 590. Autopsy studies suggest that histological changes are very frequent, but clinical renal involvement occurs in only approximately 30% of cases. All patients should have regular screening of urine for blood and protein. An asymptomatic patient with proteinuria may be in the early stages of lupus nephritis, and treatment may prevent progression to renal impairment. Proteinuria should be quantified and haematuria should prompt examination for urinary casts or fragmented red cells that suggest glomerulonephritis. Renal vein thrombosis can occur in nephrotic syndrome or associated with antiphospholipid antibodies.

The nervous system

Involvement of the nervous system occurs in up to 60% of cases and symptoms often fluctuate. There may be a mild depression but occasionally more severe psychiatric disturbances occur. Epilepsy, migraines, cerebellar ataxia, aseptic meningitis, cranial nerve lesions, cerebrovascular disease or a polyneuropathy may be seen. The pathogenic mechanism for cerebral lupus is complex. Lesions may be due to vasculitis or immune-complex deposition, thrombosis or non-inflammatory microvasculopathy. The commonest finding on MRI scan is of increased white matter signal abnormality. In patients with cerebral lupus, infection should be excluded or treated in parallel with administration of corticosteroids and immunosuppression.

The eyes

Retinal vasculitis can cause infarcts (cytoid bodies) which appear as hard exudates, and haemorrhages. There may be episcleritis, conjunctivitis or optic neuritis, but blindness is uncommon. Secondary Sjögren's syndrome is seen in about 15% of cases.

The gastrointestinal system

Mouth ulcers are common and may be a presenting feature. These may be painless or become secondarily infected and painful. Mesenteric vasculitis can produce inflammatory lesions involving the small bowel (infarction or perforation). Liver involvement and pancreatitis are uncommon.

Investigations

- Blood:
  - *A full blood count* may show a leucopenia, lymphopenia and/or thrombocytopenia. Anaemia of chronic disease or autoimmune haemolytic anaemia also occurs. The ESR is raised in proportion to the disease activity. In contrast, the CRP is usually normal but may be high when the patient has lupus pleuritis or arthritis or a coexistent infection.
  - *Urea and creatinine* only rise when renal disease is advanced. Low serum albumin or high urine protein/creatinine ratio are earlier indicators of lupus nephritis.
  - *Autoantibodies* – many different autoantibodies may be present in SLE but the most significant are ANA, anti-dsDNA, anti-Ro, anti-Sm and anti-La (see Box 10.11). Antiphospholipid antibodies are present in 25–40% of cases but not all of these patients develop antiphospholipid syndrome (see below).
  - *Serum complement* C3 and C4 levels are often reduced during active disease. The combination of high ESR, high anti-dsDNA and low C3 may herald a flare of disease. All these markers tend to return towards normal as the flare improves but in some patients anti-dsDNA levels remain high even during clinical remission.
- **Histology**. Characteristic histological and immunofluorescent abnormalities (deposition of IgG and complement) are seen in biopsies from the kidney and skin.
- **Diagnostic imaging**. CT scans of the brain sometimes show infarcts or haemorrhage with evidence of cerebral atrophy. MR can detect lesions in white matter which are not seen on CT. However, it can be very difficult to distinguish true vasculitis from small thrombi.

Box 10.11 Autoantibodies and disease associations

Antibody	Disease	Prevalence
ds DNA	SLE	70%
Anti-histone	Drug induced lupus	–
Anti-centromeric	Limited SS	70%
Anti-Ro (SS-A)	SLE	40–60%
	Primary Sjögren's	60–90%
Anti-La (SS-B)	SLE	15%
	Primary Sjögren's	35–85%
Anti-Sm	SLE	10–25% (Caucasian)
		30–50% (Black African)
Anti-UI-RNP	SLE	30%
	Overlap syndrome	
Anti-JO (anti-synthetase)	Polymyositis	30%
	Dermatomyositis	
Anti-topoisomerase 1 (Scl-70)	Diffuse cutaneous SSc	30%

SS-A, SS-B, Sjögren's syndrome -A and –B; -Ro, -La -, first two letters of name of patients; -Sm, Smith – patient name; RNP, ribonucleoprotein; SSc, systemic scleroderma.

Management

General measures

The disease and its management should be discussed with the patient, particularly the effect upon the patient's lifestyle, e.g. appearance and debility due to fatigue. Patients are advised to avoid excessive exposure to sunlight and it is also necessary to reduce cardiovascular risk factors (p. 755).

Symptomatic treatment

Many patients do not need treatment with corticosteroid tablets or immunosuppressive agents. Arthralgia, arthritis, fever and serositis all respond well to standard doses of NSAIDs (p. 517). Topical corticosteroids are effective and widely used in cutaneous lupus. Antimalarial drugs (chloroquine or hydroxychloroquine) help mild skin disease, fatigue and arthralgias that cannot be controlled with NSAIDs but patients require regular eye checks because of rare retinal toxicity (1 in 2000).

Corticosteroids and immunosuppressive drugs

Single intramuscular injections of long-acting corticosteroids or short courses of oral corticosteroids are useful in treating severe flares of arthritis, pleuritis or pericarditis. In some cases, these symptoms can only be kept under control using long-term oral corticosteroids.

Renal (p. 590) or cerebral disease and severe haemolytic anaemia or thrombocytopenia must be treated with high-dose oral corticosteroids and the first two of these require immunosuppressive drugs in addition. Cyclophosphamide was most commonly used to achieve remission in these severe forms of lupus but is being replaced by mycophenolate mofetil, which has fewer side-effects. Azathioprine is also used to maintain remission. Newer agents such as rituximab (anti-CD20) are used in refractory cases. Rituximab depletes CD20 positive B lymphocytes and thus reduces levels of autoantibodies.

Course and prognosis

An episodic course is characteristic, with exacerbations and complete remissions that may last for long periods. However, SLE can also be a chronic persistent condition. The mortality rate in SLE has fallen dramatically over the last 50 years; the 10-year survival rate is about 90%, but this is lower if major organ-based complications are present. Deaths early in the course of disease are mainly due to renal or cerebral disease or infection. Later coronary artery disease and stroke become more prevalent. Chronic progressive destruction of joints as seen in RA and OA occurs rarely, but a few patients develop deformities such as ulnar deviation. Patients with SLE have an increased long-term risk of developing some cancers, especially lymphoma.

Pregnancy and SLE

Fertility is usually normal except in severe disease and there is no major contraindication to pregnancy. Recurrent miscarriages can occur, especially in patients with antiphospholipid antibodies. Exacerbations can occur during pregnancy with frequent exacerbations of the disease postpartum. The patient's usual treatment should be continued during pregnancy. Hypertension must be controlled. Patients with anti-Ro or anti-La antibodies have a 2% risk of giving birth to babies with neonatal lupus syndrome (rash, hepatitis and fetal heart block).

ANTIPHOSPHOLIPID SYNDROME

Patients who have thrombosis (arterial or venous) and/or recurrent miscarriages and who also have persistently positive blood tests for antiphospholipid antibodies (aPL) have the antiphospholipid syndrome (APS). aPL can be detected by several different tests:

- The *anticardiolipin test*, which detects antibodies (IgG or IgM) that bind the negatively charged phospholipid, cardiolipin.
- The *lupus anticoagulant test* which detects changes in the ability of blood to clot in a test tube. Despite the name, this is not a test for lupus. It is a test for APS. The *anticoagulant* effect caused by aPL in the test tube causes an opposite *procoagulant* effect inside the body, because the factors stimulating thrombosis are different (see below).
- The *anti-β_2-glycoprotein I* test, which detects antibodies that bind β_2-glycoprotein I, a molecule that interacts closely with phospholipids.

A persistently positive test (i.e. positive on at least two occasions, 6 weeks or more apart) in one or more of these assays is needed to diagnose APS. However, some people who test positive for aPL will never get APS, i.e. not all aPLs are harmful. APS can present in patients who already have another ARD, especially SLE. APS can also occur on its own (primary APS).

Pathogenesis

Negatively charged phospholipids and β_2-glycoprotein I are present on the outer surface of apoptotic blebs and so aPL are believed to arise by a similar mechanism to the lupus autoantibodies described above. Pathogenic aPL bind to the N-terminal domain of β_2-glycoprotein I and this interaction is facilitated when the protein is bound to

phospholipid on the surface of cells such as endothelial cells, platelets, monocytes and trophoblasts. This alters the functioning of those cells leading to thrombosis and/or miscarriage.

Clinical features

Since APS is defined by the presence of thrombosis and/or pregnancy loss it is not surprising that these are the most common features. Ischaemic strokes occur in about 20% of patients and deep vein thrombosis in about 40%. Unlike most causes of thrombophilia, APS can cause either arterial or venous thrombosis (though rarely both in the same patient). Twenty-seven per cent of women who have had two or more spontaneous miscarriages have APS.

However, large studies show that patients with APS can also have many other features including:

- thrombocytopenia
- chorea, migraine and epilepsy
- valvular heart disease
- cutaneous manifestations (e.g. livedo reticularis)
- positive Coombs' test
- renal impairment due to ischaemia in the small renal vessels.

Rarely it can be catastrophic. The syndrome may also be involved in the development of accelerated atheroma.

Treatment

In patients with APS who have had one or more thrombosis, the recommended treatment to prevent further thrombosis is long-term anticoagulation with warfarin. Although it was previously thought that high-intensity anticoagulation (INR >3) was essential, this is now disputed and many patients are managed with lower target INR. Pregnant patients with APS are given oral aspirin and subcutaneous heparin from early in gestation. This reduces the chance of a miscarriage but pre-eclampsia and poor fetal growth remain common. There are no definite guidelines for managing patients with aPL who have never had thrombosis. Aspirin or clopidogrel are sometimes given prophylactically especially in those with high IgG aPL. Warfarin is given much more rarely in these circumstances.

SYSTEMIC SCLEROSIS (SCLERODERMA)

(see p. 1252)

Systemic sclerosis (SSc) is a multisystem disease. This distinguishes it from localized scleroderma syndromes, such as morphoea, that do not involve internal organ disease and are rarely associated with vasospasm (Raynaud's phenomenon). SSc has the highest case-specific mortality of any of the autoimmune rheumatic diseases. SSc occurs world-wide but there may be racial or ethnic differences in clinical features. For example, renal involvement is less frequent in Japanese cases.

The incidence of SSc is 10/million population per year with a 3:1 female to male ratio. The peak incidence is between 30 and 50 years of age. It is rare in children.

Environmental risk factors for scleroderma-like disorders include exposure to vinyl chloride, silica dust, adulterated rapeseed oil and trichlorethylene. Drugs such as bleomycin also produce a similar picture. Although unusual, familial cases are reported and twin cohorts suggest higher concordance in monozygotic pairs, consistent with genetic determinants of aetiology.

Pathology and pathogenesis

Figure 10.27 shows the pathogenic interaction and mechanisms in systemic sclerosis.

Vascular features

An early lesion is widespread vascular damage involving small arteries, arterioles and capillaries. There is initial endothelial cell damage with release of cytokines including endothelin-1, which causes vasoconstriction. There is continued intimal damage with increasing vascular permeability, leading to cellular activation, activation of adhesion molecules (E-selectin, VCAM, ICAM-1), with migration of cells into the extracellular matrix. Migrating lymphocytes are IL-2-producing cells, expressing surface antigens such as CD3, CD4 and CD5. All these factors cause release of other mediators (e.g. interleukin-1, -4, -6 and -8, TGF-β and PDGF) with activation of fibroblasts.

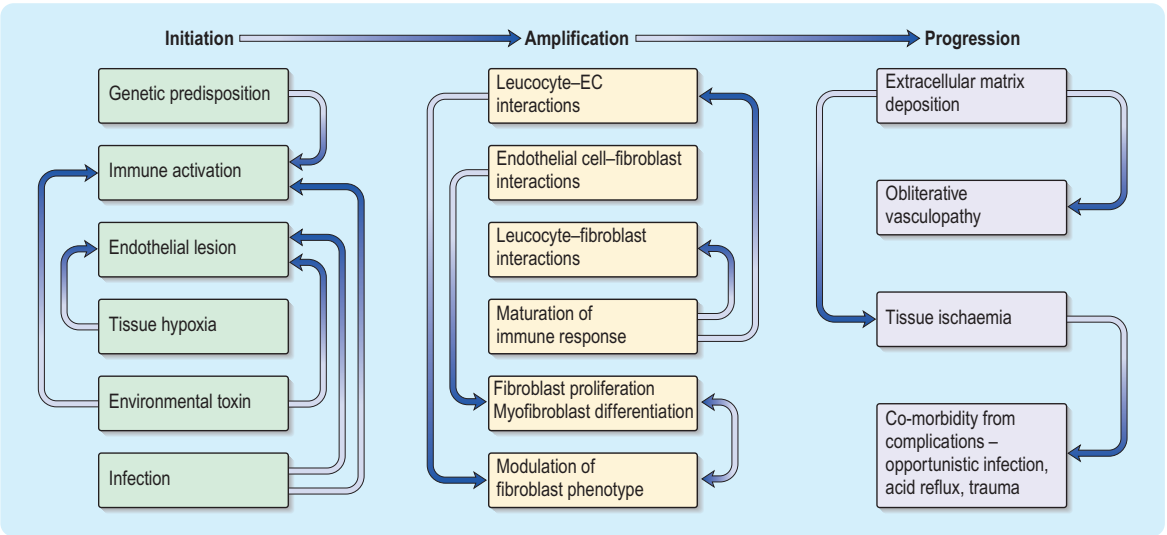


Fig. 10.27 **Pathogenic mechanisms in systemic sclerosis.** EC, endothelial cell.

The damage to small blood vessels also produces widespread obliterative arterial lesions and subsequent chronic ischaemia.

Fibrotic features

Fibroblasts synthesize increased quantities of collagen types I and III, as well as fibronectin and glycosaminoglycans, producing fibrosis in the lower dermis of the skin as well as the internal organs. It is possible that antibodies to platelet derived growth factor receptor, which have been found in blood of patients with SSc, stimulate fibroblasts to cause fibrosis.

Clinical features

Raynaud's phenomenon

Raynaud's phenomenon is seen in almost 100% of cases and can precede the onset of the full-blown disease by many years.

Limited cutaneous scleroderma (LcSSc) – 70% of cases

This usually starts with Raynaud's phenomenon many years (up to 15) before any skin changes. The skin involvement is limited to the hands, face, feet and forearms. The skin is tight over the fingers and often produces flexion deformities of the fingers. Involvement of the skin of the face produces a characteristic 'beak'-like nose and a small mouth (microstomia). Painful digital ulcers and telangiectasia with dilated nail-fold capillary loops are seen. Digital ischaemia may lead to gangrene. Gastrointestinal tract involvement is common in this group. Pulmonary hypertension (PHT) develops in 10–15% of this group and pulmonary interstitial disease may occur.

The CREST syndrome (Calcinosis, Raynaud's phenomenon, Esophageal involvement, Sclerodactyly and skin changes in the fingers, Telangiectasia) was the term previously used to describe this syndrome.

Diffuse cutaneous scleroderma (DcSSc) – 30% of cases

Initially oedematous in onset, skin sclerosis rapidly follows. Raynaud's phenomenon usually starts just before or concomitant with the oedema.

Diffuse swelling and stiffness of the fingers is rapidly followed by more extensive skin thickening which can involve most of the body in the severest cases. Later the skin becomes atrophic. Early involvement of other organs occurs with general symptoms of lethargy, anorexia and weight loss.

- Heartburn, reflux or dysphagia due to oesophageal involvement is almost invariable and anal incontinence occurs in many patients. Malabsorption from bacterial overgrowth due to dilatation and atony of the small bowel is not infrequent, and more rarely dilatation and atony of the colon occurs. Pseudo-obstruction is a known complication.
- Renal involvement is acute or chronic. Acute hypertensive renal crisis used to be the most common cause of death in systemic sclerosis. ACE inhibitors and better care along with dialysis and renal transplantation have changed this.
- Lung disease, both fibrosis and pulmonary hypertension, contributes significantly to mortality in SSc. PHT can be isolated or secondary to fibrosis, and high plasma levels of endothelin-1 are seen.
- Myocardial fibrosis leads to arrhythmias and conduction defects. Pericarditis is found occasionally.

Sometimes these systemic features occur without skin involvement (SSc sine scleroderma). Overlap syndromes with additional features of SLE, RA or inflammatory muscle disorder occur.

Investigations

- **Full blood count**. A normochromic, normocytic anaemia occurs and a microangiopathic haemolytic anaemia is seen in some patients with renal disease.
- **Urea and electrolytes**. Urea and creatinine rise with acute renal disease
- **Autoantibodies** (see Box 10.11):
  - **(a)** *In LcSSc*: speckled, nucleolar or anti-centromere antibodies (ACAs) occur in 70% of cases.
  - **(b)** *In DcSSc*: there are anti-topoisomerase-1 antibodies (called anti-ScL-70) in 30% of cases, and anti-RNA polymerase (I, II and III) antibodies in 20–25%.
  - **(c)** *Rheumatoid factor* is positive in 30%.
  - **(d)** *ANA* is positive in 95%.
- **Urine** microscopy and, if there is proteinuria, urine protein/creatinine ratio should be measured.
- **Imaging**:
  - **(a)** CXR – to exclude other pathology, for changes in cardiac size and established lung disease.
  - **(b)** Hands – deposits of calcium around fingers (in severe cases, erosion and absorption of the tufts of the distal phalanges – termed acrosteolysis).
  - **(c)** Barium swallow generally confirms impaired oesophageal motility. Scintigraphy, manometry impedance, and upper GI endoscopy are also valuable.
  - **(d)** High-resolution CT – to demonstrate fibrotic lung involvement.
- **Other investigations** of gastrointestinal tract (e.g. see Fig. 6.5, p. 250), lung, renal and cardiac as appropriate.

Management

Treatment should be organ-based in order to try to control the disease. Currently there is no cure. In contrast to many other ARD, corticosteroids and immunosuppressants are rarely used in SSc, with the exception of SSc-related pulmonary fibrosis.

- Education, counselling and family support are essential.
- Regular exercises and skin lubricants may limit contractures but no treatment has proven efficacy in reducing skin fibrosis.
- Raynaud's may be improved by hand warmers and oral vasodilators (calcium-channel blockers, ACE inhibitors). In severe cases, parenteral vasodilators (prostacyclin analogues and calcitonin gene-related peptide) are used. Lumbar sympathectomy can help foot symptoms. Radical micro-arteriolysis (digital sympathectomy) can be used where individual fingers or toes are severely ischaemic and thoracic sympathectomy under video assisted thoracic surgery is now performed.
- Oesophageal symptoms can almost always be improved by proton-pump inhibitors, and prokinetic drugs are also helpful.
- Symptomatic malabsorption requires nutritional supplements and rotational antibiotics (see p. 280) to treat small intestinal bacterial overgrowth.

- Renal involvement requires intensive control of hypertension. First drug of choice is an ACE inhibitor. Vigilance for hypertensive scleroderma renal crisis (SRC) is critical, especially in early-stage dcSSc with rapidly progressive skin and tendon friction rubs. High-dose corticosteroids (above 10 mg prednisolone daily) may increase the risk of SRC.
- Pulmonary hypertension is treated with oral vasodilators, oxygen and warfarin. Advanced cases should receive prostacyclin therapy (inhaled, subcutaneous or intravenous) or the oral endothelin-receptor antagonists (bosentan and sitaxentan). Right heart failure is treated conventionally and transplantation (heart–lung or single lung) should be considered in eligible cases.
- Pulmonary fibrosis is currently treated with immunosuppression, most often with cyclophosphamide or azathioprine combined with low-dose oral prednisolone.

Prognosis

In limited cutaneous scleroderma the disease is often milder, with much less severe internal organ involvement and a 70% 10-year survival. Pulmonary hypertension is a significant later cause of death. Lung fibrosis and severe gut involvement also determine mortality. In diffuse disease, where organ involvement is often severe at an earlier stage, many patients die of pulmonary, cardiac or renal involvement. Overall, pulmonary involvement (vascular or interstitial) accounts for around 50% of scleroderma-related deaths.

Localized forms of scleroderma occur either in patches (morphoea, p. 1252) or linear forms. These are more commonly seen in children and adolescents and do not convert into systemic forms, although ANA may occur in localized scleroderma and very occasionally there is coexistence of localized and systemic forms.

POLYMYOSITIS (PM) AND DERMATOMYOSITIS (DM)

Polymyositis is a rare disorder of unknown cause, in which the clinical picture is dominated by inflammation of striated muscle, causing proximal muscle weakness. When the skin is involved it is called 'dermatomyositis'. The incidence is about 2–10/million population per annum and it occurs in all races and at all ages. The aetiology is unknown, although viruses (e.g. Coxsackie, rubella, influenza) have been implicated and persons with HLA-B8/DR3 appear to be genetically predisposed.

Clinical features

Adult polymyositis

Women are affected three times more commonly than men.

The onset can be insidious, over months, or acute. General malaise, weight loss and fever can develop during the acute phase, but the cardinal symptom is proximal muscle weakness. The shoulder and pelvic girdle muscles may become wasted but are not usually tender. Face and distal limb muscles are not usually affected. Movements such as squatting and climbing stairs become difficult. As the disease progresses, involvement of pharyngeal, laryngeal and respiratory muscles can lead to dysphonia and respiratory failure. These severe complications are rare if the disease is treated early.

Adult dermatomyositis

This is also more common in women. Apart from muscle weakness these patients often suffer from myalgia, polyarthritis and Raynaud's phenomenon but DM is primarily distinguished from PM by the characteristic rash. This typically affects the eyelids, where heliotrope (purple) discoloration is accompanied by periorbital oedema, and the fingers where one sees purple-red raised vasculitic patches. These patches occur over the knuckles (Gottron's papules) in 70% of patients, and this appearance is highly specific for DM. Ulcerative vasculitis and calcinosis of the subcutaneous tissue occurs in 25% of cases. In the long term, muscle fibrosis and contractures of joints occur.

Other organ involvement (antisynthetase syndrome)

Twenty to thirty per cent of patients with PM or DM have antibodies to tRNA synthetase enzymes. These patients are more likely to develop pulmonary interstitial fibrosis, Raynaud's phenomenon, arthritis and hardening and fissuring of skin over the pulp surface of the fingers (mechanic's hands). This variant of PM/DM is sometimes called antisynthetase syndrome and often has a poor outcome. Respiratory muscles are affected in PM/DM and this compounds the effects of interstitial fibrosis. Dysphagia is seen in about 50% of patients owing to oesophageal muscle involvement.

Association with other ARD

There is an association with other ARD (e.g. SLE, RA and SSc) with their associated clinical features such as deforming arthritis, malar rash and skin sclerosis.

Association with malignancies

The relative risk of cancer is 2.4 for male and 3.4 for female patients, and a wide variety of cancers have been reported. The onset and clinical picture does not differ from that of typical DM/PM. The associated cancer may not become apparent for 2–3 years, and recurrent or refractory DM should prompt a search for occult malignancy.

Malignancy (e.g. lung, ovary, breast, stomach) can also predate the onset of myositis, particularly in males with DM.

Childhood dermatomyositis

This most commonly affects children between the ages of 4 and 10 years. The typical rash of DM is usually accompanied by muscle weakness. Muscle atrophy, subcutaneous calcification and contractures may be widespread and severe. Ulcerative skin vasculitis is common and recurrent abdominal pain due to vasculitis is also a feature.

Investigations

- **Serum creatine kinase (CK),** aminotransferases, lactate dehydrogenase (LDH) and aldolase are usually raised and are useful guides to muscle damage but may not reflect activity.
- The **ESR** is usually not raised.
- **Serum autoantibody studies**. Anti-nuclear antibody testing is commonly positive in patients with DM. Rheumatoid factor is present in up to 50% and many *myositis-specific* antibodies (MSAs) have been recognized and correlate with certain subsets. Anti-synthetase antibodies have been described above.

- **Electromyography (EMG)** shows a typical triad of changes with myositis: spontaneous fibrillation potentials at rest; polyphasic or short-duration potentials on voluntary contraction; and salvos of repetitive potentials on mechanical stimulation of the nerve.
- **MRI** can be used to detect abnormally inflamed muscle.
- **Needle muscle biopsy** shows fibre necrosis and regeneration in association with an inflammatory cell infiltrate with lymphocytes around the blood vessels and between muscle fibres. Open biopsy allows more thorough assessment.
- **Screening for malignancy** is usually limited to relatively non-invasive investigations such as CXR, mammography, pelvic/abdominal ultrasound, urine microscopy and a search for circulating tumour markers.
- **PET** scan.

Treatment

Bed rest may be helpful but must be combined with an exercise programme. Prednisolone is the mainstay of treatment; 0.5–1.0 mg/kg bodyweight as initial therapy continued until at least one month after myositis has become clinically and enzymatically inactive. Tapering of steroids must be slow. Early intervention with steroid-sparing agents such as methotrexate, azathioprine, ciclosporin, cyclophosphamide and mycophenolate mofetil is common, especially where there is clinical relapse or rise in CK as the dose of steroids is reduced. Intravenous immunoglobulin therapy (IVIG) is helpful in some recalcitrant cases. Treatment of childhood DM tends to be more intensive with earlier use of immunosuppressive agents.

Inclusion body myositis

Inclusion body myositis is an idiopathic inflammatory myopathy occurring usually in men over 50 years. Weakness of the pharyngeal muscles causes difficulty in swallowing in over 50%. It is a slowly progressive weakness of mainly distal muscles. In contrast to polymyositis, the creatine kinase is only slightly elevated; the EMG shows both myopathic and neuropathic changes. On MRI, the changes are often more distal but can be similar to polymyositis. A muscle biopsy shows inflammation and basophilic rimmed vacuoles with diagnostic filamentous inclusions and vacuoles on electron microscopy. A trial of corticosteroids is worthwhile but generally the response is poor.

SJÖGREN'S SYNDROME

The syndrome of dry eyes (keratoconjunctivitis sicca) in the absence of rheumatoid arthritis or any of the autoimmune diseases is known as 'primary Sjögren's syndrome'. There is an association with HLA-B8/DR3. Dryness of the mouth, skin or vagina may also be a problem. Salivary and parotid gland enlargement is seen.

Associated systemic features include:

- arthralgia and occasional non-progressive polyarthritis, like that seen in SLE (but much less common)
- Raynaud's phenomenon
- dysphagia and abnormal oesophageal motility as seen in systemic sclerosis (but less common)
- other organ-specific autoimmune disease, including thyroid disease, myasthenia gravis, primary biliary cirrhosis and autoimmune hepatitis
- renal tubular defects (uncommon) causing nephrogenic diabetes insipidus and renal tubular acidosis
- pulmonary diffusion defects and fibrosis
- polyneuropathy, fits and depression
- vasculitis
- increased incidence of non-Hodgkin's B cell lymphoma.

Pathology and investigations

Biopsies of the salivary gland or of the lip show a focal infiltration of lymphocytes and plasma cells.

- **Schirmer tear test**. A standard strip of filter paper is placed on the inside of the lower eyelid; wetting of <10 mm in 5 minutes indicates defective tear production.
- **Rose Bengal staining** of the eyes shows punctate or filamentary keratitis.
- **Laboratory abnormalities**. These include raised immunoglobulin levels, circulating immune complexes and autoantibodies. Rheumatoid factor is usually positive. Anti-nuclear antibodies are found in 80% of cases and anti-mitochondrial antibodies in 10%. Anti-Ro (SSA) antibodies are found in 60–90%, compared with 10% of cases of RA and secondary Sjögren's syndrome. This antibody is of particular interest because it can cross the placenta and cause congenital heart block.

Management

Symptomatic treatment is with artificial tears and saliva-replacement solutions. Hydroxychloroquine may help fatigue and arthralgia. Corticosteroids are rarely needed but are used to treat persistent salivary gland swelling or neuropathy.

'OVERLAP' SYNDROMES AND UNDIFFERENTIATED AUTOIMMUNE RHEUMATIC DISEASE

An *overlap syndrome* is one where the patient shows the characteristic clinical features of more than one ARD. Treatment of each ARD is usually the same as if they occurred separately.

Undifferentiated ARD is a term used for patients who have evidence of autoimmunity (e.g. positive autoantibody test) and some clinical features of such diseases (commonly Raynaud's phenomenon and/or arthralgia) but not enough to make a clear diagnosis of any individual ARD. These patients sometimes develop a clearer ARD over time, but some always remain undifferentiated and tend to have relatively mild disease without major organ problems.

SYSTEMIC INFLAMMATORY VASCULITIS

Vasculitis is a histological term describing inflammation of the vessel wall. Vasculitis can be seen in many diseases (see Table 10.18 and 10.19). The group of diseases described in this section (systemic inflammatory vasculitides) are characterized by widespread vasculitis leading to systemic symptoms and signs, generally requiring treatment with corticosteroids and/or immunosuppressive drugs. The classification remains controversial, but is usually based on the type of artery affected (Fig. 10.28 and Table 10.18).

- Large vessel vasculitis refers to the aorta and its major tributaries.
- Medium vessel vasculitis refers to medium and small-sized arteries and arterioles.
- Small vessel vasculitis refers to small arteries, arterioles, venules and capillaries.

Table 10.18 Types of systemic vasculitis

Large	Giant cell arteritis/ polymyalgia rheumatica	
	Takayasu's arteritis	
Medium	Classical polyarteritis nodosa (PAN)	
	Kawasaki's disease	
Small	Microscopic polyangiitis	ANCA associated
	Wegener's granulomatosis	ANCA associated
	Churg–Strauss syndrome	ANCA associated
	Henoch–Schönlein purpura	
	Cutaneous leucocytoclastic vasculitis	
	Essential cryoglobulinaemia	

Table 10.19 Other conditions associated with vasculitis (see also Table 10.18)

Infective	e.g. Subacute infective endocarditis
Non-infective	Vasculitis with rheumatoid arthritis
	Systemic lupus erythematosus
	Scleroderma
	Polymyositis/dermatomyositis
	Drug-induced Behçet's disease
	Goodpasture's syndrome
	Hypocomplementaemia
	Serum sickness
	Paraneoplastic syndromes
	Inflammatory bowel disease

Within this classification based on size, different types of vasculitis can be distinguished in terms of histological appearance or autoantibody status. In particular, only some vasculitides are associated with giant cells or granulomata in the vessel wall, and the presence or absence of anti-neutrophil cytoplasmic antibody (ANCA, see p. 504) is also a useful distinguishing feature.

LARGE VESSEL VASCULITIS

Polymyalgia rheumatica (PMR) and giant cell (temporal) arteritis are systemic illnesses of the elderly. Both are associated with the finding of a giant cell arteritis on temporal artery biopsy.

Polymyalgia rheumatica (PMR)

PMR causes a sudden onset of severe pain and stiffness of the shoulders and neck, and of the hips and lumbar spine; a limb girdle pattern. These symptoms are worse in the morning, lasting from 30 minutes to several hours. The clinical history is usually diagnostic and the patient is always over 50 years old.

Patients develop systemic features of tiredness, fever, weight loss, depression and occasionally nocturnal sweats if PMR is not diagnosed and treated early. A differential diagnosis is shown in Box 10.12.

Box 10.12 Symptom patterns in some muscle disorders

- **Polymyositis** – proximal muscle ache and weakness
- **Polymyalgia rheumatica** – proximal morning stiffness and pain
- **Myopathy** – weakness, but no pain or stiffness

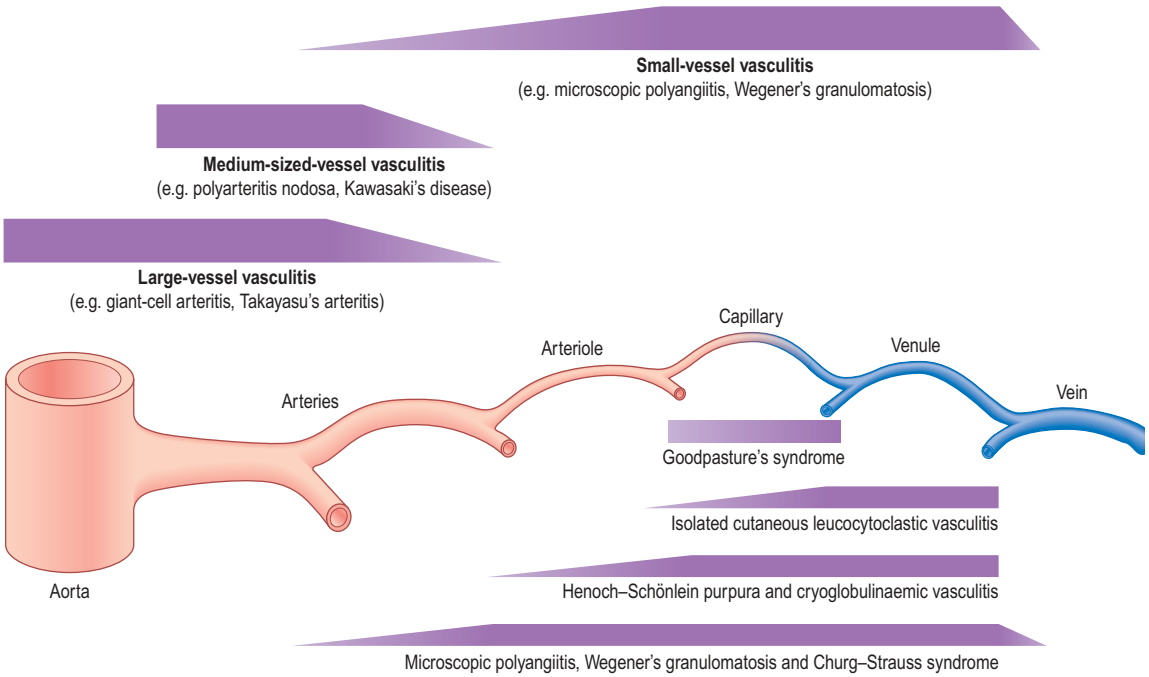


Fig. 10.28 **Sites of vascular involvement by vasculitides.** From Jeanette JC and Falk RJ. *New England Journal of Medicine* 1997; **337:** 1512–1523. Copyright © 1997 Massachusetts Medical Society. All rights reserved.

Investigation of PMR

- **A raised ESR and/or CRP** is a hallmark of this condition. It is rare to see PMR without an acute-phase response. If it is absent, the diagnosis should be questioned and the tests repeated a few weeks later before treatment is started.
- **Serum alkaline phosphatase and γ-glutamyl-transpeptidase** may be raised as markers of the acute inflammation.
- **Anaemia** (mild normochromic, normocytic) is often present.
- **Temporal artery biopsy** shows giant cell arteritis in 10–30% of cases, but is not usually performed.

Giant cell arteritis (GCA)

GCA is inflammatory granulomatous arteritis of large cerebral arteries which occurs in association with PMR. The patient may have current PMR, a history of recent PMR, or be on treatment for PMR. It is extremely rare under 50 years of age. Presenting symptoms of GCA include severe headaches, tenderness of the scalp (combing the hair may be painful) or of the temple, claudication of the jaw when eating, tenderness and swelling of one or more temporal or occipital arteries. The most feared manifestation is sudden painless temporary or permanent loss of vision in one eye due to involvement of the ophthalmic artery (see p. 1138). Systemic manifestations of severe malaise, tiredness and fever occur.

Investigation of GCA

- **Normochromic, normocytic anaemia.**
- **ESR** is usually raised (in the region of 50–120 mm/h) and the CRP very high.
- **Liver biochemistry.** Abnormalities occur, as in PMR. The albumin may be low.
- **A temporal artery biopsy** from the affected side is the definitive diagnostic test. This should be taken before, or within 7 days of starting, high doses of corticosteroids. The lesions are patchy and the whole length of the biopsy (>1 cm long) must be examined; even so, negative biopsies occur.

The histological features of GCA are:

- intimal hypertrophy
- inflammation of the intima and sub-intima
- breaking up of the internal elastic lamina
- giant cells, lymphocytes and plasma cells in the internal elastic lamina.

Treatment of PMR or GCA

Corticosteroids produce a dramatic reduction of symptoms of PMR within 24–48 hours of starting treatment, provided the dose is adequate. If this improvement does not occur, the diagnosis should be questioned. This treatment should reduce the risk of patients with PMR developing GCA. NSAIDs are less effective and should not be used.

In GCA, corticosteroids are obligatory because they significantly reduce the risk of irreversible visual loss and other focal ischaemic lesions, but much higher doses are needed than in PMR. If GCA is suspected, it may not be possible to arrange a temporal artery biopsy rapidly. In these circumstances, treatment should not be delayed, especially if there have already been episodes of visual loss or stroke.

Starting doses of prednisolone are:

- **PMR**: 10–15 mg prednisolone as a single dose in the morning
- **GCA**: 60–100 mg prednisolone, usually in divided doses.

The dose should then be reduced gradually in weekly or monthly steps. While the dose is above 20 mg, the step reductions are 5 mg, reducing the evening doses first. Between 20 mg and 10 mg the reduction can be in 2.5 mg steps, but below 10 mg the rate should be slower and the steps each of 1 mg. Most patents will eventually be able to stop corticosteroids after 12–18 months but up to 25% may need low doses long-term. Steroid-sparing immunosuppressive agents are used in refractory cases where it is hard to reduce the corticosteroid dose without causing a flare of disease or a rise in ESR or CRP.

Calcium and vitamin D supplements and sometimes bisphosphonates are necessary to prevent osteoporosis while high-dose steroids are being used (p. 563).

Takayasu's arteritis

This is a granulomatous inflammation of the aorta and its major branches and is discussed on page 809.

MEDIUM-SIZED VESSEL VASCULITIS

Polyarteritis nodosa (PAN)

Classical PAN is a rare condition which, unlike other vasculitic diseases, usually occurs in middle-aged men. It is accompanied by severe systemic manifestations, and its occasional association with hepatitis B antigenaemia suggests a vasculitis secondary to the deposition of immune complexes. Pathologically, there is fibrinoid necrosis of vessel walls with microaneurysm formation, thrombosis and infarction.

Clinical features

These include fever, malaise, weight loss and myalgia. These initial symptoms are followed by dramatic acute features that are due to organ infarction.

- *Neurological* – mononeuritis multiplex is due to arteritis of the vasa nervorum.
- *Abdominal* – pain due to arterial involvement of the abdominal viscera, mimicking acute cholecystitis, pancreatitis or appendicitis. Gastrointestinal haemorrhage occurs because of mucosal ulceration.
- *Renal* – presents with haematuria and proteinuria. Hypertension and acute/chronic renal failure occur.
- *Cardiac* – coronary arteritis causes myocardial infarction and heart failure. Pericarditis also occurs.
- *Skin* – subcutaneous haemorrhage and gangrene occur. A persistent livedo reticularis is seen in chronic cases. Cutaneous and subcutaneous palpable nodules occur, but are uncommon.
- *Lung* – involvement is rare.

Investigations and treatment

- **Blood count**. Anaemia, leucocytosis and a raised ESR occur.
- **Biopsy** material from an affected organ shows features listed above.

- **Angiography**. Demonstration of microaneurysms in hepatic, intestinal or renal vessels if necessary.
- **Other investigations** as appropriate (e.g. ECG and abdominal ultrasound), depending on the clinical problem. ANCA is positive only rarely in classic PAN.

Treatment is with corticosteroids, usually in combination with immunosuppressive drugs such as azathioprine.

Kawasaki's disease

This is an acute systemic vasculitis involving medium-sized vessels, affecting mainly children under 5 years of age. It is very frequent in Japan, and an infective trigger is suspected. It occurs world-wide and is also seen in adults.

Clinical features and treatment

The clinical features are:

- fever lasting 5 days or more
- bilateral conjunctival congestion 2–4 days after onset
- dryness and redness of the lips and oral cavity 3 days after onset
- acute cervical lymphadenopathy accompanying the fever
- polymorphic rash involving any part of the body
- redness and oedema of the palms and soles 2–5 days after onset.

Five of these six features should be present to make the diagnosis, or four of six if coronary aneurysms can be seen on two-dimensional echocardiography, MRI or angiography.

Cardiovascular changes in the acute stage include pancarditis and coronary arteritis leading to aneurysms or dilatation. Other features include diarrhoea, albuminuria, aseptic meningitis and arthralgia and, in most, there is a leucocytosis, thrombocytosis and a raised CRP. Anti-endothelial cell autoantibodies are often detectable.

Treatment is with high-dose intravenous immunoglobulin, which prevents the coronary artery disease, followed after the acute phase by aspirin 200–300 mg daily. There is no evidence that steroid treatment improves the outcome.

SMALL VESSEL VASCULITIS

This can be separated into those that are positive or negative for anti-neutrophilic cytoplasmic antibody (ANCA) (see p. 504).

The clinical features and diagnosis of small vessel vasculitis are shown in Table 10.20.

ANCA-positive small vessel vasculitis

- Wegener's granulomatosis – see page 870.
- Churg–Strauss granulomatosis – see page 870.
- Microscopic polyangiitis – see page 871.

ANCA-negative small vessel vasculitis

Henoch–Schönlein purpura – see pages 554 and 592

Cryoglobulinaemic vasculitis – see page 591

Cutaneous leucocytoclastic vasculitis (p. 1263)

This is the characteristic acute purpuric lesion which histologically involves the dermal post-capillary venules. This lesion affects only the skin and should be differentiated from similar lesions produced in systemic vasculitis. The purpura may be accompanied by arthralgia and glomerulonephritis. Hepatitis C infection is common and may be an aetiological agent. The condition can also be caused by drugs such as sulphonamides and penicillin.

Treatment of small cell vasculitis

The treatment depends on the organs involved. Vasculitis confined to the skin may not require systemic treatment

Table 10.20 Diagnosis and clinical features of small vessel vasculitis

	Wegener's granulomatosis	Churg–Strauss syndrome	Microscopic polyangiitis	Henoch–Schönlein purpura	Cryoglobulinaemic vasculitis
Features					
ANCA (in blood)	90%	+	+	–	–
PR3					
MPO	–	60%	60%		
Necrotizing granulomas	+	+	–	–	–
Cryoglobulins (in blood and vessels)	–	–	–	–	–
IgA immune deposits (mainly)	–	–	–	+	–
Asthma and eosinophilia	–	+	–	–	–
Organs involved					
Skin	40	60	40	90	90
Kidneys	80	45	90	50	55
Lungs	90	70	50	<5	<5
ENT	90	50	35	<5	<5
Musculoskeletal	60	50	60	75	70
Neurological	50	70	30	10	40
Gastrointestinal	50	50	50	60	30

Modified from Jeanette JC, Falk RJ. Small vessel vasculitis. *New England Journal of Medicine* 1997; **337:** 1512–1523. Copyright © 1997 Massachusetts Medical Society. All rights reserved.

whereas involvement of major organs (e.g. lungs or kidneys in Wegener's granulomatosis) requires high-dose corticosteroids, immunosuppression and sometimes plasma exchange.

BEHÇET'S DISEASE

Behçet's disease is an inflammatory disorder of unknown cause. There is a striking geographical distribution, it being most common in Turkey, Iran and Japan. The prevalence per 100 000 is 10–15 in Japan and 80–300 in Turkey. There is a link to the HLA-B51 allele (a split antigen of B5), with a relative risk of 5–10; this association is not seen in patients in the USA and Europe.

Clinical features

The cardinal clinical feature is recurrent oral ulceration. The international criteria for diagnosis require oral ulceration and any two of the following: genital ulcers, defined eye lesions, defined skin lesions, or a positive skin pathergy test (see below). Oral ulcers can be aphthous or herpetiform. The eye lesions include an anterior or posterior uveitis or retinal vascular lesions. Cutaneous lesions consist of erythema nodosum, pseudofolliculitis and papulopustular lesions.

Other manifestations include a self-limiting peripheral mono- or oligoarthritis affecting knees, ankles, wrists and elbows; gastrointestinal symptoms of diarrhoea, abdominal pain and anorexia; pulmonary and renal lesions; a brainstem syndrome, organic confusional states and a meningoencephalitis. All the common manifestations are self-limiting except for the ocular attacks. Repeated attacks of uveitis can cause blindness.

The pathergy reaction is highly specific to Behçet's disease. Skin injury, by a needle prick for example, leads to papule or pustule formation within 24–48 hours. Blood tests usually show raised ESR and CRP but not autoantibodies.

Treatment

Corticosteroids, immunosuppressive agents and ciclosporin are used for chronic uveitis and the rare neurological complications. Colchicine helps erythema nodosum and joint pain. Thalidomide may be useful in some cases although side-effects of drowsiness and peripheral neuropathy are common. It should not be used in pregnant women because of phocomelia (limb abnormalities).

FURTHER READING

Black CM, Denton CP. Systemic sclerosis and related disorders in adults and children. In: Isenberg DA, Maddison PJ, Woo P et al (eds) *Oxford Textbook of Rheumatology,* 3rd edn. Oxford: Oxford University Press, 2004, Section 6.7: 872–894.

Isenberg DA, Maddison PJ, Woo P et al (eds). *Oxford Textbook of Rheumatology,* 3rd edn. Oxford: Oxford University Press, 2004, Section 6.6: 819–871.

Scott DGI, Watts RA. Systemic vasculitis: epidemiology, classification and environmental factors. *Annals of Rheumatology* 2000; **59:** 161–163.

DIFFERENTIAL DIAGNOSIS OF RHEUMATIC COMPLAINTS IN THE ELDERLY (Fig.10.29)

Musculoskeletal problems

These are common at all ages. In the elderly, pain arises from a combination of age-related changes and injury.

Back and neck pain (see pp. 509 and 505)

These are commonly associated with spondylosis on X-ray. Complications include spinal stenosis or nerve root claudication. Look for locally tender areas, which might be injected. Mechanical back pain may be worse in the morning, especially when spondylosis is severe or there is marked deformity; there is a normal ESR and CRP, which distinguishes this from polymyalgia rheumatica. Rarely, elderly women present with previously undiagnosed ankylosing spondylitis.

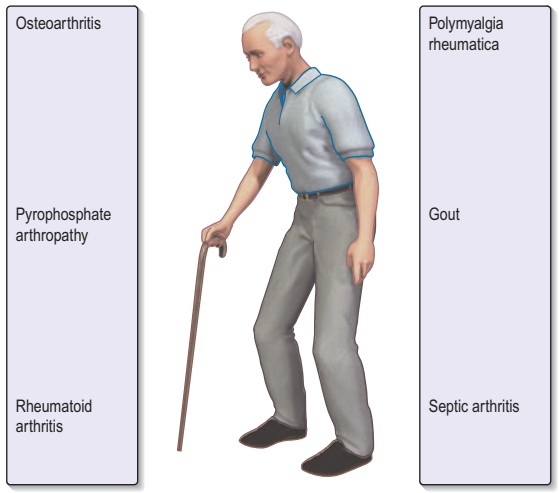


Fig. 10.29 **The differential diagnosis of arthritis in the elderly.**

Osteoporotic fractures

Osteoporotic fractures of the spine cause acute pain or deformity and subsequent postural back pain. The acute pain may warrant a period of bed rest (see p. 561). Intravenous pamidronate helps severe pain. The role of cement vertebroplasty is under investigation. The management of such problems in the elderly is the same as for younger people but a response to treatment is less predictable. Advice about posture, exercise, general fitness and about medication for osteoporosis to reduce the risk of further fractures is necessary but must be tempered by the patient's general health.

Symptomatic osteoarthritis (see p. 518)

This increases with increasing age and is uncommon below the age of 50 years. It is necessary to look for possible reversible causes of pain and disability, such as periarticular lesions or a joint effusion. Caution about drug treatment is necessary because of the increased risk of side-effects in older patients. Joint replacement surgery has transformed the outlook for many older people with OA by offering pain control and increased mobility and independence.

Rheumatoid arthritis (see p. 523)

RA may present in the elderly as a dramatic onset of symmetrical polyarthritis. This has a reasonable prognosis and responds well to low doses of prednisolone and other drug therapy. RA in the older patient may mimic polymyalgia rheumatica; the synovitis becomes apparent as the corticosteroid dose is reduced. Treat with DMARDs as for RA in order to reduce the corticosteroid dose.

Polymyalgia rheumatica and giant cell arteritis (see p. 549)

These must be recognized and treated. Failure to respond to adequate doses of prednisolone should trigger a search for an alternative disease, such as RA, vasculitis, infection or malignancy.

Systemic inflammatory vasculitis
Behçet's disease
Differential diagnosis of rheumatic complaints in the elderly
Arthritis in children
Juvenile idiopathic arthritis (JIA)

Chondrocalcinosis

This increases with increasing age and is seen on over 40% of knee X-rays in the over-80s. It may produce a variety of arthritic conditions (see p. 522).

Pseudogout (see p. 539)

This causes acute monoarthritis of the wrist or knee in older people. It responds well to aspiration and intra-articular corticosteroids or oral NSAIDs.

Gout

Gout causes acute monoarthritis in the elderly but may also present as a polyarticular inflammatory arthritis, especially in elderly women on long-term diuretic treatment, or as tophaceous gout (p. 538).

Septic arthritis (see p. 339)

Septic arthritis in the elderly and frail produces articular symptoms, and signs that may be muted. The patient usually has septicaemia and the best way to ensure the recognition of septic arthritis is to retain it in the differential diagnosis of all musculoskeletal presentations in unwell, elderly people.

Treatment

In all cases the management is similar to that described in the main part of the chapter, but special care should be taken to reduce doses of medication where possible and to think about possible adverse drug effects or interactions.

ARTHRITIS IN CHILDREN

Joint and limb pains are common in children but arthritis is fortunately rare. Babies and young children may present with immobility of a joint or a limp, but the diagnosis can be extremely difficult. Figure 10.30 summarizes the differential diagnosis.

For chronic conditions, the child and family often need a great deal of support from physiotherapists, occupational therapists, psychologists, teachers, social workers and orthopaedic surgeons. These are best obtained in specialist paediatric centres.

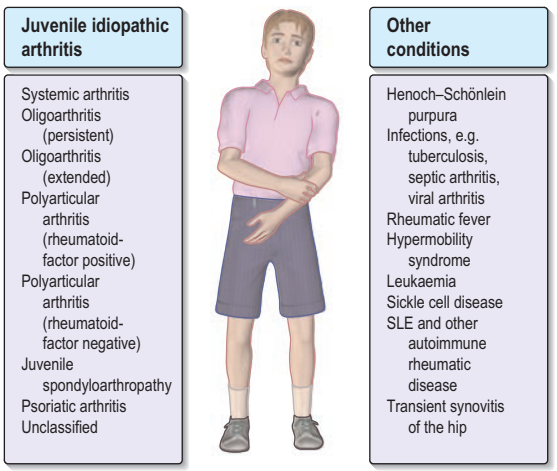


Fig. 10.30 **The differential diagnosis of arthritis in children.**

JUVENILE IDIOPATHIC ARTHRITIS (JIA)

Systemic onset JIA

Still's disease (which accounts for 10% of cases of JIA) affects boys and girls equally up to 5 years of age; then girls are more commonly affected. Adult-onset Still's disease is extremely rare.

Clinical features include a high (>39°) fever and evanescent pink maculopapular rash at the same time, with arthralgia and arthritis, myalgia and generalized lymphadenopathy. Hepatosplenomegaly, pericarditis and pleurisy occur. The differential diagnoses include malignancy, in particular leukaemia and neuroblastoma, and infection. Laboratory tests show a high ESR and CRP, neutrophilia and thrombocytosis. Autoantibodies are negative. Macrophage activation syndrome (an excessive proliferation of T cells and macrophages) is a rare but potentially fatal complication.

Oligoarthritis (persistent)

This is the most common form of JIA (50–60%) but is still a relatively uncommon condition. It affects, by definition, four or fewer joints – especially knees, ankles and wrists – often in an asymmetrical pattern. It affects mainly girls, with a peak age of 3 years. The prognosis is generally good with most going into remission. Uveitis (often with a positive ANA) occurs and requires regular screening by slit-lamp examination. Blindness can occur if it is untreated. Prognosis is generally good, with remission occurring eventually in most patients.

Oligoarthritis (extended)

In approximately 25% of patients, oligoarthritis extends to affect many more joints after around 6 months. This form of arthritis can be very destructive.

Polyarticular JIA

The rheumatoid factor-positive form (usually also anti-CCP positive) occurs in older girls, usually over 8 years. It is a systemic disease; the arthritis commonly involves the small joints of the hands, wrists, ankles and feet initially, and eventually larger joints. It can be a very destructive arthritis and needs aggressive treatment.

The rheumatoid factor-negative form is commoner. It usually affects girls under 12 years but can occur at any age. The arthritis is often asymmetrical, with a distribution similar to that seen in the RF-positive form. It may also affect the cervical spine, temporomandibular joints and elbows. Patients may be ANA positive, with a risk of chronic uveitis.

Enthesitis-related arthritis (formerly juvenile spondyloarthropathy)

This affects teenage and younger boys mainly, producing an asymmetrical arthritis of lower-limb joints and enthesitis. It is associated with HLA-B27 and a risk of iritis. It is the childhood equivalent of adult ankylosing spondylitis but spinal involvement is rare in childhood. Approximately one in three develops spinal disease in adulthood.

Psoriatic arthritis

This occurs in children and is similar in pattern to the adult form. The arthritis can be very destructive. Psoriasis may develop long after the arthritis but is found commonly in a first-degree relative.

Treatment of JIA

JIA should always be referred to a specialist paediatric rheumatology unit with facilities to assess and design treatment plans which aim to prevent long-term disability. These units also need facilities for rehabilitation, education and surgical intervention. NSAIDs reduce pain and stiffness but disease-modifying agents such as methotrexate are used to control moderate and severe disease. Corticosteroids are often required in systemic disease: intravenous pulsed methylprednisolone is used, followed by methotrexate (10–15 mg/m^2) weekly to control disease and prevent growth suppression. Anti-TNF agents (see Table 10.15) are used if methotrexate fails, and are highly effective in all types except systemic-onset disease where the results are variable. In the UK etanercept was approved by the National Institute for Health and Clinical Excellence for the treatment of JIA in 2002. Sulfasalazine is used only in enthesitis-related JIA. Aspirin may be a cause of Reye's syndrome and should not be used under the age of 12 years.

Prognosis

Before biological therapy, up to 50% of children developed long-term disability; 25% continued to have active arthritis into adult years. Death was due to infection or systemic disease with pericarditis or amyloidosis. The prognosis has now improved but long-term studies are awaited.

OTHER TYPES

Henoch–Schönlein purpura (see also p. 592)

This is the commonest systemic vasculitis seen in children. IgA immune complexes deposit in the small vessels. It often occurs after upper respiratory tract infections. Other manifestations include lower limb purpura, a transient non-migratory polyarthritis, and abdominal pain. Fifty per cent of these patients will have haematuria and proteinuria, due to a glomerulonephritis; treatment of this is discussed on page 592. The prognosis is excellent, although 1% develop chronic renal damage.

Rheumatic fever

Rheumatic fever still occurs occasionally in developed countries but is more common in developing countries. It is described on page 136.

The arthritis affects large joints and migrates between joints, each being affected for a few days at a time. This is unlike systemic onset JIA, where arthritis is usually much more persistent in each affected joint. The fever is persistent but rarely as high as in systemic onset JIA, and the temperature often remains above normal. A child may not volunteer a history of sore throat and the carditis may be silent. Isolated arthritis is the presenting symptom in 14–42%. The disease is easily missed if not included in the differential diagnosis of acute childhood arthritis.

Treatment is described on page 137.

Hypermobility syndrome

Five to ten per cent of children are hypermobile. A proportion of them will develop various musculoskeletal complaints in early childhood, such as late walking, flat feet, or nocturnal leg pains, probably due to hypermobile ankles and knees suffering recurrent sprains and strains after exercise. Joint effusions, subluxation, dislocation and ligamentous injuries may occur throughout childhood. Low back pain may develop in affected adolescents. There may be some risk of the early development of osteoarthritis in adulthood. More severe hypermobility is also seen in Ehlers–Danlos and Marfan's syndromes (see pp. 568 and 780).

Treatment is with exercise directed at improving the strength of muscles that cross affected joints, as well as overall fitness and endurance. It may be necessary to reduce or change sporting and other activities.

Miscellaneous conditions

Idiopathic musculoskeletal pain can become chronic in children. Management requires exclusion of the causes shown in Figure 10.30, but without performing unnecessary laboratory investigations. Nocturnal musculoskeletal pains are episodic and may be associated with hypermobility. They are called 'growing pains'. They often last 15–30 minutes and awaken the child from sleep, and may require physiotherapy and analgesics, together with advice and support to the parents.

Low back pain in children may reflect psychosocial problems at home or school as much as any obvious musculoskeletal pathology.

Osteochondritis can affect the ossification centre of the ends of bones. A typical condition is *Osgood–Schlatter disease*, which is characterized by localized pain and swelling over the tibial tubercle or at the patellar tendon insertion. It is usually seen in athletic teenagers and responds to local treatment and changes of sporting activities. Sever's disease is an osteochondritis of the insertion of the Achilles tendon into the calcaneum.

Perthes' disease is an idiopathic, possibly avascular, necrosis of the proximal femoral epiphysis, of unknown aetiology. It presents as a painless limp, usually in boys aged 3–12 years, and is occasionally bilateral. If severe it may require surgical correction.

Transient synovitis of the hip (irritable hip) causes painful limitation of movement, usually of one hip, after an upper respiratory infection in young children (usually boys). Symptoms usually resolve within a few weeks (2–3% develop Perthes' disease) but other more serious causes of hip pain should be excluded. Treatment is with rest and analgesia until the pain resolves.

RHEUMATOLOGICAL PROBLEMS SEEN IN OTHER DISEASES

Gastrointestinal and liver disease

- *Enteropathic synovitis* – see page 536.
- *Autoimmune hepatitis* (see p. 344) may be accompanied by an arthralgia similar to that seen in systemic lupus erythematosus. Joint pain occurs in a bilateral, symmetrical distribution, with the small joints of the hands being predominantly affected. Joints usually look

normal but sometimes there is a slight soft-tissue swelling. These patients often have positive tests for anti-nuclear antibodies.
- *Primary biliary cirrhosis* patients occasionally have a symmetrical arthropathy.
- *Hereditary haemochromatosis* is associated with arthritis in 50% of cases; this is often the first sign of the disease and chondrocalcinosis is common.
- *Whipple's disease* (see p. 281) is accompanied by fever and arthralgia.

Malignant disease

It is not uncommon for malignant diseases to present with musculoskeletal symptoms. Bone pain may be due to multiple myeloma, lymphoma, a primary tumour of bone or secondary deposits. The pain is typically unremitting, worse at night and there are other clinical clues such as weight loss or ill-health. Secondary gout occurs in conditions such as chronic myeloid leukaemia.

Neoplastic disease of bone

Malignant tumours of bone are shown in Table 10.21. The most common tumours are metastases from the bronchus, breast and prostate. Metastases from kidney and thyroid are less common. Primary bone tumours are rare and usually seen only in children and young adults.

Symptoms are usually related to the anatomical position of the tumour, with local bone pain. Systemic symptoms (e.g. malaise and pyrexia) and aches and pains occur and are occasionally related to hypercalcaemia (see p. 1020). The diagnosis of metastases can often be made from the history and examination, particularly if the primary tumour has already been diagnosed. Symptoms from bony metastases may, however, be the first presenting feature.

Investigations

- **Skeletal isotope scans** show bony metastases as 'hot' areas before radiological changes occur.
- **X-rays** may show metastases as osteolytic areas with bony destruction. Osteosclerotic metastases are characteristic of prostatic carcinoma.
- **Serum alkaline phosphatase** (from bone) is usually raised.
- **Hypercalcaemia** is seen in 10–20% of patients with malignancies. It is associated chiefly with metastases, but can also result from ectopic parathormone or parathyroid hormone-related protein secretion.
- **Prostate-specific antigen** (PSA) and serum acid phosphate are raised in the presence of prostatic metastases.

Table 10.21 Malignant neoplasms of bone

Metastases (osteolytic)
Bronchus
Breast
Prostate (often osteosclerotic as well)
Thyroid
Kidney
Multiple myeloma
Primary bone tumours (rare; seen in the young) e.g.
Osteosarcomas
Fibrosarcomas
Chondromas
Ewing's tumour

Treatment

Treatment is usually with analgesics and anti-inflammatory drugs. Local radiotherapy to bone metastases relieves pain and reduces the risk of pathological fracture. Some tumours respond to chemotherapy; others are hormone-dependent and respond to hormonal therapy. Biphosphonates (p. 564) can help symptomatically. Occasionally pathological fractures require internal fixation.

Hypertrophic pulmonary osteoarthropathy

Hypertrophic osteoarthropathy is most often associated with carcinoma of the bronchus. It is a paraneoplastic, non-metastatic complication and may be the presenting feature of the disease. It occurs only rarely with other conditions that also cause clubbing. It is seen most often in middle-aged men, who present with pain and swelling of the wrists and ankles. Other joints are involved occasionally.

The diagnosis is made on the presence of clubbing of the fingers, which is usually gross, and periosteal new bone formation along the shafts of the distal ends of the radius, ulna, tibia and fibula on X-ray. A chest X-ray usually shows the malignancy.

Treatment should be directed at the underlying carcinoma; if this can be removed, the arthropathy disappears. NSAIDs relieve the symptoms.

Paraneoplastic polyarthritis

This is seen with carcinoma of the breast in women and of the lung in men, and also with renal cell carcinoma. The neoplasm may be occult at onset and the diagnosis is then difficult to make.

Skin disease

Psoriatic arthritis

This is discussed on page 1242.

Erythema nodosum (p. 1250)

This is accompanied by arthritis in over 50% of cases. The knees and ankles are particularly affected, being swollen, red and tender. The arthritis subsides, along with the skin lesions, within a few months. Treatment is with NSAIDs or occasionally steroids.

Neurological disease

Neuropathic joints (Charcot's joints) are joints damaged by trauma as a result of the loss of the protective pain sensation. They were first described by Charcot in relation to tabes dorsalis. They are also seen in syringomyelia, diabetes mellitus and leprosy. The site of the neuropathic joint depends upon the localization of the pain loss:

- In tabes dorsalis, the knees and ankles are most often affected.
- In diabetes mellitus, the joints of the tarsus are involved.
- In syringomyelia, the shoulder is involved.

Ask the authors
www.studentconsult.com

Neuropathic joints are not painful, although there may be painful episodes associated with crystal deposition. Presentation is usually with swelling and instability. Eventually severe deformities develop.

The characteristic finding is a swollen joint with abnormal but painless movement. This is associated with neurological findings that depend upon the underlying disease (e.g. dissociated sensory loss in syringomyelia or polyneuropathy in diabetes). X-ray changes are characteristic, with gross joint disorganization and bony distortion.

Treatment is symptomatic. Surgery may be required in advanced cases.

Blood disease

Arthritis due to haemarthrosis is a common presenting feature of *haemophilia* (see p. 438). Attacks begin in early childhood in most cases and are recurrent. The knee is the most commonly affected joint but the elbows and ankles are sometimes involved. The arthritis can lead to bone destruction and disorganization of joints. Apart from replacement of factor VIII, affected joints require initial immobilization followed by physiotherapy to restore movement and measures to prevent and correct deformities.

Sickle cell crises (p. 411) are often accompanied by joint pain that particularly affects the hands and feet in a bilateral, symmetrical distribution. Affected joints usually look normal but are occasionally swollen. This condition may also be complicated by avascular necrosis (see p. 410) and by osteomyelitis.

Arthritis can also occur in *acute leukaemia*; it may be the presenting feature in childhood. The knee is particularly affected and is very painful, warm and swollen. Treatment is directed at the underlying leukaemia. Arthritis may also occur in chronic leukaemia, with leukaemic deposits in and around the joints.

Individuals with *thalassaemia major* (see p. 408) are living longer and are presenting with back pain due to premature disc degeneration, secondary spondylosis and crush fractures due to osteoporosis. There is marked discal calcification.

Endocrine and metabolic disorders

Hypothyroid patients may complain of pain and stiffness of proximal muscles, resembling polymyalgia rheumatica. They may also have carpal tunnel syndrome. Less often, there is an arthritis accompanied by joint effusions, particularly in the knees, wrist and small joints of the hands and feet. These problems respond rapidly to thyroxine.

Hyperparathyroidism may be complicated by chondrocalcinosis and acute pseudogout.

In *acromegaly*, arthralgia occurs in about 50% of patients. It particularly affects the small joints of the hands and knees. There may be carpal tunnel syndrome.

In *Cushing's disease*, back pain is common.

Joint disorders related to *diabetes mellitus* are described on page 1057.

Familial hypercholesterolaemia is associated with oligo- or polyarthritis usually with tendon xanthomata. Arthritis also occurs in combined hyperlipidaemia.

MISCELLANEOUS ARTHROPATHIES

Familial Mediterranean Fever (FMF)

FMF is inherited as an autosomal recessive condition and occurs in certain ethnic groups, particularly Arabs, Turks, Armenians and Sephardic Jews. The gene, called *MEFV*, has been localized to chromosome 16. It encodes for pyrin or marenostrin, which activate the biosynthesis of a chemotactic-factor inactivator in neutrophils. Failure to produce this leads to FMF attacks.

These are characterized by recurrent attacks of fever, arthritis and serositis. Abdominal or chest pain due to peritonitis or pleurisy occurs. The arthritis is usually monoarticular and attacks last up to 1 week. The condition may be mistaken for palindromic rheumatism (p. 525), but such attacks are not usually accompanied by fever.

The diagnosis can be made by PCR, if available, but usually it is based on the clinical picture and exclusion of other conditions.

Treatment. Regular colchicine 1000–1500 μg daily can usually prevent the attacks. In resistant patients, thalidomide (p. 552) and anakinra (p. 530) can be tried. In general the disorder is benign but in 25% of cases renal amyloidosis develops.

Sarcoidosis (see also p. 868)

The most common type of arthritis is that associated with erythema nodosum, which occurs in 20% of cases of sarcoidosis at or soon after the onset of the disease. The most useful diagnostic test is a chest X-ray, which shows hilar lymphadenopathy in 80% of cases.

Other patterns of arthritis occur later in the disease. These include a transient rheumatoid-like polyarthritis and an acute monoarthritis that can be mistaken for gout.

Treatment is with NSAIDs, but if these fail to control the symptoms, corticosteroids are usually very effective.

Osteochondromatosis

In this condition, foci of cartilage form within the synovial membrane. These foci become calcified and then ossified (osteochondromas). They may give rise to loose bodies within the joint. The condition occurs in a single joint of a young adult and X-rays are usually diagnostic.

Treatment involves removal of loose bodies and synovectomy.

Pigmented villonodular synovitis

This is characterized by exuberant synovial proliferation that occurs either in joints or in tendon sheaths. The main manifestation in joints is recurrent haemarthrosis. It may produce progressive local bone destruction. A malignant form is seen occasionally.

Treatment is synovectomy or radiotherapy. In tendon sheaths, the condition gives rise to a nodular mass that requires excision.

Relapsing polychondritis

Relapsing polychondritis is a rare inflammatory condition of cartilage. It occurs equally in males and

females, usually the elderly. Tenderness, inflammation and eventual destruction of cartilage occur, mainly in the ear, nose, larynx or trachea. A seronegative polyarthritis occurs, as well as episcleritis and evidence of a vasculitis (e.g. glomerulonephritis). The diagnosis is clinical with laboratory evidence of acute inflammation.

Treatment involves corticosteroids and immunosuppressive agents.

DISEASES OF BONE

Bone is a specialized connective tissue serving three major functions:

- *mechanical* – structure and muscle attachment for movement
- *metabolic* – providing the body's primary store of calcium and phosphate
- *protective* – enclosing the marrow and other vital organs.

STRUCTURE AND PHYSIOLOGY

(Fig. 10.31)

Bone structure

Bone is comprised of cells and a matrix of organic protein and inorganic mineral. Long bones (femur, tibia, humerus) and flat bones (skull, scapula) have different embryological templates, with varying proportions of cortical and trabecular bone.

- *Cortical (compact) bone* forms the shaft of long bones and the outer shell of flat bones. Formed of concentric rings of bone, it is particularly adapted to withstand bending strain.

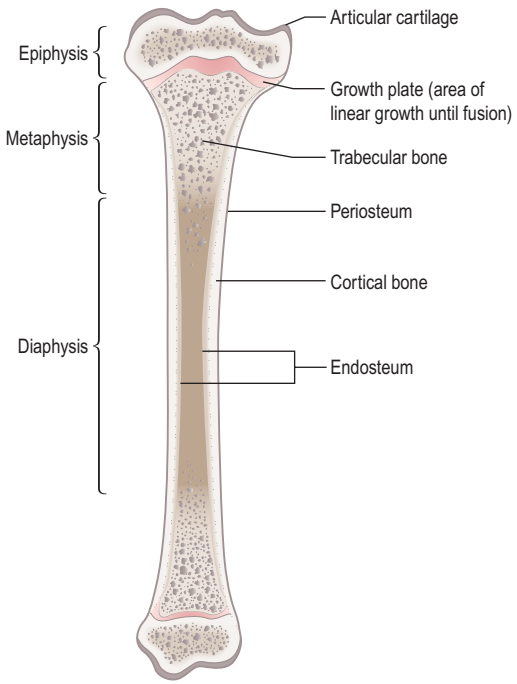


Fig. 10.31 **Diagram of a longitudinal section of a growing long bone.**

- *Trabecular (cancellous) bone* is found at the ends of long bones and inside flat bones. Comprised of a network of interconnecting rods and plates of bone, it offers resistance to compressive loads. It is also the main site of bone remodelling for mineral homeostasis.
- *Woven bone* lacks an organized structure. It appears in the first few years of life, at sites of fracture repair and in high-turnover bone disorders such as Paget's disease.
- *Lamellar bone*, in which the collagen fibres are arranged in parallel bundles, forms the bone in adult life.

Matrix components

Type I collagen is the main protein, forming parallel lamellae of differing density (which impairs spreading of cracks). In cortical bone, concentric lamellae form around a central blood supply (Haversian system) which communicates via transverse (Volkmann's) canals. Non-collagen proteins include osteopontin, osteocalcin and fibronectin. Bone mineral largely consists of calcium and phosphate in the form of hydroxyapatite.

Bone cells

Osteocytes

These are small cells, derived from osteoblasts, embedded in bone and interconnected with each other and with bone lining cells through cytoplasmic processes. They respond to mechanical strain by undergoing apoptosis or through altered cell signalling.

Osteoblasts

Derived from local mesenchymal stem cells, these cells synthesize matrix (osteoid) and regulate its mineralization. After bone formation, the majority of osteoblasts are removed by apoptosis, others remaining at the bone/marrow interface as lining cells or within the bone as osteocytes. Osteoblasts critically regulate bone resorption through the balance in expression of the stimulatory RANKL (the *l*igand for *r*eceptor *a*ctivator of *n*uclear factor *k*appaB) and its antagonist, osteoprotegerin (OPG). Osteoblasts are rich in alkaline phosphatase and express receptors for PTH, oestrogen, glucocorticoids, vitamin D, inflammatory cytokines and the transforming growth factor-β family, all of which may therefore influence bone remodelling.

Osteoclasts

These are cells with the unique capacity to resorb bone and are derived from haematopoietic precursors of the macrophage lineage. In response to RANKL and macrophage colony stimulating factor (M-CSF), they attach to bone, creating a ruffled border which forms a number of extracellular lysosomal compartments. Hydrogen ions are actively

secreted into these spaces and the acid environment removes the mineral phase before specialized cysteine proteases (e.g. cathepsin K) resorb the collagen matrix.

Bone growth and remodelling

Longitudinal growth occurs at the epiphyseal growth plate, a cartilage structure between the epiphysis and metaphysis (Fig. 10.31). Cartilage production is tightly regulated, with subsequent mineralization and growth finally arrested at between 18 and 21 years when the epiphysis and metaphysis fuse.

In adults, bone is regularly remodelled to ensure repair of microdamage and turnover of calcium and phosphate for homeostasis. This is carried out by the *basic multicellular unit* (BMU) (Fig. 10.32). Retraction of bone lining cells and altered expression of RANKL and OPG and other regulators of bone remodelling lead to induction of osteoclast formation and bone resorption. Unknown factors limit the amount of bone resorbed, after which osteoblasts fill in the resorption cavity. Bone remodelling is said to be coupled as formation normally follows resorption. New bone formation without resorption may, however, occur in the adult skeleton in response to anabolic therapy such as parathyroid hormone peptides. The cellular regulation of remodelling is incompletely defined. Recognized signals include changes in osteocytes (apoptosis or altered signalling of sclerostin, prostaglandins and other molecules), regulation of bone formation (reciprocal effects of *wnt* and of *dickkopf (Dkk)* and sclerostin on the LRP5/6–β-catenin pathway), and altered balance of RANKL and OPG. Additional influences include systemic hormones of which oestrogen (in both sexes) is particularly involved, promoting survival of osteocytes and inhibiting osteoclastogenesis.

CALCIUM HOMEOSTASIS AND ITS REGULATION

Calcium homeostasis is regulated by the effects of parathyroid hormone (PTH) and 1,25-dihydroxyvitamin D (1,25$(OH)_2D_3$) on gut, kidney and bone. Calcium-sensing receptors are present in the parathyroid glands, kidney, brain and other organs.

CALCIUM ABSORPTION AND DISTRIBUTION (Fig. 10.33)

Daily calcium consumption, primarily from dairy foods, should ideally be around 20–25 mmol (800–1000 mg). The combined effect of calcium and vitamin D deficiency contributes to bone fragility in some older persons. Intestinal

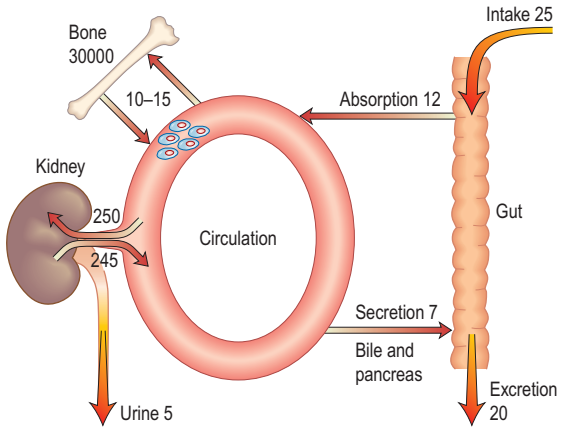


Fig. 10.33 **Calcium exchange in the normal human.** The amounts are shown in mmol per day.

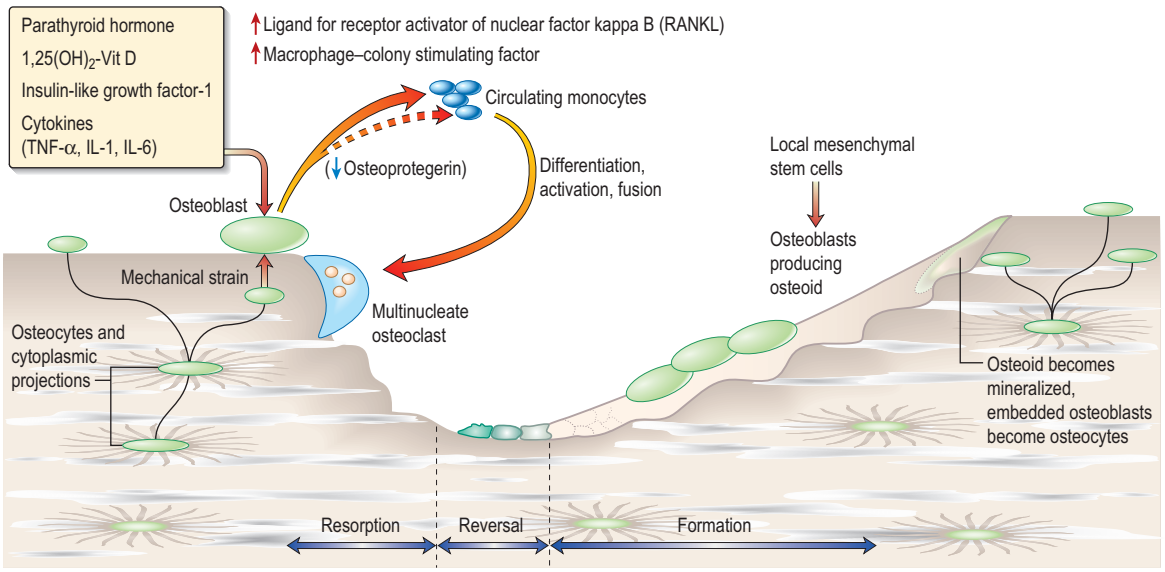


Fig. 10.32 **A diagram of a basic multicellular unit on a trabecular bone surface.** *Resorption*, reversal and bone formation are occurring simultaneously at different locations within the remodelling space. The sequence of events over time is shown, from left to right. A multinucleate osteoclast has formed in response to stimuli including increased RANKL and M-CSF and reduced osteoprotegerin (OPG) expression by osteoblasts lining the bone surface. The osteoclast removes an area of bone, with unknown signals limiting the depth of resorption. During the *reversal phase* (centre) macrophages appear in the osteoclasts' wake, removing apoptotic cells and preparing the bone surface to receive newly formed osteoid from osteoblasts (right) which form some time after the osteoclasts, possibly in response to the microenvironment of healthy bone exposed by resorption. This osteoid will gradually be fully mineralized, completing the cycle over approximately 120 days.

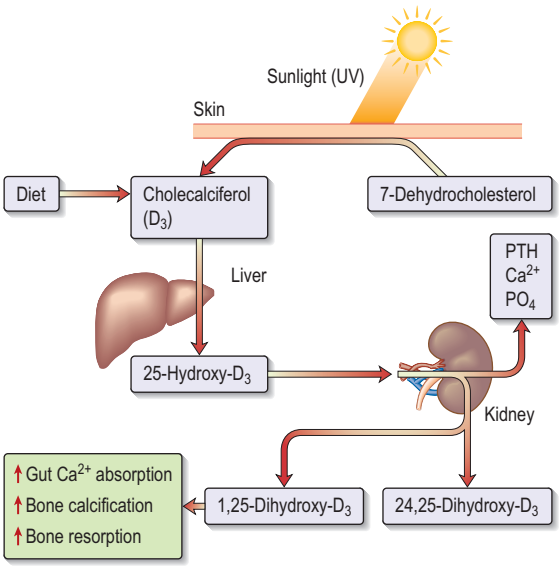


Fig. 10.34 **The metabolism and actions of vitamin D.**

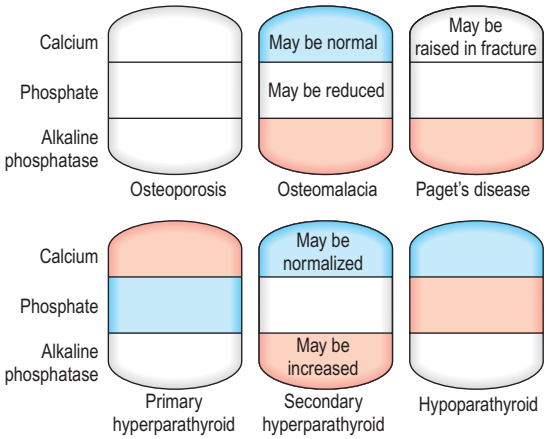


Fig. 10.35 **Changes in serum calcium, phosphate and alkaline phosphatase in main bone disorders.** Shading represents increased (red), reduced (blue) or normal (white) levels.

absorption of calcium is reduced by vitamin D deficiency, and in malabsorption states.

Vitamin D metabolism (Fig. 10.34)

The primary source of vitamin D in humans is photoactivation in the skin of 7-dehydrocholesterol to cholecalciferol, which is then converted first in the liver to 25-hydroxyvitamin D ($25(OH)D_3$) and subsequently in the kidney (by the enzyme 1α hydroxylase) to $1,25(OH)_2D_3$. Regulation of the latter step is by PTH, phosphate and feedback inhibition by $1,25(OH)_2D_3$.

Parathyroid hormone (PTH)

PTH, an 84 amino-acid hormone, is secreted from the chief cells of the parathyroid gland, which bear calcium-sensing and vitamin D receptors. PTH increases renal phosphate excretion and increases plasma calcium by:

- increasing osteoclastic activity (a rapid response)
- increasing intestinal absorption of calcium (a slower response)
- increasing 1α-hydroxylation of vitamin D (the rate-limiting step)
- increasing renal tubular reabsorption of calcium.

Hypomagnesaemia can suppress the normal PTH response to hypocalcaemia.

Calcitonin

Calcitonin is produced by thyroid C cells. Although calcitonin inhibits osteoclastic bone resorption and increases the renal excretion of calcium and phosphate, neither excess calcitonin (in medullary carcinoma of the thyroid) nor its deficiency following thyroidectomy has significant skeletal effects in humans.

INVESTIGATION OF BONE AND CALCIUM DISORDERS (Fig. 10.35)

Total plasma calcium (2.2–2.6 mmol/L)

About 40% is ionized and physiologically active: the remainder is complexed or protein bound. As ionized calcium is difficult to measure, normal practice is to measure total calcium, correcting the value to allow for protein binding according to the following formula: add or subtract 0.02 mmol/L for each gram per litre of a simultaneous albumin level below or above 40 g/L. For critical measurements, samples should be taken in the fasting state and without a tourniquet (the latter may increase local plasma calcium concentration).

Plasma phosphate (0.8–1.4 mmol/L)

Phosphate is essential to most biological systems. High levels are found in renal failure and hypoparathyroidism, while low levels are associated with primary hyperparathyroidism, hypophosphataemic rickets and osteomalacia and other disorders associated with reduced renal tubular phosphate reabsorption.

Plasma PTH (reference range 10–65 ng/mL, need to check with your local laboratory)

The PTH assay measures the intact hormone. In hypercalcaemia due to causes other than hyperparathyroidism, serum PTH levels are suppressed. Lithium toxicity may be associated with raised PTH levels, and in familial hypocalciuric hypercalcaemia (FHH) serum PTH may be normal or marginally elevated.

Serum 25-hydroxyvitamin D (reference ranges are seasonal; generally >20 ng/mL or 50 nmol/L are considered normal)

Serum $25-(OH)D_3$ provides an assessment of vitamin D status, provided there is no reason to suspect defective 1α-hydroxylation in the kidney. Very low levels (e.g. ≤7.5 ng/mL) indicate the presence of severe vitamin D deficiency whereas reductions of a lesser magnitude are referred to as vitamin D *insufficiency*, in which the main functional consequence is secondary hyperparathyroidism and increased bone turnover, particularly in cortical bone.

24-hour urinary calcium (normal range 2.5 up to 6.25 (female) and 7.5 (male) mmol/24 h)

This is increased where renal tubular reabsorption of calcium is decreased, and in hypercalcaemia. One exception

is familial hypocalciuric hypercalcaemia where the genetic defect leads to inappropriately reduced calcium excretion. Measurement of 24-hour urinary calcium excretion should be routinely performed in the assessment of hypercalcaemic patients.

Biochemical markers of bone formation and resorption

While these are available in many laboratories, their use is limited by large biovariability and measurement variance. Serial measurements at the same time of day in individual patients are useful in assessing response to treatment of metabolic bone diseases. In addition, measurements of bone turnover markers may have a role in the assessment of fracture risk.

- *Bone-specific alkaline phosphatase.* Circulating alkaline phosphatase is derived from bone, liver and placenta. The bone-specific isoenzyme can be measured as a marker of formation, although there is some overlap with the liver isoenzyme. Elevated serum levels occur during bone growth, fracture repair, and in high bone turnover states (Fig. 10.35).
- *Type 1 collagen propeptides* are by-products of collagen synthesis. Serum levels of both the carboxyterminal (P1CP) and aminoterminal (P1NP) propeptides reflect bone formation.
- *Serum osteocalcin* is another bone formation marker.
- *Serum or urine levels of N-terminal (NTX) and C-terminal (CTX) cross-linked telopeptides* reflect bone resorption. They may change rapidly in response to anti-resorptive drugs or in disease states.

Diagnostic imaging

- *Plain radiographs* identify fractures, tumours and infections. Other specific features may be seen (see following sections).
- *Radionucleotide imaging*. Technetium-99m-labelled methylene bisphosphonate uptake in bone reflects bone turnover and blood flow. Increased uptake is therefore seen in fracture, tumour and metastatic deposits, infection and Paget's disease of bone.
- *Magnetic resonance imaging* is rarely used specifically to image bone. However, variation in technique to suppress the high signal associated with bone marrow (such as STIR sequences, see p. 505) allows highly sensitive recognition of 'bone marrow oedema', a non-specific feature of a number of bone disorders including avascular necrosis. High-resolution MRI provides information about bone microarchitecture, but this is not yet applied in the clinical setting.
- *Bone biopsy* (Fig. 10.36). A core of bone is removed, including both cortices of the iliac crest, using a trephine. The non-decalcified specimen is examined for static and dynamic (bone turnover) indices. An oral tetracycline is given to the patient prior to the biopsy, for 2 days on two occasions 10 days apart, allowing assessment of the rate of bone turnover and mineralization. Biopsy is most commonly used in assessment of suspected renal bone disease and osteomalacia.
- *Bone densitometry measurements* (p. 562).

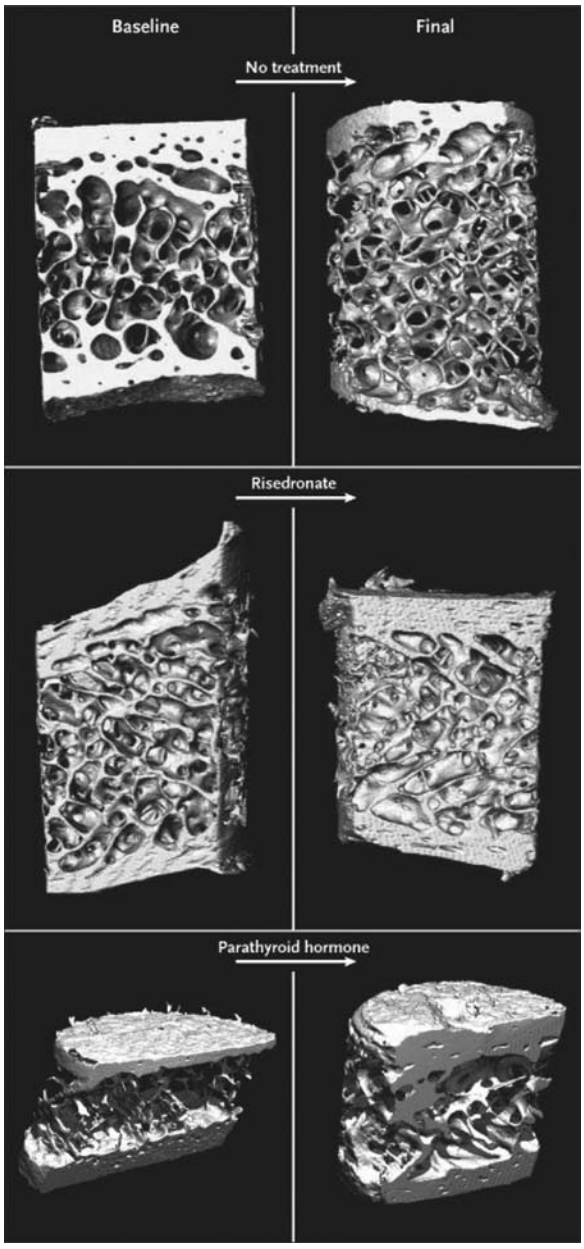


Fig. 10.36 **Computerized images of bone biopsies in osteoporosis demonstrate changes in bone architecture over time.** *Upper pair*: without treatment, the bone cortex becomes thinner, and numerous trabecular rods and plates are lost. *Centre pair*: in response to bisphosphonate treatment, resorption is reduced, with preservation of bone mineral and structure. *Lower pair*: in response to anabolic therapy, bone is increased at both trabecular and cortical sites. Reproduced from *Journal of Bone and Mineral Research* 2003; **18**: 1932–1941 with permission of the American Society for Bone and Mineral Research.

OSTEOPOROSIS

Definition and incidence

Osteoporosis is defined as 'a disease characterized by low bone mass and micro-architectural deterioration of bone tissue, leading to enhanced bone fragility and an increase in fracture risk'.

The World Health Organization (WHO) defines osteoporosis as a bone density of 2.5 standard deviations (SDs)

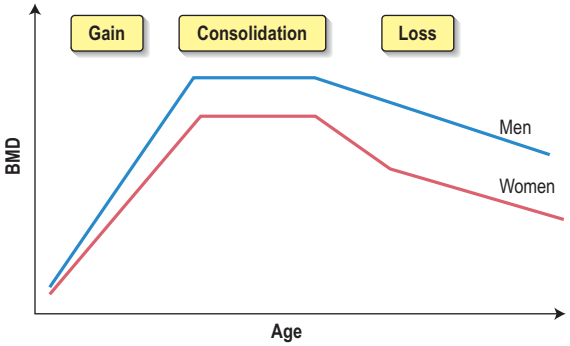


Fig. 10.37 **Lifetime changes in bone mineral density (BMD).** Peak bone mass is achieved between 20 and 30 years of age (gain), and consolidated up to around the age of 40 years. Then, age-related bone loss occurs in both men and women, with an accelerated loss in women starting around the time of menopause which lasts between 5 and 10 years.

below the young healthy adult mean value (T-score ≤ –2.5) or lower. Values between –1 and –2.5 SDs below the young adult mean are termed 'osteopenia'. The rationale for this definition is the inverse relationship between bone mineral density and fracture risk in postmenopausal women and also older men. However, this definition should not be applied to younger women, men or children.

Fractures due to osteoporosis are a major cause of morbidity and mortality in elderly populations. One in two women and one in five men aged 50 years will have an osteoporotic fracture during their remaining lifetime. Caucasian and Asian races are particularly at risk. The annual cost to the health services in the UK is estimated at around £1.8 billion. As the risk of fracture increases exponentially with age, changing population demographics will increase the burden of disease.

Pathogenesis

Osteoporosis results from increased bone breakdown by osteoclasts and decreased bone formation by osteoblasts leading to loss of bone mass.

Bone mass decreases with age (Fig. 10.37) but will depend on the 'peak' mass attained in adult life and on the rate of loss in later life. Genetic factors are the single most significant influence on peak bone mass, but multiple genes are involved, including collagen type1A1 (p. 568), vitamin D receptor and oestrogen receptor genes. Nutritional factors, sex hormone status and physical activity also affect peak mass.

Risk factors

Oestrogen deficiency is a major factor in the pathogenesis of accelerated bone loss. In the elderly, vitamin D insufficiency and consequent hyperparathyroidism are pathogenetic factors.

Additional risk factors are associated with increased bone loss, or with increased bone fragility but are independent of effects on the bone mineral density (Table 10.22). For example, hyperparathyroidism, hyperthyroidism and malabsorption each increase the risk of the person having a low bone mass and are BMD dependent. However, other risk factors such as previous fracture, increasing age, glucocorticoid therapy, smoking and falls increase the risk of fracture, on top of the risk associated with their particular low bone mass. The effect of many of these risk factors is particularly notable in terms of hip fracture risk.

Treatment also can depend on the type of risk factors, for if they are recognized as 'skeletal', they respond to bone-directed treatment and if as 'non-skeletal', they require other intervention (e.g. reduction of falls risk).

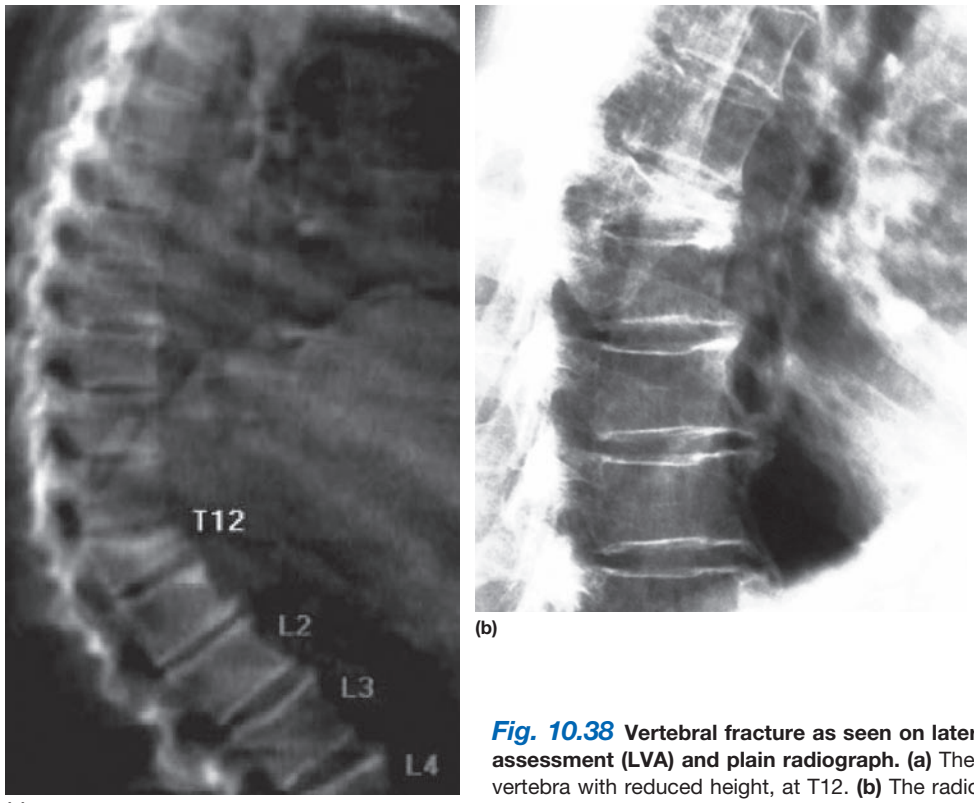
Not all causes of osteoporosis affect bone remodelling and architecture in the same way. For example, oestrogen deficiency results in increased numbers of remodelling units, and increased resorption depth exceeding osteoblast synthethic capacity, with a loss of resistance to fracture that is not fully reflected in the bone density measurement. Glucocorticoids induce a high turnover state initially, with increased fracture risk evident within three months of starting therapy. More prolonged use leads to a reduced turnover state but with a net loss due to reduced synthesis (through increased inhibition of the *wnt*–LRP5/6 axis).

Table 10.22 Risk factors for osteoporosis

BMD-dependent	BMD-independent
Female sex	Increasing age
Caucasian/Asian	Previous fragility fracture
Gastrointestinal disease	Family history of hip fracture
Hypogonadism	Low body mass index
Immobilization	Smoking
Chronic liver disease	Alcohol abuse
Chronic renal disease	Glucocorticoid therapy
Low dietary calcium intake	High bone turnover
Vitamin D insufficiency	Increased risk of falling
Chronic obstructive pulmonary disease	Rheumatoid arthritis
Drugs	
Heparin	
Calcineurin inhibitors e.g. ciclosporin	
Anticonvulsants	
Thiazolidinediones	
Aromatase inhibitors	
Anti-androgens	
GnRH analogues	
Proton pump inhibitors	
?Selective serotonin reuptake inhibitors	
Endocrine disease	
Cushing's syndrome	
Hyperthyroidism	
Hyperparathyroidism	
Other diseases	
Diabetes mellitus	
Mastocytosis	
Multiple myeloma	
Osteogenesis imperfecta	

Clinical features

Fracture is the only cause of symptoms in osteoporosis. Sudden onset of severe pain in the spine, often radiating around to the front, suggests *vertebral crush fracture*. However, only about one in three vertebral fractures is symptomatic. Pain from mechanical derangement, increasing kyphosis, height loss and abdominal protuberance follow crushed vertebrae. *Colles' fractures* typically follow a fall on an outstretched arm. *Fractures of the proximal femur* usually



(a) (b)

Fig. 10.38 **Vertebral fracture as seen on lateral vertebral assessment (LVA) and plain radiograph. (a)** The LVA image shows a vertebra with reduced height, at T12. **(b)** The radiograph shows a typical wedge compression fracture.

occur in older individuals falling on their side or back. Other causes of low-trauma fractures must not be overlooked, including metastatic disease and myeloma.

Investigations

Plain radiographs usually show a fracture and may reveal previously asymptomatic vertebral deformities. Such clinically silent fractures may also be detected during the DXA scan with an additional analysis (called *lateral vertebral assessment,* Fig. 10.38) carried out with a much lower radiation dose than conventional imaging.

Bone density

- *Dual energy X-ray absorptiometry* (DXA) measures areal bone density (mineral per surface area rather than a true volumetric density), usually of the lumbar spine and proximal femur. It is precise, accurate, uses low doses of radiation and is the gold standard in osteoporosis diagnosis (Fig. 10.39). Because of osteophytes, spinal deformity and vertebral fractures, spinal values may be artefactually elevated and should be interpreted with caution in the elderly.
- *Quantitative ultrasound* of the calcaneum. This does not require ionizing radiation and is cheaper than other methods. It cannot be used for diagnostic purposes but is useful as a screening procedure prior to DXA assessment. The WHO T-score definition of osteoporosis should only be applied to DXA carried out at the hip and spine, and not to ultrasound, where a term such as 'low bone mass' is preferable.
- *Quantitative CT scanning* allows true volumetric assessment, and distinction between trabecular and cortical bone. However, it is more expensive, requires higher radiation than other techniques, and to date offers no clinical advantage.

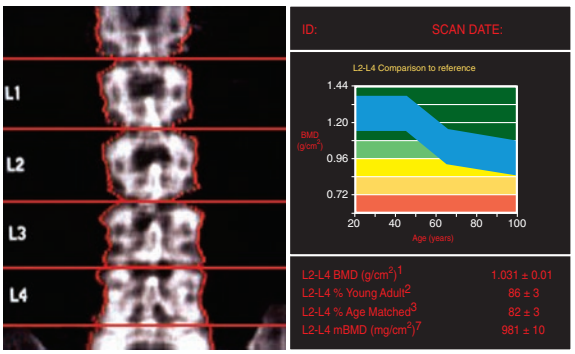


Fig. 10.39 **Dual energy X-ray absorptiometry of the lumbar spine.** The AP image of the lumbar spine is shown on the left and on the right, the patient's value is expressed in relation to the reference range.

Associated disease and risk factors

Investigations to exclude other diseases or identify contributory factors associated with osteoporosis should be per-

formed and are particularly necessary in men, in whom secondary causes are more common (see Table 10.22).

Selection of individuals for treatment: risk assessment

The purpose of treatment is to reduce the risk of fractures. Thus, assessment of absolute fracture risk should be made in every case. Although bone mineral density measurements in the spine and proximal femur provide useful information about fracture risk, they have a relatively low sensitivity and the majority of fragility fractures occur in women with a T-score ≥–2.5. Prediction of fracture risk can be improved by the addition of risk factors that are at least partially independent of bone mineral density (see Table 10.22). This forms the basis of a 'fracture risk tool' for clinical practice that has been developed under the auspices of the WHO, which provides an algorithm for calculation of 10-year fracture probabilities, based on independent clinical risk factors with and without bone mineral density values. The intervention threshold can then be determined by the cost-effectiveness of treatment and by clinical judgement.

All postmenopausal women and older men with a history of fragility fracture should be reviewed for treatment. In those aged ≥75 years, DXA is often not necessary prior to treatment but in those ≤75 years, DXA is useful in guiding treatment decisions (see Table 10.23) . However, other risk factors should also be taken into account when deciding whether to treat, particularly age and a history of previous fracture. Thus individuals with higher BMD values but with other clinical risk factors have a greater fracture probability than those with low BMD values in the absence of risk factors (Fig. 10.40).

Prevention and treatment

- *Symptomatic management.* New vertebral fractures require bed rest for 1–2 weeks with strong analgesia (e.g. meptazinol 200 mg every 3–6 hours i.m. or i.v. – 100 mg) and transcutaneous electrical nerve stimulation (TENS) may be helpful. Muscle relaxants (e.g. diazepam 2 mg three times daily), subcutaneous calcitonin (50 IU daily) or a single intravenous infusion of pamidronate (60–90 mg) may be useful for pain relief, and physiotherapy helps restore confident mobilization. Non-spinal fractures should be treated by conventional orthopaedic means.
- *Calcium and vitamin D.* Daily intakes of 700–1000 mg of calcium (ideally 1500 mg postmenopausally) and 400–800 IU of vitamin D are recommended. In those with low dietary calcium intake and at risk from vitamin D insufficiency, calcium and vitamin D supplements should be advised. Vitamin D and calcium supplements should also be given to patients receiving bone protective medication.
- *Exercise.* Thirty minutes of weight-bearing exercise three times a week may increase BMD, though this has not been a universal finding in all studies. Gentle exercise in the elderly may reduce the risk of falls and improve the protective responses to falling.
- *Smoking cessation.* Smoking is associated with lower BMD and increased fracture risk. Alcohol abuse (>3 units/day) should be avoided.
- *Reduce falls.* Physiotherapy and assessment of home safety are helpful. Hip protectors do reduce fractures in the elderly in residential care when worn correctly, but compliance is poor.

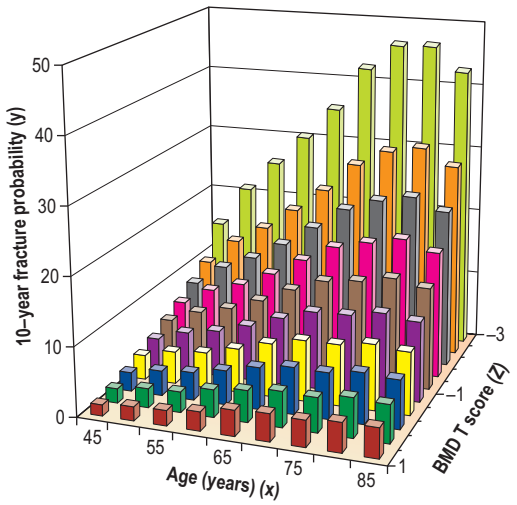


Fig. 10.40 **A graph illustrating the combined effects of age (x-axis) and reduced bone mass (expressed as T-scores on the z-axis) on the 10-year probability of fracture (y-axis) in a population of women.** Reduced bone mass osteopenia, T-scores between –1 and –2.5. Osteoporotic bone (T-scores less than –2.5) is shown in pale green. Note that the risk of fracture may be greater in an older woman with osteopenia than in a young woman with osteoporosis. Data from Kanis JA. Diagnosis of osteoporosis and assessment of fracture risk. *Lancet* 2002; **359**: 1929–1936 with permission.

Table 10.23 Indications for DXA scanning

Radiographic osteopenia
Previous fragility fracture (in those aged less than 75 years)
Glucocorticoid therapy (in those aged less than 65 years)
Body mass index below 19 (kg/m^2)
Maternal history of hip fracture
BMD-dependent risk factors in Table 10.22

In patients presenting with height loss and/or kyphosis, lateral thoracic spine X-ray should be the initial investigation.

Pharmacological intervention (see Fig. 10.36)

Most interventions used act by inhibiting bone resorption (anti-resorptives) although parathyroid hormone peptides stimulate bone formation. The mechanism of action of strontium ranelate remains incompletely defined.

The evidence base for *anti-fracture efficacy* of interventions varies. Adequately powered randomized controlled trials, with fracture as the primary endpoint, exist for alendronate, risedronate, ibandronate, zoledronate, raloxifene, hormone replacement therapy, strontium ranelate, teriparatide (recombinant human PTH peptide 1-34), human recombinant parathyroid hormone 1-84 and combined calcium/vitamin D (the latter in frail older individuals only).

Some interventions have been shown to reduce fracture at vertebral and non-vertebral sites, including the hip, whereas others have not been demonstrated to be effective at all sites (Table 10.24). Since a fracture at one site increases the risk of subsequent fracture at any site, treatments with efficacy at all major fracture sites (particularly spine and hip) are preferable. Hence the bisphosphonates and strontium ranelate are generally regarded as first-line options in the majority of postmenopausal women with osteoporosis.

Table 10.24 Medications to reduce fracture risk in postmenopausal women

Intervention	Vertebral fracture	Non-vertebral fracture	Hip fracture
Alendronate	+	+	+
Etidronate	+	ND	ND
Ibandronate	+	+*	ND
Raloxifene	+	ND	ND
Risedronate	+	+	+
Strontium ranelate	+	+	+*
Zoledronate	+	+	+
Teriparatide	+	+	ND
PTH (1–84)	+	ND	ND

*Demonstrated only in high-risk subgroup.
ND, not demonstrated.

- *Bisphosphonates*, synthetic analogues of bone pyrophosphate, adhere to hydroxyapatite and inhibit osteoclasts. *Alendronate, risedronate and zoledronate* increase bone mineral density at the hip and spine, and reduce the incidence of fractures at vertebral, hip and other sites, particularly in women who have already sustained a vertebral fracture. Alendronate and risedronate are given as once daily or, most commonly, once weekly doses (70 mg and 35 mg, respectively) and zoledronate is given as a once yearly infusion of 5 mg. *Ibandronate* is a recent addition, available as a once monthly oral therapy (150 mg per month) or as a three-monthly intravenous injection (3 mg per 3 months). However, anti-fracture efficacy at non-vertebral sites has only been shown in high-risk subgroups.
  - Oral bisphosphonates should be taken in the fasting state with a large drink of water, while the patient is standing or sitting upright. The patient should subsequently remain upright and avoid food and drink for at least 30 minutes.
  - Bisphosphonates are generally well tolerated but may be associated with upper gastrointestinal side-effects such as oesophagitis, particularly if the dosing instructions are not closely followed. Bisphosphonates should be used with caution in patients with chronic kidney disease. Osteonecrosis of the jaw is seen following high dose i.v. (rarely oral) nitrogen-containing bisphosphonate in patients with malignant disease. It is associated with poor dental hygiene. An association with bisphonate use in osteoporosis has not been clearly eatablished.
  - The optimal duration of bisphosphonate therapy is unknown; prolonged suppression of bone turnover may have adverse effects, and it is currently advised to reassess treatment after 5–10 years.
- *Strontium ranelate* consists of two stable strontium atoms linked to an organic acid, ranelic acid. Its mechanism of action remains uncertain although it has weak anti-resorptive activity whilst maintaining bone formation. It reduces the risk of vertebral, hip and other non-vertebral fractures in postmenopausal women with osteoporosis. The dose is 2 g daily, given as granules dissolved in water at night. It is a useful alternative option to the bisphosphonates particularly in the frail elderly population, in whom oral bisphosphonates are often contraindicated because of difficulty in complying with the dosing regimen. Nausea, diarrhoea and headaches are infrequent side-effects, and there is a small increase in the risk of venous thromboembolism.
- *Raloxifene* 60 mg daily is a selective oestrogen-receptor modulator (SERM). It has no stimulatory effect on the endometrium but activates oestrogen receptors in bone. It prevents BMD loss at the spine and hip in postmenopausal women, though fracture rates were only reduced in the spine. It also reduces the incidence of oestrogen-receptor-positive breast carcinoma in women treated for up to 4 years. Leg cramps and flushing may occur and the risk of thromboembolic complications is also increased to a degree similar to that seen with HRT. Its use is associated with a small increase in the risk of stroke.
- *Recombinant human parathyroid hormone peptide 1–34 (teriparatide)* and *recombinant human parathyroid hormone 1–84* (anabolic agents that stimulate bone formation). *Teriparatide* has been shown to reduce vertebral and non-vertebral fractures in postmenopausal women with established osteoporosis, although data on hip fracture are not available. It is given by daily subcutaneous injection in a dose of 20 μg for 18 months. *Recombinant human parathyroid hormone* 1–84 is also administered by once daily subcutaneous injection at a dose of 100 μg and has been shown to reduce vertebral, but not non-vertebral fractures. An anti-resorptive drug may be given after parathyroid hormone peptide therapy to maintain the increase in bone mineral density. Parathyroid hormone peptide therapy is mainly indicated in severe cases of vertebral osteoporosis or in women who fail to respond to other therapies. Teriparatide causes only mild transient hypercalcaemia and routine monitoring is not required. Nausea and headache may occur. Recombinant human parathyroid hormone 1–84 is associated with a higher incidence of hypercalcaemia and hypercalciuria and routine monitoring is advised.
- *Calcium given with vitamin D* has been shown to reduce non-vertebral fractures, including hip fractures, in elderly women living in residential care (but not in free-living older community dwellers). The recommended dose is 800 IU of vitamin D and 1–1.2 g calcium daily.
- *Hormone replacement therapy* (HRT). Because of adverse effects on breast cancer and cardiovascular disease risk, HRT is a second-line option for osteoporosis except in early postmenopausal women at high fracture risk who also have perimenopausal symptoms.
- *Calcitriol* ($1,25\text{-}(OH)_2D_3$) may reduce vertebral fracture rate, although the data are inconsistent. Hypercalcaemia and hypercalciuria may occur and monitoring is necessary.
- *Calcitonin*. Nasal or subcutaneous calcitonin may be used; vertebral fracture rates may be reduced, although this has not been a universal finding.
- *Combination therapies*, either with two anti-resorptive agents or an anti-resorptive and an anabolic agent, often produce larger increases in BMD than monotherapy but have not been shown to result in greater fracture reduction than with monotherapy.

Glucocorticoid-induced osteoporosis

Individuals requiring continuous oral glucocorticoid therapy for 3 months or more (at any dose) should be assessed for coexisting risk factors (age, previous fracture, hormone status). Men and women aged over 65 years and individuals who have sustained a fragility fracture aged 40 years or more should receive treatment without delay for DXA scanning. DXA results guide treatment for other patients. Bisphosphonates (etidronate, alendronate and risedronate) and teriparatide are the approved agents for this indication. Calcium and vitamin D supplementation should also be given (see Table 10.23).

Osteoporosis in men

Alendronate (10 mg daily) and risedronate (5 mg daily) have been shown to increase BMD and to reduce vertebral fractures in men with osteoporosis. Teriparatide is also used with osteoporosis and increased risk of fracture. In men with osteoporosis who have clinical and biochemical evidence of hypogonadism, testosterone replacement is used.

OSTEONECROSIS

This is also known as aseptic avascular or ischaemic necrosis of the bone. It is due to a multitude of factors including haemoglobinopathies (e.g. sickle cell disease), deep sea diving (Caisson's disease), medication (e.g. corticosteroids, bisphosphonates), endocrine or metabolic disorders (e.g. Cushing's, diabetic mellitus), trauma, HIV infection, irradiation and alcohol abuse. Osteonecrosis of the jaw is rarely seen in patients on nitrogen-containing bisphosphonates, e.g. alendronic acid for malignant disease.

It causes symptoms similar to arthritis with pain and arthropathy, although it can be asymptomatic. If undiagnosed, it can lead to bone collapse.

MRI best confirms the diagnosis by showing bone marrow oedema. If advanced, it can be seen on plain X-rays. *Treatment* depends on the cause and preventative measures should always be taken.

PAGET'S DISEASE

Osteitis deformans or Paget's disease is a focal disorder of bone remodelling. The initial event of excessive resorption is followed by a compensatory increase in new bone formation, increased local bone blood flow and fibrous tissue in adjacent bone marrow. Ultimately, formation exceeds resorption but the new bone is structurally abnormal.

Epidemiological studies are difficult because most affected individuals are asymptomatic. Paget's disease is most often seen in Europe and particularly in northern England. It affects men and women (2:3) over the age of 40 years. The incidence approximately doubles per decade thereafter, with up to 10% of individuals radiologically affected by the age of 90. A positive family history is noted in about 14%.

Aetiology and pathogenesis

A number of genes have been implicated in Paget's disease, including nuclear factor kappa B (NFκB), sequestosome p62 (which results in activation of NFκB), osteoprotegerin and *6Cl2* (an anti-apoptotic gene). Intracellular inclusions in the osteoclasts in pagetic lesions are believed to be paramyxovirus nucleocapsid (e.g. canine distemper virus, measles or respiratory syncytial virus). However, similar microfilaments are seen in other bone disorders, and theories of a viral aetiology in Paget's remain contentious. Altered expression of *c-fos* (an oncogene) is one suggested mechanism linking viral infection with the pathogenic changes in osteoclasts, which are more numerous and contain an increased number of nuclei (up to 100). Increased osteoclastic bone resorption is followed by formation of woven bone, which is weaker than normal bone, which leads to deformity and increased fracture risk. Unaffected bone remains normal throughout life (i.e. Paget's disease does not spread, but can become symptomatic at previously silent sites).

Clinical features (Fig. 10.41)

Most (60–80%) patients with radiologically identified Paget's disease are asymptomatic. Diagnosis often follows the finding of an asymptomatic elevation of serum alkaline phosphatase, or a plain X-ray performed for other indications. The disease may involve one bone (monostotic, in 15%) or many (polyostotic). The most common sites in order of frequency are pelvis, lumbar spine, femur, thoracic spine, sacrum, skull and tibia. Small bones of the feet and hands are rarely involved.

Symptoms include the following:

- bone pain, most often in the spine or the pelvis
- joint pain when an involved bone is close to a joint, leading to cartilage damage and osteoarthritis
- deformities, in particular bowed tibia and skull changes
- complications from:
  - **(a)** nerve compression (deafness from VIIIth cranial nerve involvement; also cranial nerves II, V, VII may be involved; spinal stenosis, hydrocephalus)
  - **(b)** increased bone blood flow (myocardial hypertrophy and high-output cardiac failure)
  - **(c)** pathological fractures
- osteogenic sarcoma in pagetic bone (fewer than 1% of cases, but a 30-fold increased risk compared with non-pagetic patients).

Investigations

- *X-ray* features (Fig. 10.41b) vary from predominantly lytic lesions (osteoporosis circumscripta in the skull is characteristic), through a mixed phase, to a mainly sclerotic phase of bone expansion, thickening of trabeculae and loss of distinction between cortex and trabeculae (dedifferentiation).
- *Isotope bone scans* show the extent of skeletal involvement, but are unable to distinguish between Paget's disease and sclerotic metastatic carcinoma (especially breast and prostate).
- *Increased serum alkaline phosphatase* with normal serum calcium and phosphate reflects increased bone turnover. Levels may be normal with limited or monostotic Paget's. Levels are reduced with treatment and increase during relapse. Mild hypercalcaemia follows immobilization only when there is very extensive disease. *Urinary hydroxyproline* excretion is increased and may also be used as a marker of disease activity.

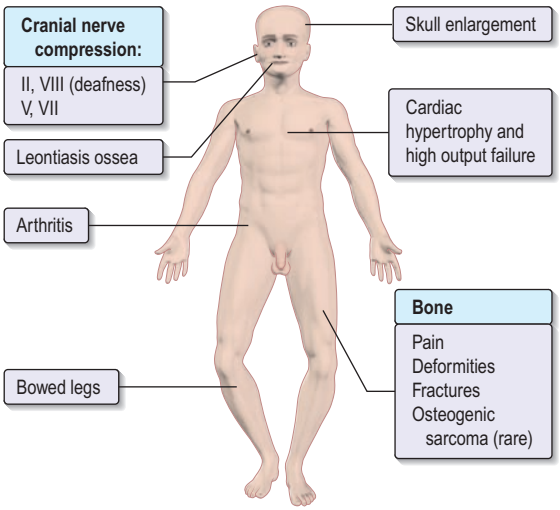
Osteonecrosis
Paget's disease

FURTHER READING

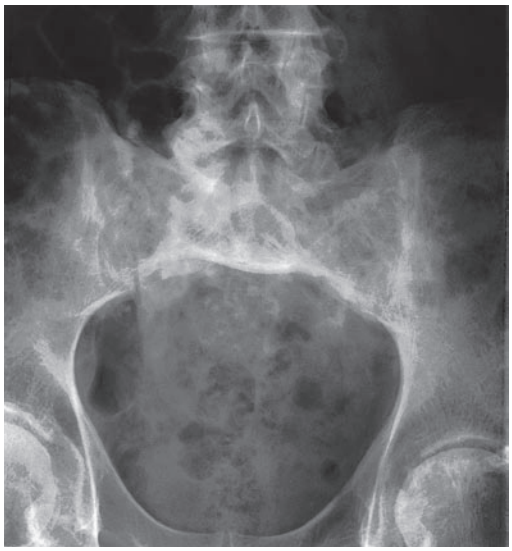
Kanis J, Oden A, Johnell O et al. The use of clinical risk factors enhances the performance of bone mineral density in prediction of hip and osteoporotic fractures in men and women. *Osteoporosis International* 2007; **18**(8): 1033–1046.

Poole KES, Compston JE. Osteoporosis and its management. *British Medical Journal* 2006; **333**: 1251–1256.

Seeman E, Delmas P. Bone quality – the material and structural basis of bone strength and fragility. *New England Journal of Medicine* 2006; **354**: 2250–2261. www.shef.ac.uk/FRAX



(a)



(b)

(c)

Fig. 10.41 **Paget's disease.**
(a) Clinical features.
(b) X-ray appearance of the pelvis, showing osteolytic and osteosclerotic lesions.
(c) Legs showing bowing of the tibia caused by increased bone growth. Note the erythema ab igne on the medial aspect of the thigh.

FURTHER READING

Whyte MP. Paget's disease of bone. *New England Journal of Medicine* 2006; **335**: 595–600.

Treatment

Bisphosphonates are the mainstay of treatment. New bone formed after treatment is lamellar, not woven (reflecting normalization of bone turnover rather than a direct effect on osteoblasts). In addition to treating symptomatic patients (i.e. bone pain or pain in joints adjacent to pagetic bone), treatment of asymptomatic lesions is appropriate if there is a significant risk of potential complications, e.g. fracture in weight-bearing long bones or the spine, nerve entrapment or deafness with skull involvement, and before orthopaedic procedures in involved bone (to reduce vascularity).

Intravenous bisphosphonates

Intravenous pamidronate has been largely replaced by the more potent zoledronate. Repeat courses of treatment are guided by symptoms and by recurrence in elevation of alkaline phosphatase or urinary hydroxyproline. Both drugs are amino-bisphosphonates and can be associated with a first-dose reaction characterized by 'flu-like' symptoms including transient pyrexia over 24–48 hours.

Oral bisphosphonates

Oral bisphosphonates are typically used at doses higher than those for osteoporosis. Options include:

- tiludronate (400 mg daily for 12 weeks, repeated as necessary after 6 weeks or longer)
- ibandronate (30 mg daily for 2 months repeated if necessary after at least 2 months)
- risedronate (30 mg/day for 2 months, repeated as necessary after at least 2 months)
- etidronate is also approved but is less often used because of its lower potency.

Surgery

Joint replacement or osteotomy is sometimes necessary to correct deformity or pain due to associated degenerative joint disease (osteoarthrosis). Intra-articular injection of lidocaine can be useful to differentiate joint or bone disease. Neurosurgery may be required where there is spinal disease. Osteosarcoma usually requires amputation, though wide excision and limb-salvage can be successful at distal sites.

RICKETS AND OSTEOMALACIA

Rickets (in children) and osteomalacia (in adults) result from inadequate mineralization of bone matrix (osteoid). They are usually caused by a defect in vitamin D availability or metabolism.

Pathology

In children the growth plate is elongated with distortion of the arrangement of chondrocytes. Calcification is delayed and vascularization impaired. In adults, osteomalacia is characterized on bone biopsy by increased osteoid width (>15 μm), increased mineralization lag time, and lack of uptake of double tetracycline labelling in osteoid seams. Changes of secondary hyperparathyroidism are often present, except in hypophosphataemic osteomalacia.

Aetiology (Table 10.25)

Vitamin D deficiency is usually due to inadequate sunlight exposure and/or dietary insufficiency (Fig. 10.34). Asian women and elderly individuals, particularly those who are housebound, are particularly at risk. Anticonvulsant therapy (especially phenytoin and phenobarbitone) affects vitamin D metabolism and predisposes to vitamin D deficiency. Malabsorption of vitamin D occurs in some gastrointestinal diseases, particularly coeliac disease and small intestinal Crohn's disease.

Chronic kidney disease results in reduced 1α-hydroxylation of 25OHD. Mutations in the *p*450c1 α gene (chromosome 12) cause 1α-hydroxylase deficiency, also known as *vitamin D-dependent rickets type I (VDDR type 1)*, an autosomal recessive disease. In *type II vitamin D-dependent rickets* there is a defect in the intracellular $1,25(OH)_2D_3$ receptor.

X-linked hypophosphataemic rickets, results in lower-limb deformities and stunted growth rate in affected males, (see p. 670). In the Fanconi syndrome and in renal tubular acidosis, osteomalacia can also develop. Phosphaturia is the common pathway underlying these conditions, caused by mutations in the *PEX* gene or excess FGF-23 production by tumours in hypophosphataemic osteomalacia. In the latter, treatment of the underlying (usually mesenchymal type) tumour results in remineralization of bone. However, location of the tumour is often difficult, and treatment with phosphate and high-dose vitamin D may be necessary.

Table 10.25 Causes of rickets/osteomalacia

Deficient intake or absorption of vitamin D
Inadequate synthesis in skin
Low dietary intake
Malabsorption:
Coeliac disease
Intestinal resection
Defective 25-hydroxylation
Chronic cholestasis (e.g. primary biliary cirrhosis)
Anticonvulsant therapy e.g. phenytoin
Defective 1-alpha hydroxylation
Chronic kidney disease
Tubular disorders (renal tubular acidosis, Fanconi's syndrome)
Vitamin D-dependent rickets types I and II
Inhibitors of mineralization
Fluoride, aluminium, bisphosphonates
Miscellaneous
X-linked hypophosphataemia (vitamin D-resistant rickets or Dent's disease)
Tumour-induced hypophosphataemic osteomalacia and rickets)
Enzyme deficiency e.g. hypophosphatasia from mutations in alkaline phosphatase gene

Clinical features

Adult osteomalacia produces vague symptoms of bone or muscle pain and tenderness. Pathological fractures occur. Occasionally a marked proximal myopathy leads to a characteristic 'waddling' gait. Deformity is uncommon. In modern practice many cases are detected biochemically in high-risk patients, especially the elderly and those with gastrointestinal disease or surgery, before clear symptoms are present. Occasionally, tetany or other hypocalcaemic features may occur.

At birth, neonatal rickets may present as craniotabes (thin deformed skull). In the first few years of life there may be widened epiphyses at the wrists and beading at the costochondral junctions, producing the 'rickety rosary', or a groove in the rib cage (Harrison's sulcus). In older children, lower limb deformities are seen. A myopathy may also occur. Hypocalcaemic tetany may occur in severe cases.

Investigations

- **Increased serum alkaline phosphatase**, indicating increased osteoblast activity, is the most common abnormality (note: alkaline phosphatase is elevated during pubertal skeletal growth).
- **Plasma calcium** is usually normal, in association with secondary hyperparathyroidism and a raised PTH, but may be low in severe cases.
- **Serum phosphate** may be low, owing to increased PTH-dependent phosphaturia, though this is variable.
- **Serum 25OHD** is usually low (the exception being vitamin D-resistant rickets). Serum $1,25(OH)_2D$ levels are also usually low, especially in VDDR type 1.
- **X-rays** are often normal in adults, but may show defective mineralization, especially in the pelvis, long bones and ribs, with pseudofractures or 'Looser's zones' – linear areas of low density surrounded by sclerotic borders.
- **Iliac crest biopsy** with double tetracycline labelling (see above) is occasionally necessary if biochemical tests are equivocal.
- **Serum fibroblast FGF-23** is sometimes elevated in tumour-associated osteomalacia.

Treatment

Treatment should be directed towards correction of the cause where possible, with increase in vitamin D intake and sunlight exposure.

Multiple formulations of vitamin D and its metabolites are available. When deficiency is nutritional, 'replacement' doses of the native vitamin are needed (400–800 IU daily). Higher 'pharmacological' doses, sometimes administered parenterally, may be needed in some patients with gastrectomy, malabsorption or liver disease, although many respond to more conventional doses. Treatment with calcitriol or alfacalcidol (see p. 633) is indicated where there is defective

FURTHER READING

Wharton B, Bishop N. Rickets. *Lancet* 2003; **362:** 1389–1400.

1α-hydroxylation, e.g. chronic kidney disease, vitamin D dependency and hypophosphataemic rickets with osteomalacia.

Monitoring of serum calcium, alkaline phosphatase and renal function should be undertaken regularly to screen for hypercalcaemia. Normalization of alkaline phosphatase is a good measure of healing.

SKELETAL DYSPLASIAS

These include a large group of heterogeneous disorders of bone and connective tissue.

Collagen defects

Collagen is responsible for many of the structural, tensile and load-bearing properties in the various tissues where it is found. The structure of collagen is discussed on page 500. Thirty or more dispersed genes encode for more than 19 different types of collagen:

- The fibrillar collagens, e.g. types I, II, III, V and XI, are encoded by *COL1A1–2, COL2A1, COL3, COL5, COL11*. Mutations of these genes produce osteogenesis imperfecta and Ehlers–Danlos syndrome.
- Basement membrane collagen, type IV, is encoded by *COL4A1–5*. Mutations lead to Alport's disease (see p. 594).
- Fibril-associated collagens with interrupted triple helices (FACIT), e.g. types IX, XII and XIV, are encoded by *COL9, COL12*.
- Filament-producing collagen type VI is encoded by *COL6A1, 2, 3*.
- Network-forming collagens types VIII and X are encoded by *COL8A1, COL10A1*.
- Anchoring fibril collagen, e.g. type VII, is encoded by *COL7A1*. Mutations of this gene produce epidermolysis bullosa (p. 1257).

Ehlers–Danlos syndrome

This is a heterogeneous group of disorders of collagen. Ten different types have been recognized with varying degrees of skin fragility, skin hyperextensibility and joint hypermobility.

Types I, II and III are inherited in an autosomal dominant fashion; the biochemical basis is unknown. No abnormalities in *COL1A1, COL1A2* and *COL2A1* genes have been found.

Type IV (vascular type) is also autosomal dominant and involves arteries, the bowel and uterus, as well as the skin. Mutations in *COL3A1* gene produce abnormalities in structure, synthesis or secretion of type III collagen.

Type VI is a recessively inherited disorder and results from a mutation in the gene that encodes lysyl hydroxylase.

Type VII is an autosomal dominant disorder where there is a defect in the conversion of procollagen to collagen; *COL1A1* and *COL1A2* mutations delete the N-proteinase cleavage sites.

Other forms of Ehlers–Danlos are very rare and their defects have not been elucidated. The clinical features are described on page 1263.

Osteogenesis imperfecta (fragilitas ossium, brittle bone syndrome)

This is a heterogeneous group of mainly autosomally dominant inherited disorders. In the majority there are mutations in the genes encoding the chains in type I collagen, i.e. *COL1A1, COL1A2*. There are four main types with very variable clinical pictures ranging from death in the perinatal period (type II), severe bone deformity (type III), to a normal lifespan (types I and IV). Recently three separate clinical subtypes have been identified (V, VI and VII).

The major clinical feature is very fragile and brittle bones but other collagen-containing tissues are also involved, such as tendons, the skin and the eyes. Osteogenesis imperfecta tarda (type I) has mild bony deformities, blue sclerae, defective dentine, early-onset deafness, hypermobility of joints, and heart valve disorders. More severe forms present with multiple fractures and gross deformities.

Treatment is with bisphosphonates, which enhances bone cortical thickness. Prognosis is variable, depending on the severity of the disease. Stem cell therapy is being used.

Miscellaneous defects

Osteopetrosis (marble bone disease)

This condition may be inherited in either an autosomal dominant or an autosomal recessive manner; the recessive type is severe and the dominant type is mild. In addition, another recessive form associated with renal tubular acidosis is due to carbonic anhydrase II deficiency.

In the severe form, which is due to a mutation in the gene, encoding for a chloride channel is necessary for osteoclast activity. Bone density is increased throughout the skeleton but bones tend to fracture easily. Involvement of the bone marrow leads to a leucoerythroblastic anaemia. There is mental retardation and early death.

In the mild form there may be only X-ray changes, but fractures and infection can occur. The acid phosphate level is raised. Stem cell transplantation has been successful.

Marfan's syndrome

This is described on page 780.

Fibroblast growth factor receptor defect

Achondroplasia ('dwarfism') is diagnosed in the first years of life. The disease is inherited in an autosomal dominant manner and is caused by a defect in the fibroblast growth factor receptor-3 gene. The trunk is of normal length but the limbs are very short and broad due to abnormal endochondrial ossification. The vault of the skull is enlarged, the face is small and the nose bridge is flat. Intelligence is normal.

FURTHER READING

Rauch F, Glorieux FH. Osteogenesis imperfecta. *Lancet* 2004; **363**: 1377–1385.

Tolar J, Teitelbaum SL, Orchard PJ. Osteopetrosis. *New England Journal of Medicine* 2004; **351**: 2839–2849.

CHAPTER BIBLIOGRAPHY

Favus MJ (ed.). *Primer on the Metabolic Bone Diseases and Disorders of Mineral Metabolism,* 6th edn. Philadelphia: Lippincott Raven, American Society for Bone and Mineral Research 2006.

Scriver CR, Beaudet AL, Sly WS, Valle D. *The Metabolic Basis of Inherited Disease,* 8th edn. New York: McGraw-Hill 2000.

SIGNIFICANT WEBSITES

http://www.rheumatology.org
American College of Rheumatology

http://www.arc.org.uk
UK Arthritis Research Campaign

http://www.rheumatology.org.uk/
British Society of Rheumatology – useful, patient-oriented information

http://www.nos.org.uk/
UK National Osteoporosis Society – useful information and reviews of ongoing research

http://www.osteo.org/
US National Institute of Health's bone diseases page, with useful links from there for osteoporosis, Paget's and osteomalacia

http://www.cbcu.cam.ac.uk/calreviews
Cambridge University site – links to quizzes, pathology images, etc.

http://www.rcplondon.ac.uk
Royal College of Physicians Guidelines

http://www.catchword.com
Clinical Medicine journal

This page intentionally left blank

11 Renal disease

Functional anatomy 571

Renal function 572
Physiology 572
Glomerular filtration rate (GFR) 574
Tubular function 575
Endocrine function 576

Investigations 578
Examination of the urine 578
Blood and quantitative tests 579
Imaging techniques 579
Transcutaneous renal biopsy 582

Glomerular diseases 582

Glomerulopathies 583
Nephrotic syndrome 584
Acute glomerulonephritis (acute nephritic syndrome) 592
Asymptomatic urinary abnormalities 593
Rapidly progressive glomerulonephritis (RPGN) 595
Other glomerular disorders 597

Renal involvement in other diseases 598

Urinary tract infection 599

Tubulointerstitial nephritis (TIN) 604
Hyperuricaemic nephropathy 606

Hypertension and the kidney 606

Other vascular disorders of the kidney 609

Renal calculi and nephrocalcinosis 609

Urinary tract obstruction 614

Drugs and the kidney 617

Acute renal failure (ARF)/acute kidney injury (AKI) 619

Chronic kidney disease (CKD) 625
Complications of chronic kidney disease 628
Progression of chronic renal impairment 631
Management of chronic kidney disease 632
Early referral of patients with chronic kidney disease 634
Renal replacement therapy 634

Cystic renal disease 642

Tumours of the kidney and genitourinary tract 644
Malignant renal tumours 644
Urothelial tumours 645

Diseases of the prostate gland 645

Renal disease in the elderly 646

FUNCTIONAL ANATOMY

The kidneys are paired organs, 11–14 cm in length in adults, 5–6 cm in width and 3–4 cm in depth. The kidneys lie retroperitoneally on either side of the vertebral column at the level of T12 to L3. The renal parenchyma comprises an outer cortex and an inner medulla. The functional unit of the kidney is the nephron, of which each contains approximately one million. Each *nephron* is made up of a glomerulus, proximal tubule, loop of Henle, distal tubule and collecting duct. The renal capsule and ureters are innervated via T10–12 and L1 nerve roots, and renal pain is felt over the corresponding dermatomes.

Arterial blood is supplied to the kidneys via the renal arteries, which branch off the abdominal aorta, and venous blood is conveyed to the inferior vena cava via the renal veins. Approximately 25% of humans possess dual or multiple renal arteries on one or both sides. The left renal vein is longer than the right and for this reason the left kidney, where possible, is usually chosen for live donor transplant nephrectomy.

The *renal artery* undergoes a series of divisions within the kidney (Fig. 11.1) forming successively the *interlobar arteries*, which run radially to the corticomedullary junction, *arcuate arteries*, which run circumferentially along the corticomedullary junction, and interlobular arteries which run radially through the renal cortex towards the surface of the kidney. Afferent glomerular arterioles arise from the interlobular arteries to supply the glomerular capillary bed, which drains into *efferent glomerular arterioles*. Efferent arterioles from the outer cortical glomeruli drain into a peritubular capillary network within the renal cortex and then into increasingly large and more proximal branches of the renal vein. By contrast, blood from the juxtamedullary glomeruli passes via the vasa recta in the medulla and then turns back towards the area of the cortex from which the vasa recta originated.

Vasa recta possess fenestrated walls, which facilitates movement of diffusible substances. The collecting ducts merge in the inner medulla to form the ducts of Bellini, which empty at the apices of the papillae into the calyces. The calyces, in common with the renal pelvis, ureter and bladder, are lined with transitional cell epithelium.

The *glomerulus* comprises four main cell types: (1) endothelial cells which are fenestrated with 500–1000 Å pores; (2) visceral epithelial cells (podocytes) which support the delicate glomerular basement membrane by means of an extensive trabecular network (foot processes); (3) parietal epithelial

Ask the authors

www.studentconsult.com

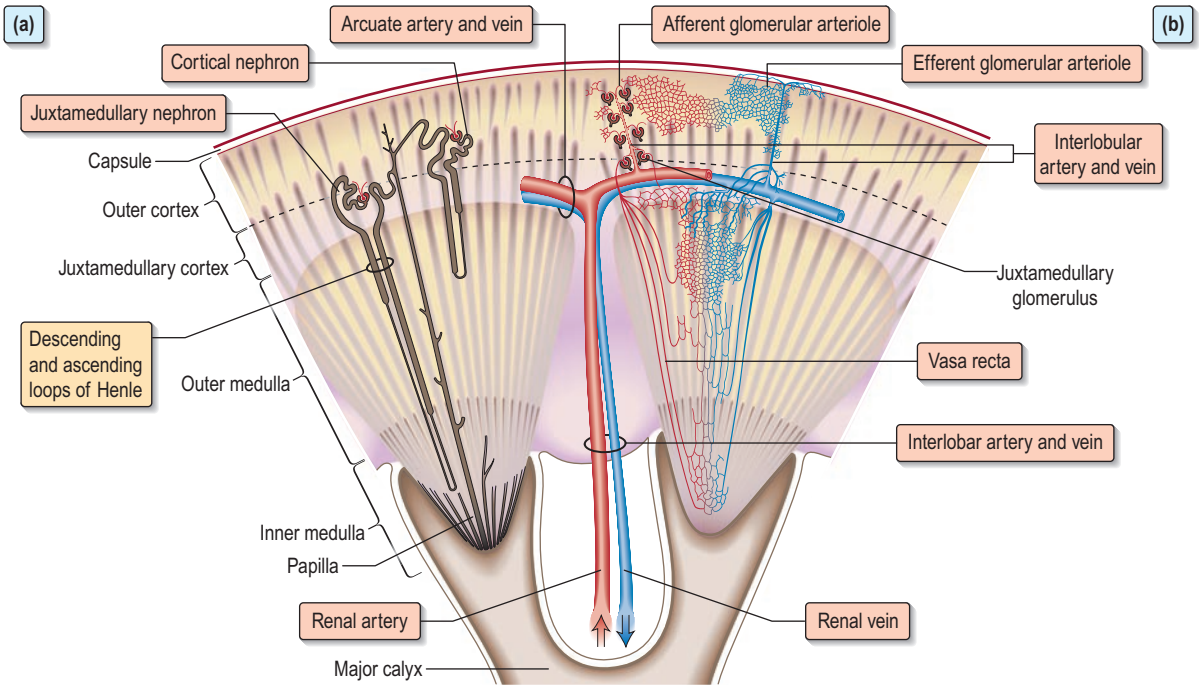


Fig. 11.1 **Functional anatomy of the kidney. (a)** The nephrons. **(b)** Arterial and venous supply. After Williams PL (ed.) (1995) Gray's Anatomy, 38th edn. Edinburgh: Churchill Livingstone, with permission from Elsevier.

cells which cover the Bowman's capsule; (4) mesangial cells (see Fig. 11.11). Mesangial cells are believed to be related to macrophages of the reticuloendothelial system and have a phagocytic function and contractile capabilities that can control blood flow and filtration surface area along the glomerular capillaries in response to a host of mediators. They also secrete the mesangial matrix, which provides a skeletal framework for the glomerular capillaries. The glomerular capillary basement membrane lies between the endothelial and the visceral epithelial cells. The latter put out multiple long foot processes which interdigitate with those of adjacent epithelial cells. Together the endothelial cells, basement membrane and epithelial cells form the filtration barrier or sieve (see Fig. 11.12a).

The *renal tubules* are lined by epithelial cells, which are cuboidal except in the thin limb of the loop of Henle where they are flat. Proximal tubular cells differ from other cells of the system as they have a luminal brush border. The cortical portion of the collecting ducts contains two cell types with different functions, namely principal cells and intercalated cells (see p. 653). Fibroblast-like cells in the renal cortical interstitium have been shown to produce erythropoietin in response to hypoxia (p. 576).

The *juxtaglomerular apparatus* comprises the macula densa, the extraglomerular mesangium and the terminal portion of the afferent glomerular arteriole (which contains renin-producing granular cells) together with the proximal portion of the efferent arteriole. The macula densa is a plaque of cells containing large, tightly packed cell nuclei (hence the name macula densa) within the thick ascending limb of the loop of Henle. This anatomical arrangement is such as to allow changes in the renal tubule to influence behaviour of the adjacent glomerulus (tubulo-glomerular feedback).

RENAL FUNCTION

PHYSIOLOGY

A conventional diagrammatic representation of the nephron is shown in Figure 11.2a and a physiological version in Figure 11.2b.

An essential feature of renal function is that a large volume of blood – 25% of cardiac output or approximately 1300 mL per minute – passes through the two million glomeruli.

A hydrostatic pressure gradient of approximately 10 mmHg (a capillary pressure of 45 mmHg minus 10 mmHg of pressure within Bowman's space and 25 mmHg of plasma oncotic pressure) provides the driving force for ultrafiltration of virtually protein-free and fat-free fluid across the glomerular capillary wall into Bowman's space and so into the renal tubule (Fig. 11.3).

The *ultrafiltration rate* (glomerular filtration rate; GFR) varies with age and sex but is approximately 120–130 mL/min per 1.73 m^2 surface area in adults. This means that, each day, ultrafiltration of 170–180 L of water and unbound small-molecular-weight constituents of blood occurs. If these large volumes of ultrafiltrate were excreted unchanged as urine, it would be necessary to ingest huge amounts of water and electrolytes to stay in balance. This is avoided by the selective reabsorption of water, essential electrolytes and other blood constituents, such as glucose and amino acids, from the filtrate in transit along the nephron. Thus, 60–80% of filtered water and sodium are reabsorbed in the proximal tubule along with virtually all the potassium, bicarbonate, glucose and amino acids (Fig. 11.2b). Additional water and sodium chloride are reabsorbed more distally, and fine tuning of salt and water balance is achieved in the distal tubules

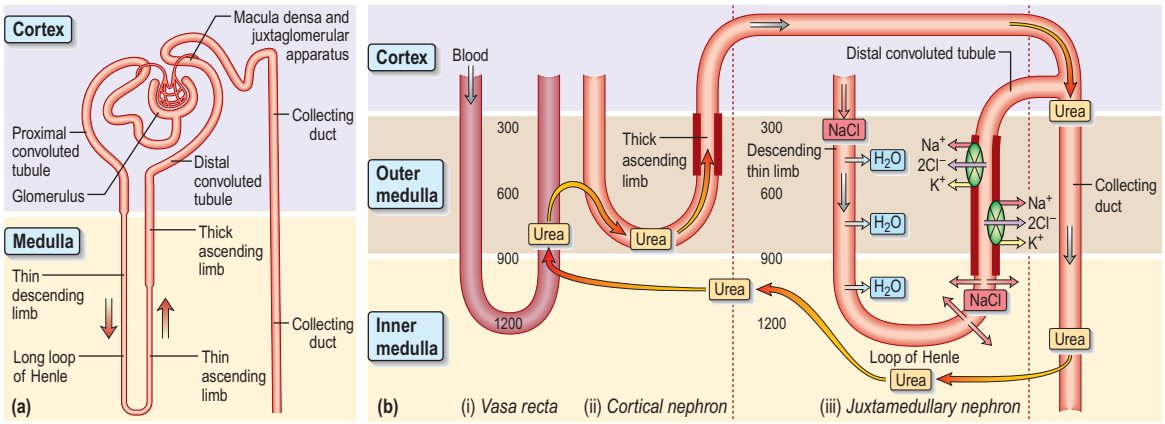


Fig. 11.2 **(a) Principal parts of the nephron.** The point where the distal tubule is in close proximity to its own glomerulus is called the juxtaglomerular apparatus. This contains the macula densa.
(b) Diagrammatic representation of the countercurrent system. (i) *Vasa recta:* these vessels descend from the cortex into the medulla and then turn back towards the cortex. (ii) *Cortical nephron:* these have short descending limbs extending into the outer medulla. (iii) *Juxtamedullary nephron:* the descending limb dips deeply into the hypertonic inner medulla. Numbers indicate approximate osmolalities.

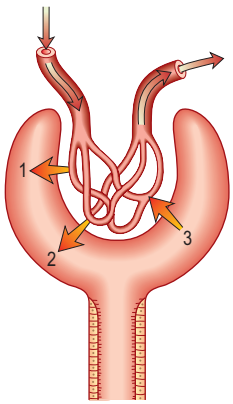


Fig. 11.3 **Pressures controlling glomerular filtration.**
1, capillary hydrostatic pressure (45 mmHg);
2, hydrostatic pressure in Bowman's space (10 mmHg);
3, plasma protein oncotic pressure (25 mmHg). Arrows (1, 2, 3) indicate the direction of a pressure gradient.

and collecting ducts under the influence of aldosterone and antidiuretic hormone (ADH). The final urine volume is thus 1–2 L daily. Calcium, phosphate and magnesium are also selectively reabsorbed in proportion to the need to maintain a normal electrolyte composition of body fluids.

The urinary excretion of some compounds is more complicated. For example, potassium is freely filtered at the glomerulus, almost completely reabsorbed in the proximal tubule, and secreted in the distal tubule and collecting ducts. A clinical consequence of this is that the ability to eliminate unwanted potassium is less dependent on GFR than is the elimination of urea or creatinine. Other compounds filtered and reabsorbed or secreted to a variable extent include urate, many organic acids and many drugs or their metabolic breakdown products. The more tubular secretion of a compound that occurs, the less dependent elimination is on the GFR; penicillin and cefradine are examples of drugs secreted by the tubules.

Urine concentration and the countercurrent system

Urine is concentrated by a complex interaction between the loops of Henle, the medullary interstitium, medullary blood vessels (vasa recta) and the collecting ducts (see p. 651). The proposed mechanism of urine concentration is termed 'the countercurrent mechanism'. The countercurrent hypothesis states that: 'a small difference in osmotic concentration at any point between fluid flowing in opposite directions in two parallel tubes connected in a hairpin manner is multiplied many times along the length of the tubes'. Tubular fluid moves from the renal cortex towards the papillary tip of the medulla via the proximal straight tubule and the thin descending limb of the loop of Henle, which is permeable to water and impermeable to sodium. The tubule then loops back towards the cortex so that the direction of the fluid movement is reversed in the ascending limb, which is impermeable to water but permeable to sodium. This results in a large osmolar concentration difference between the corticomedullary junction and the hairpin loop at the tip of the papilla, and hence countercurrent multiplication. There is an analogy with heat exchangers.

Since the urine that emerges from the proximal tubule is iso-osmotic, the first nephron segment actually involved in urinary concentration is the descending limb of Henle's loop. There are two types of descending limbs (Fig. 11.2b). The short loops originate in superficial and midcortical glomeruli, and turn in the outer medulla. The long loops, which originate in the deep cortical and juxtamedullary glomeruli, penetrate the outer medulla up to the tip of the papilla. Approximately 15% of nephrons have long loops and the remaining 85% have short loops. Both the ascending limb in the outer and inner medulla and the first part of the distal tubule are impermeable to water and urea. Through the $Na^+/K^+/2Cl^-$ cotransporter, the thick ascending limb actively transports sodium chloride, increasing the interstitial tonicity, resulting in tubular dilution with no net movement of water and urea on account of low permeability. The hypotonic fluid under ADH action undergoes osmotic equilibration with the interstitium in the late distal and the cortical and outer medullary collecting duct, resulting in water removal. Urea concentration in the

tubular fluid rises on account of low urea permeability. At the inner medullary collecting duct, which is highly permeable to urea and water, especially in response to ADH, the urea enters the interstitium down its concentration gradient, preserving interstitial hypertonicity and generating high urea concentration in the interstitium.

The hypertonic interstitium pulls water from the descending limb of the loop of Henle, which is relatively impermeable to NaCl and urea. This makes the tubular fluid hypertonic with high NaCl concentration as it arrives at the bend of the loop of Henle. Urea plays a key role in the generation of medullary interstitial hypertonicity. The urea that is reabsorbed into the inner medullary stripe from the terminal inner medullary collecting duct is carried out of this region by ascending vasa recta, which deposit urea into the adjacent descending limb of both short and long loops of Henle, thus recycling the urea to the inner medullary collecting tubule. This process is facilitated by the close anatomical relationship that the hairpin loop of Henle and the vasa recta share.

GLOMERULAR FILTRATION RATE (GFR)

In health, the GFR remains remarkably constant owing to intrarenal regulatory mechanisms. In disease, (e.g. a reduction in intrarenal blood flow, damage to or loss of glomeruli, or obstruction to the free flow of ultrafiltrate along the tubule), the GFR will fall. The ability to eliminate waste material and to regulate the volume and composition of body fluid will decline. This will be manifest as a rise in the plasma urea or creatinine and as a reduction in measured GFR.

The concentration of urea or creatinine in plasma represents the dynamic equilibrium between production and elimination. In healthy subjects there is an enormous reserve of renal excretory function, and serum urea and creatinine do not rise above the normal range until there is a reduction of 50–60% in the GFR. Thereafter, the level of urea depends both on the GFR and its production rate (Table 11.1). The latter is heavily influenced by protein intake and tissue catabolism. The level of creatinine is much less dependent on diet but is more related to age, sex and muscle mass. Once it is elevated, serum creatinine is a better guide to GFR than urea and, in general, measurement of serum creatinine is a good way to monitor further deterioration in the GFR.

It must be re-emphasized that a normal serum urea or creatinine is not synonymous with a normal GFR.

Table 11.1 Factors influencing serum urea levels

Production	Elimination
Increased by	**Increased by**
High-protein diet	Elevated GFR, e.g. pregnancy
Increased catabolism	
Surgery	**Decreased by**
Infection	Glomerular disease
Trauma	Reduced renal blood flow
Corticosteroid therapy	Hypotension
Tetracyclines	Dehydration
Gastrointestinal bleeding	Urinary obstruction
Cancer	Tubulointerstitial nephritis
Decreased by	
Low-protein diet	
Reduced catabolism, e.g. old age	
Liver failure	

GFR, glomerular filtration rate

Measurement of the glomerular filtration rate

Measurement of the GFR is necessary to define the exact level of renal function. It is essential when the serum (plasma) urea or creatinine is within the normal range. The most widely used measurement is the creatinine clearance (Fig. 11.4).

Creatinine clearance is dependent on the fact that daily production of creatinine (principally from muscle cells) is remarkably constant and little affected by protein intake. Serum creatinine and urinary output thus vary very little throughout the day.

Creatinine excretion is, however, by both glomerular filtration and tubular secretion, although at normal serum levels the latter is relatively small. As most laboratory methods for measurement of serum creatinine give slight overestimates, the calculation of clearance fortuitously gives a value close to that of inulin.

With progressive renal failure, creatinine clearance may overestimate GFR but, in clinical practice, this is seldom significant. Certain drugs – for example cimetidine, trimethoprim, spironolactone and amiloride – reduce tubular secretion of creatinine, leading to a rise in serum creatinine and a fall in measured clearance.

Given these observations, creatinine clearance, nevertheless, is a reasonably accurate measure of GFR in those situations in which it is most required – **normal or near normal renal function.** Urine is collected over 24 hours for measurement of urinary creatinine. A single plasma level of creatinine is measured some time during the 24-hour period.

$$\text{Creatinine clearance} = U \times V/P$$

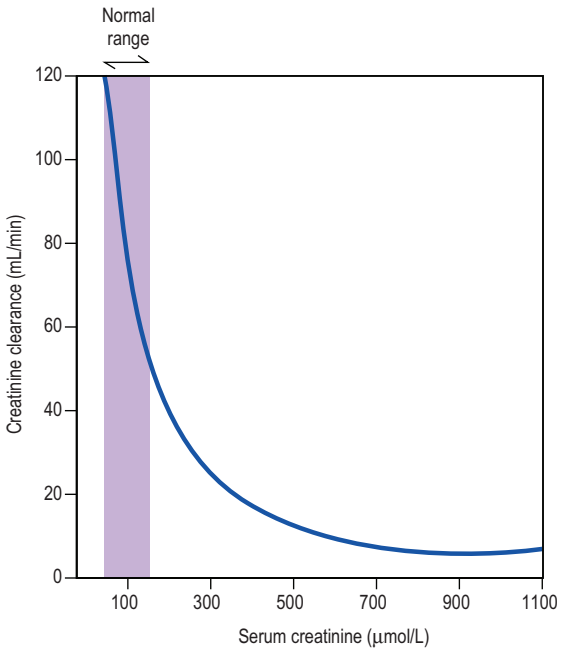


Fig. 11.4 **Creatinine clearance versus serum creatinine.** Note that the serum creatinine does not rise above the normal range until there is a reduction of 50–60% in the glomerular filtration rate (creatinine clearance).

where U = urine concentration of creatinine; V = rate of urine flow in mL/min; P = plasma concentration of creatinine. Normal ranges: men 90–140 mL/min; women 80–125 mL/min.

Where urine collections are difficult (e.g. with ileal conduits) or deemed inaccurate, the GFR may be measured by the single injection of compounds such as [^{51}Cr]EDTA (ethylenediaminetetraacetic acid), [^{99m}Tc]DTPA (diethylenetriaminepentaacetic acid) or [^{125}I]iothalamate, their excretion being primarily by glomerular filtration. Following intravenous injection of the compound, three blood samples are obtained at 2, 3 and 4 hours (or rather longer intervals if the patient is oedematous or if renal failure is suspected). The GFR may then be calculated from the slope of the exponential fall in blood level of the compound. *Iohexol* (a low-osmolality, non-ionic contrast medium) is also used as it is non-radioactive and can be measured reliably by a routine biochemical HPLC (high-pressure liquid chromatography) technique. Iohexol clearance correlates with radioisotopic techniques and can be assessed from finger-prick blood samples.

Cystatin C is an endogenous cysteine protease inhibitor. It is a low-molecular-weight protein filtered freely by the glomerulus and is produced at a constant rate. It is not secreted by tubules and therefore correlates with GFR. The cystatin C level is unaffected by gender, and correction for surface area is not needed. Studies suggest that cystatin C levels correlate better with radioisotope estimation of GFR than do creatinine levels and are more sensitive to mild renal impairment, particularly when the creatinine level is still normal.

Calculated GFR. Measurement of true GFR is cumbersome, time consuming and may be inaccurate if 24-hour urine collections are incomplete. Therefore, several formulae have been developed that allow a prediction of creatinine clearance or GFR from serum creatinine and demographics. The Cockroft–Gault formula is used mostly to adjust the dose of drugs by physicians and pharmacists and takes into account the patient's age, gender, weight and plasma creatinine (Box 11.1). A prediction equation has been developed based on the data derived from the Modification of Diet in Renal Disease (MDRD) study in patients with chronic renal failure (Box 11.1). This equation is based on age, sex, creatinine and ethnicity which makes it more reliable than the Cockroft–Gault equation in ethnically diverse groups of individuals. It is very popular in large population-based epidemiological studies where weight of the patient is usually not available. A modification of MDRD equation is used by most chemical pathology laboratories to calculate eGFR but it is less reliable if actual GFR is ≤60 mL/min and can result in inappropriate referral to renal physicians. Another clinical practice guideline formula is produced by the National Kidney Foundation. All these equations have not, however, been validated across all ranges of renal impairment, weights or body mass index (BMI), or ethnic groups; this makes them unreliable in the monitoring of patients with acute or chronic kidney renal disease whilst being treated and many clinicians still rely on measured creatinine clearance.

Box 11.1 Estimation of creatinine clearance

Cockroft–Gault equation

$$\text{Creatinine clearance} = \frac{(140 - \text{age}) \times \text{weight (kg)} \times (\text{constant})}{\text{Serum creatinine } [\mu\text{mol/L}]}$$

Constant = 1.23 for males and 1.04 for women.

Modification of diet in renal disease (MDRD) equation

Calculation of Estimated GFR by 4 variables:

$$\text{Estimated GFR } (\text{mL/min/1.73 m}^2) = 186 \times (S_{Cr})^{-1.154} \times (\text{age})^{-0.203} \times \text{constant } (0.742 \text{ if female}) \times (1.210 \text{ if Black African})$$

To convert creatinine values in μmol/L to mg/dL multiply by 0.0113.

MacGregor MS, Boag DE, Innes A. Chronic kidney disease. *Quarterly Journal of Medicine* 2006; **99**: 365–375.

TUBULAR FUNCTION

The major function of the tubule is the selective reabsorption or excretion of water and various cations and anions to keep the volume and electrolyte composition of body fluid normal (see Ch. 12).

The active reabsorption from the glomerular filtrate of compounds such as glucose and amino acids also takes place. Within the normal range of blood concentrations these substances are completely reabsorbed by the proximal tubule. However, if blood levels are elevated above the normal range, the amount filtered (filtered load = GFR × plasma concentration) may exceed the maximal absorptive capacity of the tubule and the compound 'spills over' into the urine. Examples of this occur with hyperglycaemia in diabetes mellitus or elevated plasma phenylalanine in phenylketonuria.

Conversely, inherited or acquired defects in tubular function may lead to incomplete absorption of a normal filtered load, with loss of the compound in the urine (a lowered 'renal threshold'). This is seen in renal glycosuria, in which there is a genetically determined defect in tubular reabsorption of glucose. It is diagnosed by demonstrating glycosuria in the presence of normal blood glucose levels. Inherited or acquired defects in the tubular reabsorption of amino acids, phosphate, sodium, potassium and calcium also occur, either singly or in combination. Examples include cystinuria and the Fanconi syndrome (see p. 1069 and Ch. 12). Tubular defects in the reabsorption of water result in nephrogenic diabetes insipidus (p. 1018). Under normal circumstances, antidiuretic hormone induces an increase in the permeability of water in the collecting ducts by attachment to receptors with subsequent activation of adenyl cyclase. This then activates a protein kinase, which induces preformed cytoplasmic vesicles containing water channels (termed 'aquaporins') to move to and insert into the tubular luminal membrane. This allows water entry into tubular cells down a favourable osmotic gradient. Water then crosses the basolateral membrane and enters the bloodstream. When the effect of ADH wears off, water channels return to the cell cytoplasm (Fig. 12.5).

Acid–base balance

Tubular function is also critical to the control of acid–base balance. Thus, filtered bicarbonate is largely reabsorbed and

hydrogen ions are excreted mainly buffered by phosphate (see p. 714).

Investigation of tubular function in clinical practice

Various tubular mechanisms could theoretically be investigated, but, in clinical practice, tests of tubular function are required less often than glomerular function.

Twenty-four-hour sodium output may be helpful in determining whether a patient is complying with a low-salt diet and in the management of salt-losing nephropathy. Tests of proximal tubular function may be required in the diagnosis of Fanconi's syndrome or isolated proximal tubular defects (e.g. urate clearance). Bicarbonate, glucose, phosphate and amino acid are all reabsorbed in the proximal tubule. Their presence in the urine is abnormal, and though formal methods of measuring maximal reabsorption are available, they are seldom necessary. *Retinol-binding protein and β_2-microglobulin* are normally reabsorbed by the proximal tubule, and their urinary excretion is non-specifically increased by diseases of the proximal tubule.

Two tests of distal tubular function are commonly applied in clinical practice: measurement of urinary concentrating capacity in response to water deprivation, and measurement of urinary acidification. These tests are dealt with on pages 1017 and 676.

Protein and polypeptide metabolism

The kidney is a major site for the catabolism of many small-molecular-weight proteins and polypeptides, including many hormones such as insulin, parathyroid hormone (PTH) and calcitonin, by endocytosis carried out by the megalin–cubilin complex in the brush border of proximal tubular cells. In renal failure the metabolic clearance of these substances is reduced and their half-life is prolonged. This accounts, for example, for the reduced insulin requirements of diabetic patients as their renal function declines.

Drug and toxicant elimination

A substantial fraction of prescription drugs are handled and eliminated by the kidney. Many of these medications (e.g. penicillins, cephalosporins, diuretics, NSAIDs, antivirals and methotrexate) circulate in the plasma as small organic anions. These organic anions, which are often bound to albumin, are actively eliminated by the proximal tubule of the nephron by an organic anion transporter (OAT) system. The OAT system translocates drugs as well as endogenous substances and toxins.

ENDOCRINE FUNCTION

Renin–angiotensin system (see also p. 1023)

The juxtaglomerular apparatus is made up of specialized arteriolar smooth muscle cells that are sited on the afferent glomerular arteriole as it enters the glomerulus. These cells synthesize prorenin, which is cleaved into the active proteolytic enzyme renin. Active renin is then stored in and released from secretory granules. Prorenin is also released in the circulation and comprises 50–90% of circulating renin, but its physiological role remains unclear as it cannot be converted into active renin in the systemic circulation. Renin converts angiotensinogen in blood to angiotensin I. Angiotensin-converting enzyme (ACE), which is located in the lung, luminal border of endothelial cells, glomeruli and other organs, converts angiotensin I (decapeptide) to angiotensin II (octapeptide). Renin release is controlled by:

- pressure changes in the afferent arteriole
- sympathetic tone
- chloride and osmotic concentration in the distal tubule via the macula densa (Fig. 11.2a)
- local prostaglandin and nitric oxide release.

Angiotensin II has two major systemic effects; systemic vasoconstriction and sodium and water retention. Both of these actions will tend to reverse the hypovolaemia or hypotension that is usually responsible for the stimulation of renin release. Angiotensin II promotes renal NaCl and water absorption by direct stimulation of Na^+ reabsorption in the early proximal tubule and by increased adrenal aldosterone secretion which enhances Na^+ transport in the collecting duct. In addition to influencing systemic haemodynamics, angiotensin II also regulates GFR. Although it constricts both afferent and efferent arterioles, vasoconstriction of efferent arterioles is three times greater than that of afferent, resulting in increase of glomerular capillary pressure and maintenance of GFR. In addition, angiotensin II constricts mesangial cells, reducing the filtration surface area, and sensitizes the afferent arteriole to the constricting signal of tubuloglomerular feedback (see p. 572). The net result is that angiotensin II has opposing effects on the regulation of GFR: (a) an increase in glomerular pressure and consequent rise in GFR; (b) reduction in renal blood flow and mesangial cell contraction, reducing filtration (see Fig. 11.49). In renal artery stenosis with resultant low perfusion pressure, angiotensin II maintains GFR. However, in cardiac failure and hypertension, GFR may be reduced by angiotensin II.

Erythropoietin (see also p. 390)

Erythropoietin is the major stimulus for erythropoiesis. It is a glycoprotein produced principally by fibroblast-like cells in the renal interstitium.

- Under hypoxic conditions both the alpha and beta subunits of hypoxia inducible factor 1 (HIF-1) are expressed forming a heterodimer, causing erythropoietin gene transcription via the combined effects of hepatic nuclear factor 4 (HNF-4) and coactivator p300. Erythropoietin, once formed, binds to its receptors on erythroid precursor cells.
- Under *normal oxygen* conditions, only the HIF-1-beta subunit is constitutively expressed. The alpha subunit undergoes proline hydroxylation in the presence of iron and oxygen by prolyl hydroxylase.
- The hydroxylated HIF-1-alpha subunit binds to von Hippel–Lindau protein and a ubiquitin ligase E3 complex is activated. This leads to ubiquitination (p. 20) and subsequent degradation of HIF-1-alpha via proteosomes so that no erythropoietin is transcribed. In normoxic conditions HIF-1-alpha also undergoes asparaginyl hydroxylation which prevents HIF complex from recruiting coactivators. These hydroxylation steps have absolute requirement for molecular oxygen which forms the basis of oxygen sensing.

Loss of renal substance, with decreased erythropoietin production, results in a normochromic, normocytic anaemia.

Conversely, erythropoietin secretion may be increased, with resultant polycythaemia, in patients with polycystic renal disease, benign renal cysts or renal cell carcinoma. Recombinant human erythropoietin has been biosynthesized and is available for clinical use, particularly in patients with chronic renal disease (see p. 633).

Vitamin D metabolism (see also p. 559)

Naturally occurring vitamin D (cholecalciferol) requires hydroxylation in the liver at position 25 and again by a 1α-hydroxylase enzyme (mitochondrial cytochrome P450) mainly in the distal convoluted tubule, the cortical and inner medullary part of the collecting ducts and the papillary epithelia of the kidney to produce the metabolically active 1,25-dihydroxycholecalciferol (1,25-$(OH)_2D_3$). The 1α-hydroxylase activity is increased by high plasma levels of parathyroid hormone (PTH), low phosphate and low 1,25-$(OH)_2D_3$. 1,25-dihydroxycholecalciferol and 25-hydroxycholecalciferol are degraded in part by being hydroxylated at position 24 by 24-hydroxylase. The activity of this enzyme is reduced by PTH and increased by 1,25-$(OH)_2D_3$ (which therefore promotes its own inactivation).

Reduced 1α-hydroxylase activity in diseased kidneys results in relative deficiency of 1,25-$(OH)_2D_3$. As a result, gastrointestinal calcium and to a lesser extent phosphate absorption is reduced and bone mineralization impaired. Receptors for 1,25-$(OH)_2D_3$ exist in the parathyroid glands, and reduced occupancy of the receptors by the vitamin alters the set-point for release of PTH in response to a given decrement in plasma calcium concentration. Gut calcium malabsorption, which induces hypocalcaemia, and relative lack of 1,25-$(OH)_2D_3$, contribute therefore to the hyperparathyroidism seen regularly in patients with renal impairment, even of modest degree.

Autocrine function

Endothelins

The endothelins ET-1, ET-2 and ET-3 are a family of similar potent vasoactive peptides that also influence cell proliferation and epithelial solute transport. They do not circulate but act locally. ETs are produced by most types of cells in the kidney. The vascular actions are mediated by two receptors, with ETA (specific for ET-1) mediating vasoconstriction and ETB (responsive to all ETs) causing vasodilatation. Endothelins inhibit sodium and water absorption by suppressing Na^+/K^+-ATPase and Na^+/H^+ antiporter activity in the proximal tubule and antagonizing the action of ADH and aldosterone in the collecting duct. Tubular transport actions are mediated by ETB. Endothelins, through vasoconstriction by ETA and salt and water retention via ETB receptors, cause hypertension. Endothelins, mainly through ETA receptors, can also alter cell proliferation and matrix accumulation by increasing tissue inhibitor of metalloproteinase (TIMP), cytokines, fibronectin and collagen. These peptides also stimulate the proliferation of a variety of renal cell types.

Prostaglandins

Prostaglandins are unsaturated, oxygenated fatty acids, derived from the enzymatic metabolism of arachidonic acid, mainly by constitutively expressed cyclo-oxygenase-1 (COX-1) or inducible COX-2 (see Fig. 14.32). COX-1 is highly expressed in the collecting duct, while COX-2 expression is restricted to the macula densa. Both COX isoforms convert arachidonic acid to the same product, the bioactive but unstable prostanoid precursor, prostaglandin H_2 (PGH_2). PGH_2 is converted to:

- PGE_2 (formed by PDE_2 synthase in the collecting duct, responsible for natriuretic and diuretic effects)
- PGD_2 (undetermined significance, produced in proximal tubule)
- prostacyclin (PGI_2) (mainly synthesized in the interstitial and vascular compartment)
- thromboxane A_2 (vasoconstrictor, mainly synthesized in glomerulus).

They all act through G-coupled transmembrane receptors, maintaining renal blood flow and glomerular filtration rate in the face of reductions induced by vasoconstrictor stimuli such as angiotensin II, catecholamines and α-adrenergic stimulation. In the presence of renal underperfusion, inhibition of prostaglandin synthesis by non-steroidal anti-inflammatory drugs results in a further reduction in GFR, which is sometimes sufficiently severe as to cause acute renal failure (acute kidney injury). Renal prostaglandins also have a natriuretic renal tubular effect and antagonize the action of antidiuretic hormone. Renal prostaglandins do not regulate salt and water excretion in normal subjects, but in some circumstances, such as chronic renal failure (chronic kidney disease), prostaglandin-induced vasodilatation is involved in maintaining renal blood flow. Patients with chronic kidney disease are thus vulnerable to further deterioration in renal function on exposure to non-steroidal anti-inflammatory drugs, as are elderly patients in many of whom renal function is compromised by renal vascular disease and/or the effects of ageing upon the kidney. Moreover, in conditions such as volume depletion, which are associated with high renin release (facilitated by prostaglandins), inhibition of prostaglandin synthesis may lead to hyperkalaemia due to hyporeninaemic hypoaldosteronism (since angiotensin II is the main stimulus for aldosterone).

Urodilatin: renal natriuretic peptide

(see also p. 1023)

A 32-amino-acid atrial natriuretic-like peptide (ANP-like peptide), synthesized by different post-translational processing of pro-ANP in the connecting and collecting ducts in the kidney, has been isolated from human urine. Its natriuretic potency exceeds that of atrial ANP by increasing cGMP production in the collecting duct. It is postulated that cardiac ANP is primarily a regulator of the cardiovascular system through its vascular effects and that renal natriuretic peptide participates in the intrarenal regulation of sodium and chloride transport. These potent natriuretic properties are being utilized by the use of synthetic urodilatin infusion in patients with cirrhosis to enhance sodium and water excretion.

Nitric oxide and the kidney

Nitric oxide (Fig. 15.10), a molecular gas, is formed by the action of three isoforms of nitric oxide synthase (NOS). All three enzymes, neuronal (nNOS or NOS1), inducible (iNOS or NOS2) and endothelial (eNOS or NOS3), which are cytochrome P450-like proteins, facilitate the addition of the guanidine nitrogen of the amino acid arginine to molecular oxygen, producing nitric oxide and water. In general nNOS and eNOS are constitutively active, producing low levels of nitric oxide dependent upon intracellular calcium elevation. In contrast, the transcriptional regulation of iNOS can be

markedly induced, particularly by inflammatory cytokines, resulting in extremely large amounts of nitric oxide. The most recognized cellular target of nitric oxide is soluble guanylate cyclase. The stimulation of this enzyme enhances the synthesis of cyclic GMP from GTP. All three isoforms are expressed in the kidney with eNOS in the vascular compartment, nNOS mainly in the macula densa and inner medullary collecting duct, and iNOS in several tubule segments. Nitric oxide mediates the following physiological actions in the kidney:

- regulation of renal haemodynamics
- natriuresis by inhibiting Na^+/K^+-ATPase and Na^+/H^+ antiporter and antagonizing ADH
- modulation of tubuloglomerular feedback so that the composition of tubular fluid delivered to the macula densa changes the filtration rate of the associated glomerulus.

FURTHER READING

Almond A, Siddiqui S, Robertson S et al. Comparison of combined urea and creatinine and prediction equations as measures of residual renal function when GFR is low. *Quarterly Journal of Medicine* 2008; **101:** 619–624.

Hao CM, Breyer MD. Physiological regulation of prostaglandins in the kidney. *Annual Review of Physiology* 2008; **70:** 357–377.

Pollock JS, Pollock DM. Endothelin and NOS1/nitric oxide signaling and regulation of sodium homeostasis. *Current Opinion in Nephrology and Hypertension* 2008; **17**(1): 70–75.

Stevens LA, Coresh J, Greene T et al. Assessing kidney function – measured and estimated glomerular filtration rate. *New England Journal of Medicine* 2006; **354**: 2473–2483.

INVESTIGATIONS

EXAMINATION OF THE URINE

Appearance

This is of little value in the differential diagnosis of renal disease except in the diagnosis of haematuria. Overt 'bloody' urine is usually unmistakable but should be checked using dipsticks (Stix testing). Very concentrated urine may also appear dark or smoky. Other causes of discoloration of urine include cholestatic jaundice, haemoglobinuria, drugs such as rifampicin, use of fluorescein or methylthioninium chloride (methylene blue), and ingestion of beetroot. Discoloration of urine after standing for some time occurs in porphyria, alkaptonuria and in patients ingesting the drug L-dopa.

Volume

In health, the volume of urine passed is primarily determined by diet and fluid intake. In temperate climates it lies within the range 800–2500 mL per 24 hours. The minimum amount passed to stay in fluid balance is determined by the amount of solute – mainly urea and electrolytes – being excreted and the maximum concentrating power of the kidneys would be approximately 650 mL.

In diseases such as chronic kidney disease (CKD) or diabetes insipidus, impairment of concentrating ability requires increased volumes of urine to be passed, given the same daily solute output. An increased solute output, such as in glycosuria or increased protein catabolism following surgery, also demands increased urine volumes.

Specific gravity and osmolality

Urine specific gravity is a measure of the weight of dissolved particles in urine, whereas urine osmolality reflects the number of such particles. Usually the relationship between the two is close. Measurement of urine specific gravity or osmolality is required only in the differential diagnosis of oliguric renal failure or the investigation of polyuria or inappropriate ADH secretion. Specific gravity is usually fixed at 1.010 in CKD or acute tubular necrosis as compared to prerenal acute kidney injury and inappropriate ADH secretion where specific gravity is very high – close to 1.025.

Urinary pH

Measurement of urinary pH is unnecessary except in the investigation and treatment of renal tubular acidosis (see p. 676).

Chemical (Stix) testing

Routine Stix testing of urine for blood, protein and sugar is obligatory in all patients suspected of having renal disease.

Blood

Haematuria may be overt, with bloody urine, or microscopic and found only on chemical testing. A positive Stix test must always be followed by microscopy of fresh urine (with the exception of menstruating women) to confirm the presence of red cells and so exclude the relatively rare conditions of haemoglobinuria or myoglobinuria. Bleeding may come from any site within the urinary tract (Fig. 11.5):

- *Overt bleeding from the urethra* is suggested when blood is seen at the start of voiding and then the urine becomes clear.
- *Blood diffusely present* throughout the urine comes from the bladder or above.
- *Blood only at the end of micturition* suggests bleeding from the prostate or bladder base.

Urine microscopy is used to detect red-cell casts, which are diagnostic of glomerulonephritis. In the absence of red-cell casts, further investigations, such as urine cytology, renal imaging and cystoscopy, are required to define the site of bleeding. Renal biopsy may be required (see p. 582).

Protein

Proteinuria is one of the most common signs of renal disease. Detection is primarily by Stix testing. Most reagent strips can detect protein if albuminuria exceeds 300 mg/d. They react primarily with albumin and are relatively insensitive to globulin and Bence Jones proteins.

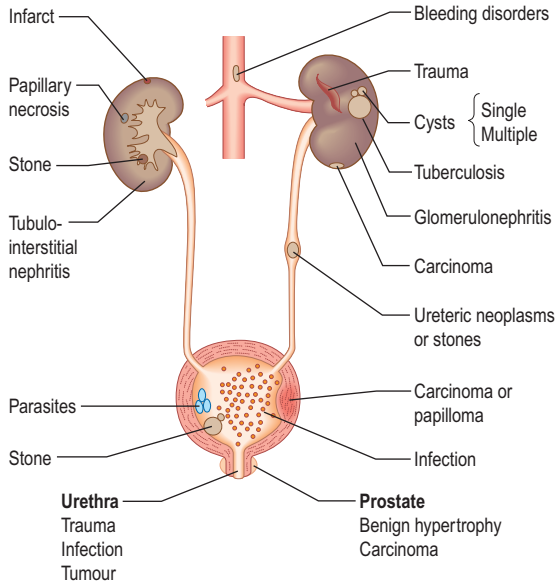


Fig. 11.5 **Sites and causes of bleeding from the urinary tract.**

If proteinuria is confirmed on repeated Stix testing, protein excretion in 24-hour urine collections should be measured (but see below). Healthy adults excrete up to 30 mg daily of albumin. Pyrexia, exercise and adoption of the upright posture (*postural proteinuria*) all increase urinary protein output but are benign.

Microalbuminuria

Normal individuals excrete less than 20 µg of albumin per minute (30 mg in 24 hours). Dipsticks, however, detect albumin only in a concentration above 200 µg (300 mg per 24 hours if urine volume is normal). An albumin excretion between these two levels – so-called 'microalbuminuria' – is now known to be an early indicator of diabetic glomerular disease and systemic endothelial dysfunction and is a useful prognostic marker for future cardiovascular events.

Timed 24-hour urinary excretion rates provide the most precise measure of microalbuminuria. However, in clinical practice it is more convenient to test for microalbuminuria using random urine samples in which albumin concentration is related to urinary creatinine concentration. Generally an albumin : creatinine ratio of 2.5 to 20 corresponds to albuminuria of 30–300 mg daily respectively. Kits are available to test for microalbuminuria.

Glucose

Renal glycosuria is uncommon, so that a positive test for glucose always requires exclusion of diabetes mellitus.

Bacteriuria

Dipstick tests for bacteriuria are based on the detection of *nitrite* produced from the reduction of urinary nitrate by bacteria and also for the detection of leucocyte esterase, an enzyme specific for neutrophils. Although each test on its own has limitations, a positive reaction with both tests has a high predictive value for urinary tract infection (p. 602).

Microscopy

Urine microscopy should be carried out in all patients suspected of having renal disease, on a 'clean' sample of mid-stream urine. The presence of numerous skin squames suggests a contaminated, poorly collected sample that cannot be properly interpreted.

If a clean sample of urine cannot be obtained, suprapubic aspiration is required in suspected urinary tract infections, particularly in children.

- *White blood cells*. The presence of 10 or more WBCs per cubic millimetre in fresh unspun mid-stream urine samples is abnormal and indicates an inflammatory reaction within the urinary tract such as urinary tract infection (UTI), stones, tubulointerstitial nephritis, papillary necrosis, tuberculosis and interstitial cystitis.
- *Red cells*. The presence of one or more red cells per cubic millimetre in unspun urine samples results in a positive Stix test for blood and is abnormal.
- *Casts* (see Fig. 11.6) are cylindrical bodies, moulded in the shape of the distal tubular lumen, and may be hyaline, granular or cellular. Coarse granular casts occur with pathological proteinuria in glomerular and tubular disease. Red-cell casts – even if only single – always indicate renal disease. White cell casts may be seen in acute pyelonephritis. They may be confused with the tubular cell casts that occur in patients with acute tubular necrosis.

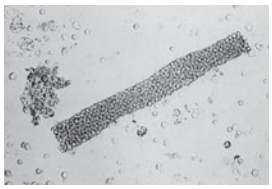


Fig. 11.6 **Red-cell cast.** Note aggregation of red cells as a 'cast' of the tubule.

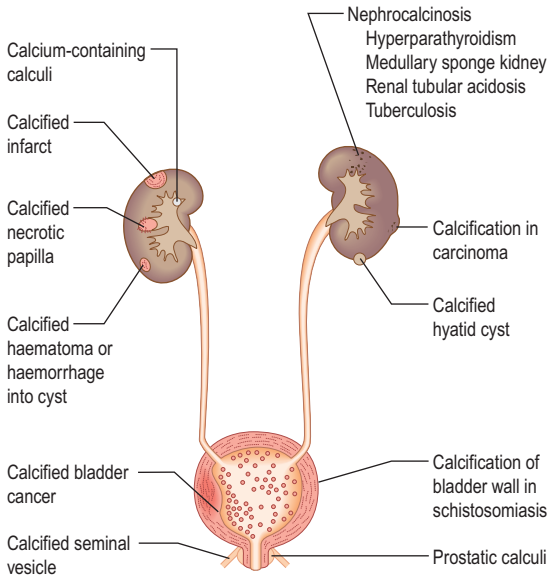


Fig. 11.7 **Calcification in the renal tract.** Calculi can occur at any site.

- *Bacteria* – see p. 602. Always culture urine prior to starting antibiotic therapy for sensitivities. Stix testing for blood or protein is of no value in the diagnosis of a UTI as both can be absent in the urine of many patients with bacteriuria.

BLOOD AND QUANTITATIVE TESTS

The use of serum urea, creatinine and GFR as measures of renal function is discussed on page 574. Other quantitative tests of disturbed renal function are described under the relevant disorders, as are diagnostic tests, e.g. ANCA, immunofluorescence, and complement.

IMAGING TECHNIQUES

Plain X-ray

A plain radiograph of the abdomen is valuable to identify renal calcification or radiodense calculi in the kidney, renal pelvis, line of the ureters or bladder (Fig. 11.7).

Ultrasonography

Ultrasonography of the kidneys and bladder has the advantage over X-ray techniques of avoiding ionizing radiation and intravascular contrast medium. In renal diagnosis it is the method of choice for:

- Renal measurement and for renal biopsy or other interventional procedures.
- Checking for pelvicalyceal dilatation as an indication of renal obstruction when chronic renal obstruction is suspected. (In suspected acute ureteric obstruction, unenhanced spiral CT is the method of choice.)
- Characterizing renal masses as cystic or solid.
- Diagnosing polycystic kidney disease.
- Detecting intrarenal and/or perinephric fluid (e.g. pus, blood).
- Demonstrating renal arterial perfusion or detecting renal vein thrombosis using Doppler. *Doppler ultrasonography* is based on the principle that, when incident sound waves are reflected from a moving structure, their frequency is shifted by an amount proportional to the velocity of the reflector (e.g. an RBC); this shift can be quantified and displayed as a spectral Doppler scan or colour overlay (colour Doppler).
- Measurement of bladder wall thickness in a distended bladder and to check for bladder tumours and stones. A scan obtained after voiding allows bladder emptying to be assessed.

The disadvantages of using ultrasonography to assess the urinary tract are:

- It does not show detailed pelvicalyceal anatomy.
- It does not fully visualize the normal adult ureter.
- It may miss small renal calculi and does not detect the majority of ureteric calculi.
- It is operator-dependent.

In patients with suspected benign prostatic hypertrophy, examination of the bladder before and after voiding, with measurement of the prostate, and examination of the kidneys to check for pelvicalyceal dilatation suffice. If prostate cancer is suspected, more detailed ultrasound examination of the prostate with a transrectal transducer, usually with transrectal prostate biopsy, is necessary.

Computed tomography (CT)

Computed tomography is used as a first-line investigation in cases of *suspected ureteric colic*. Spiral CT has both improved image resolution and allows reconstruction of the imaging data in a variety of planes. CT is also used:

- to characterize renal masses which are indeterminate at ultrasonography
- to stage renal tumours
- to detect 'lucent' calculi; low-density calculi which are lucent on plain films (e.g. uric acid stones) are well seen on CT
- to evaluate the retroperitoneum for tumours, retroperitoneal fibrosis (periaortitis) and other causes of ureteric obstruction
- to assess severe renal trauma
- to visualize the renal arteries and veins by CT angiography
- to stage bladder and prostate tumours; MRI is, however, increasingly used to stage prostate cancer
- *Disadvantages* include radiation and contrast nephrotoxicity (p. 624).

The use of spiral unenhanced CT in suspected ureteric colic permits diagnosis of causes of pain other than calculi more readily than does urography.

Magnetic resonance imaging (MRI)

MRI is used:

- to characterize renal masses as an alternative to CT
- to stage renal, prostate and bladder cancer
- to demonstrate the renal arteries by magnetic resonance *angiography* with gadolinium as contrast medium. In experienced hands its sensitivity and specificity approaches renal angiography.

Magnetic resonance urography is preferred over intravenous urography (IVU) in patients with chronic urolithiasis or intrinsic or extrinsic ureteric tumour, and in paediatric uroradiology. It is analogous to IVU. Gadolinium is used as contrast medium and is less nephrotoxic than iodine-containing agents used in IVU. However, the Federal Drug Administration (FDA) advises not using gadolinium in patients with renal insufficiency because of development of nephrogenic systemic fibrosis (p. 1254).

Excretion urography

Excretion urography (also known as IVU or intravenous pyelography (IVP)) has largely been replaced by ultrasonography and CT scanning.

Antegrade pyelography (Fig. 11.8)

Antegrade pyelography involves percutaneous puncture of a pelvicalyceal system with a needle and the injection of contrast medium to outline the pelvicalyceal system and

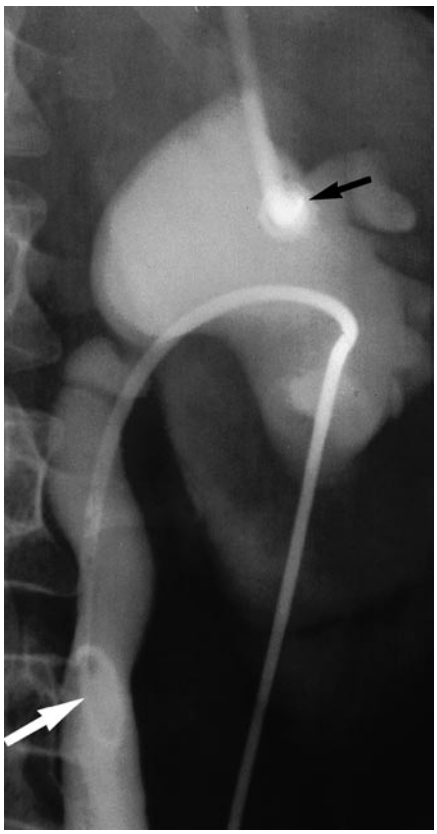


Fig. 11.8 **Antegrade pyelography via percutaneous catheter (small arrow) of obstructed system.** Percutaneous drainage catheter (large arrow) has been inserted.

ureter to the level of obstruction. It is used when ultrasonography has shown a dilated pelvicalyceal system in a patient with suspected obstruction. Antegrade pyelography is the preliminary to percutaneous placing of a drainage catheter or ureteric stent in the obstructed pelvicalyceal system (percutaneous nephrostomy).

Retrograde pyelography

Following cystoscopy, preferably under screening control, a catheter is either impacted in the ureteral orifice or passed a short distance up the ureter, and contrast medium is injected. Retrograde pyelography is mainly used to investigate lesions of the ureter and to define the lower level of ureteral obstruction shown on excretion urography or ultrasound plus antegrade studies. It is invasive, commonly requires a general anaesthetic, and may result in the introduction of infection.

Micturating cystourethrography (MCU)

This involves catheterization and the instillation of contrast medium into the bladder. The catheter is then removed and the patient screened during voiding to check for vesicoureteric reflux and to study the urethra and bladder emptying. It is used in children with recurrent infection (see p. 601).

MCU is not an appropriate investigation in adults because vesicoureteric reflux tends to disappear by the time adulthood is reached. The presence or absence of vesicoureteric reflux may also be investigated by scintigraphy (see below).

Aortography or renal arteriography

Conventional or digital subtraction angiography (DSA) is used. The latter allows the use of smaller doses of contrast medium which can be injected via a central venous catheter (venous DSA) or via a fine transfemoral arterial catheter (arterial DSA). Angiography is mainly used to define extrarenal or intrarenal arterial disease. Arteriography is still the 'gold standard' method of renal artery imaging but magnetic resonance angiography and spiral CT angiography are being used increasingly (Fig. 11.9). Complications include cholesterol embolizations (p. 609) and contrast-induced kidney damage (contrast nephropathy).

Venography is only used occasionally to exclude renal vein thrombosis. In most instances this can be done less invasively with Doppler ultrasonography, CT or MRI.

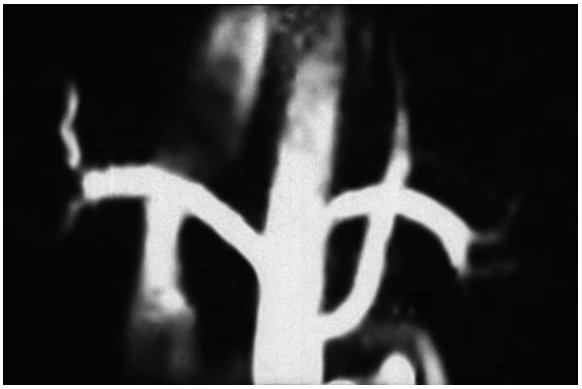


Fig. 11.9 **Magnetic resonance angiogram of normal renal arteries.**

Renal scintigraphy

Renal scintigraphy using a gamma camera is divided into:

- dynamic studies in which the function of the kidney is examined serially over a period of time, most often using a radiopharmaceutical excreted by glomerular filtration
- static studies involving imaging of tracer that is taken up and retained by the renal tubule.

Dynamic scintigraphy

The radiopharmaceutical technetium-labelled diethylenetriaminepentaacetic acid, [^{99m}Tc]DTPA, is excreted by glomerular filtration. ^{123}I-labelled ortho-iodohippuric acid (Hippuran) is both filtered and secreted by the tubules, and mercaptoacetyltriglycine (MAG3) labelled with technetium (^{99m}Tc) is excreted by renal tubular secretion. Following venous injection of a bolus of tracer, emissions from the kidney can be recorded by gamma camera. This information allows examination of blood perfusion of the kidney, uptake of tracer as a result of glomerular filtration, transit of tracer through the kidney, and the outflow of tracer-containing urine from the collecting system.

Renal blood flow. Dynamic studies can be used to investigate patients in whom renal artery stenosis is suspected as a cause for hypertension and patients with severe oliguria (post-traumatic, post-aortic surgery, or after a kidney transplant) to establish whether, and to what extent, there is renal perfusion. In patients with unilateral renal artery stenosis there is, typically, a slowed and reduced uptake of tracer with delay in reaching a peak. Studies carried out before and after administration of an ACE inhibitor may demonstrate a fall in uptake that is suggestive of functional arterial stenosis. Both false-positive and false-negative results occur, particularly in patients with renal impairment, and renal arteriography remains the 'gold standard' in the diagnosis of renal artery stenosis. In patients with total renal artery occlusion, no kidney uptake of tracers is observed.

Investigation of obstruction. Renal scintigraphy provides functional evidence of obstruction. After injection usually of (^{99m}Tc)MAG3 a rise in resistance to flow in the pelvis or ureter prolongs the parenchymal transit of tracer and there is usually a delay in emptying the pelvis. On whole-kidney renograms, the time–activity curve fails to fall after an initial peak, or continues to rise (Fig. 11.10).

When the possibility of obstruction is suspected, a dynamic renal scintigram is performed with diuresis. Furosemide (0.5 mg/kg, adult dose 40 mg) is given intravenously about 18–20 minutes into the study. Time–activity curves show an immediate fall after furosemide in the absence of obstruction but the retention of activity in the pelvis persists in the presence of obstruction. A decision as to whether conservative surgery or nephrectomy should be carried out in unilateral obstruction is facilitated by renographic assessment of the contribution of each kidney.

At the end of dynamic studies, bladder emptying may be investigated and any postmicturition residual urine measured.

Glomerular filtration rate. This is discussed on page 574.

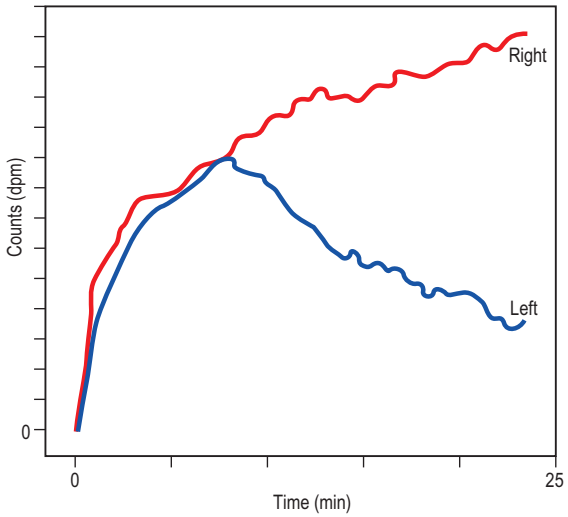


Fig. 11.10 **Dynamic scintigram.** Note the progressive rise of the right kidney curve to a plateau (in contrast to the normal left kidney curve) owing to urinary tract obstruction on the right side.

Static scintigraphy

This is usually performed using [^{99m}Tc]DMSA (dimercaptosuccinic acid), which is taken up by tubular cells. Uptake is proportional to renal function.

Relative renal function. Function is normally evenly divided between the kidneys, with a range of 45–55%. Static studies are particularly useful in unilateral renal disease, where the relative uptake of the two kidneys can be calculated.

Kidney visualization. Normal kidneys show a uniform uptake with a smooth renal outline. Scars can be identified as photon-deficient 'bites'. Static scintigraphy is of considerable value in identifying ectopic kidneys or 'pseudotumours' of the kidneys (i.e. normally functioning renal tissue abnormally placed within the kidney).

Localization of infection. The use of citrate labelled with gallium-67 or isotopically labelled leucocytes that are taken up by inflammatory tissue may be of value in defining localized infection, such as renal abscesses or infection within a renal cyst.

TRANSCUTANEOUS RENAL BIOPSY

(Practical box 11.1)

Renal biopsy is carried out under ultrasound control in specialized centres and requires interpretation by an experienced pathologist. Renal biopsy is helpful in the investigation of the nephritic and nephrotic syndromes, acute and chronic renal failure, haematuria after urological investigations and renal graft dysfunction. Native renal biopsy material must be examined by conventional histochemical staining, by electron microscopy, and by immunoperoxidase or immunofluorescence. Techniques like in situ hybridization and polymerase chain reaction analysis are also widely used in renal biopsy specimens.

The complications of transcutaneous renal biopsy are shown in Table 11.2.

Practical Box 11.1 Transcutaneous renal biopsy

Before biopsy

1. A coagulation screen is performed. It must be normal.
2. The serum is grouped and saved for crossmatching.
3. The patient is given a full explanation of what is involved and consent obtained.

During biopsy

1. The patient lies prone with a hard pillow under the abdomen.
2. The kidney is localized by ultrasound.
3. Local anaesthetic is injected along the biopsy track.
4. The patient holds a breath when the biopsy is performed.

After biopsy

1. A pressure dressing is applied to the biopsy site and the patient rests in bed for 24 hours.
2. The fluid intake is maximized to prevent clot colic.
3. The pulse and blood pressure are checked regularly.
4. The patient is advised to avoid heavy lifting or gardening for 2 weeks.

Table 11.2 Complications of transcutaneous renal biopsy

Macroscopic haematuria – about 20%
Pain in the flank, sometimes referred to shoulder tip
Perirenal haematoma
Arteriovenous aneurysm formation – about 20%, almost always of no clinical significance
Profuse haematuria demanding blood transfusion – 1–3%
Profuse haematuria demanding occlusion of bleeding vessel at angiography or nephrectomy – approximately 1 in 400
Introduction of infection
The mortality rate is about 0.1%

GLOMERULAR DISEASES

A glomerulus consists of a collection of capillaries which come from the afferent arteriole and are confined within the urinary space (Bowman's capsule); this is continuous with the proximal tubule. The capillaries are partially attached to mesangium, a continuation of the arteriolar wall consisting of mesangial cells and the matrix. The free wall of glomerular capillaries (across which filtration takes place) consists of basement membrane covered by visceral epithelial cells with individual foot processes and lined by endothelial cells (Fig. 11.11). The normal thickness of the basement membrane equals about 250–300 nm. The spaces between foot processes, with diameters of 20–60 nm, are called filtration pores, by which filtered fluid reaches the urinary space. The endothelial cells on the luminal aspect of the basement membrane are fenestrated (diameter 70–100 nm). The basement membrane is arranged in three zones – lamina rara externa, lamina rara densa and lamina rara interna – and is composed of type IV collagen and negatively charged proteoglycans (heparan sulphate).

Filtration barrier (slit diaphragm)

(Figure 11.12)

The glomerular filtration barrier consists of the fenestrated endothelium, the glomerular basement membrane and the

terminally differentiated visceral epithelial cells known as podocytes. Podocytes dictate the size-selective nature of the filtration barrier. Foot processes extend from the cell body to adhere to the glomerular basement membrane and establish a zipper-like molecular diaphragm (slit diaphragm) with a foot process from a neigbouring cell. Discovery of nephrin as the key protein of the slit diaphragm led to the identification of other proteins, such as CD2AP (CD2-associated protein), canonical TRPC6 (transient receptor potential channel 6), podocin, P-cadherin, α- and β-catenin, ZO-1 (zonula occludens-1) which co-localize within the subcellular domain to function as a molecular sieve. These proteins, in addition to providing structural support to the cytoskeletal proteins like filamentous actin, also have signalling functions in order to maintain the normal function of podocytes. Abnormalities in any of these proteins result in the breakdown of the filtration barrier with consequent torrential leak of macromolecules.

Glomerular disease includes glomerulonephritis, i.e. inflammation of the glomeruli, and glomerulopathies when there is no evidence of inflammation. There is an overlap between these terms.

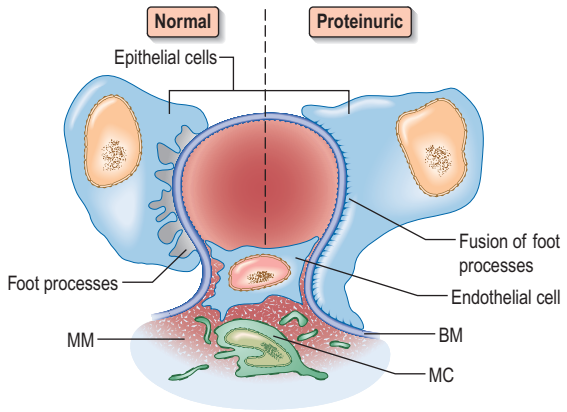
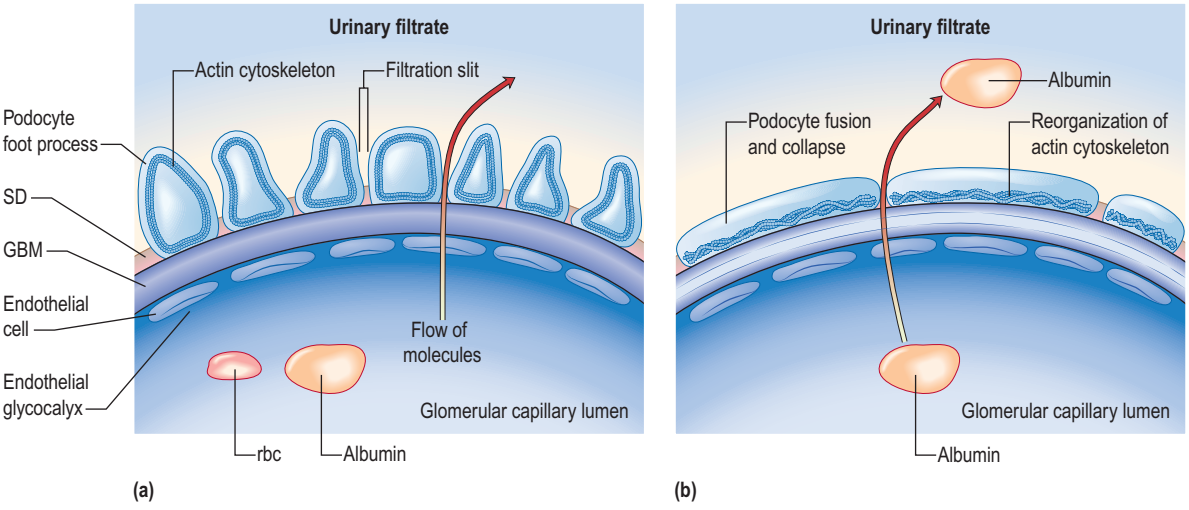


Fig. 11.11 **Diagram showing a normal (left) and proteinuric (right) glomerulus.** A capillary loop showing normal glomerular morphology on the left with an epithelial cell with pseudopodia (foot processes). In the proteinuric diagram (right) there is fusion of the foot processes characteristic of many diseases. BM, basement membrane; MC, mesangial cell; MM, mesangial matrix. After Marsh FP *Postgraduate Nephrology*. Butterworth Heinemann 1985.

GLOMERULOPATHIES

Glomerulopathies are the third most common cause of end-stage renal disease (after diabetes and hypertension) in Europe and the USA, accounting for some 10–15% of such patients.

Glomerulopathy (GN) is a general term for a group of disorders in which:



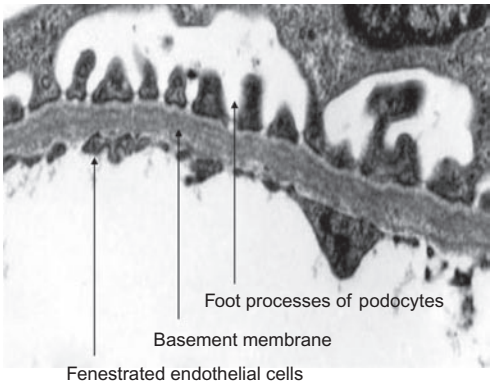


Fig. 11.12 **The glomerular filtration barrier.**
(a) Blood enters the glomerular capillaries and is filtered across the endothelium and the glomerular basement membrane and through the filtration slits between podocyte foot processes to produce the primary urine filtrate. In healthy glomeruli, this barrier restricts the passage of macromolecules. The proteins which form the slit diaphragm are essential for the normal functioning of the filtration barrier.
(b) Loss of these proteins, either genetically or acquired, leads to foot process effacement and breakdown of the barrier and leakage of albumin. **(a, b)** Adapted from Quaggin S. Sizing up sialic acid in glomerular disease. *Journal of Clinical Investigation* **117**: 1480–1483.
(c) Electron micrograph of normal filtration barrier. SD, slit diaphragm.

- There is primarily an immunologically mediated injury to glomeruli, although renal interstitial damage is a regular accompaniment.
- The kidneys are involved symmetrically.
- Secondary mechanisms of glomerular injury come into play following an initial immune insult such as fibrin deposition, platelet aggregation, neutrophil infiltration and free radical-induced damage.
- Renal lesions may be part of a generalized disease (e.g. systemic lupus erythematosus, SLE).

Pathogenesis

GN is considered to be an immunologically mediated disorder with involvement of cellular immunity (T lymphocytes, macrophages/dendritic cells), humoral immunity (antibodies, immune complexes, complement), and other inflammatory mediators (including cytokines, chemokines and the coagulation cascade). The immune response can be directed against known target antigens, particularly when GN complicates infections, neoplasia or drugs. More frequently the underlying antigenic target is unknown. Primary GN may occur in genetically susceptible individuals following an environmental insult. The genetic susceptibility is usually determined by major histocompatibility complex (HLA) genes (e.g. HLA-A1, B8, DR2, DR3). The environmental factors may be drugs (e.g. hydralazine), chemicals (e.g. gold, silica, hydrocarbons) or infectious agents. The known predisposing factors are discussed in more detail below. The physical evidence of immune reactions is indicated by the presence of circulating autoantibodies and/or abnormalities in serum complement and glomerular deposition of antibodies, immune complexes, complement and fibrin.

Pathological terms in glomerular disease

The most commonly used terms are:

- *Focal*: some but not all the glomeruli contain the lesion.
- *Diffuse* (global): most of the glomeruli (>75%) contain the lesion.
- *Segmental*: only a part of the glomerulus is affected (most focal lesions are also segmental, e.g. focal segmental glomerulosclerosis).
- *Proliferative*: an increase in cell numbers due to hyperplasia of one or more of the resident glomerular cells with or without inflammation.
- *Membrane alterations*: capillary wall thickening due to deposition of immune deposits or alterations in basement membrane.
- *Crescent formation*: epithelial cell proliferation with mononuclear cell infiltration in Bowman's space.

Classification of glomerulopathies

There is no complete correlation between the histopathological types of GN and the clinical features of disease. Glomerular diseases have been classified in numerous ways. Here they are organized and discussed as they relate to four major glomerular syndromes:

- *Nephrotic syndrome* – massive proteinuria (>3.5 g/day), hypoalbuminaemia, oedema, lipiduria and hyperlipidaemia.
- *Acute glomerulonephritis (acute nephritic syndrome)* – abrupt onset of glomerular haematuria (RBC casts or dysmorphic RBC), non-nephrotic range proteinuria, oedema, hypertension and transient renal impairment.
- *Rapidly progressive glomerulonephritis* – features of acute nephritis, focal necrosis with or without crescents and rapidly progressive renal failure over weeks.
- *Asymptomatic haematuria, proteinuria* or both.

Table 11.3 Investigation of glomerular diseases

Investigations	Positive findings
Urine microscopy	Red cells, red-cell casts
Urinary protein	Nephrotic or sub-nephrotic range proteinuria
Serum urea	May be elevated
Serum creatinine	May be elevated
Culture (throat swab, discharge from ear, swab from inflamed skin)	Nephritogenic organism (not always)
Antistreptolysin-O titre	Elevated in post-streptococcal nephritis
C3 and C4 levels	May be reduced
Antinuclear antibody	Present in significant titre in systemic lupus erythematosus
ANCA	Positive in vasculitis
Anti-GBM	Positive in Goodpasture's syndrome
Cryoglobulins	Increased in cryoglobulinaemia
Creatinine clearance	Normal or reduced
Chest X-ray	Cardiomegaly, pulmonary oedema (not always)
Renal imaging	Usually normal
Renal biopsy	Any glomerulopathy

Certain types of GN, particularly those that are a part of a systemic disease, can present as more than one syndrome, e.g. lupus nephritis, cryoglobulinaemia, and Henoch–Schönlein purpura, but more typically they are associated with the nephrotic syndrome and will be discussed below. Investigation of glomerular diseases is shown in Table 11.3.

NEPHROTIC SYNDROME

Pathophysiology

Hypoalbuminaemia. Urinary protein loss of the order 3.5 g daily or more in an adult is required to cause hypoalbuminaemia. In children, proportionately less proteinuria results in hypoalbuminaemia. The normal dietary protein intake in the UK is of the order 70 g daily and the normal liver can synthesize albumin at a rate of 10–12 g daily. How then does a urinary protein loss of the order of 3.5 g daily result in hypoalbuminaemia? This can be partly explained by increased catabolism of reabsorbed albumin in the proximal tubules during the nephrotic syndrome even though actual albumin synthesis rate is increased. However, in addition, dietary intake of protein increases albuminuria, so that the plasma albumin concentration tends to decrease during consumption of a high-protein diet. If the increase in urinary albumin excretion that follows dietary augmentation is prevented by

administration of ACE inhibitors (ACEI), a high-protein diet causes an increase in plasma albumin concentration in the nephrotic syndrome. Therefore, to maximize serum albumin concentration in nephrotic patients, a reduction in urinary albumin excretion with an ACEI is always necessary.

Proteinuria. The mechanism of the proteinuria is complex. It occurs partly because structural damage to the glomerular basement membrane leads to an increase in the size and number of pores, allowing passage of more and larger molecules. Electrical charge is also involved in glomerular permeability. Fixed negatively charged components are present in the glomerular capillary wall, which repel negatively charged protein molecules. Reduction of this fixed charge occurs in glomerular disease and appears to be a key factor in the genesis of heavy proteinuria.

Hyperlipidaemia. The characteristic disorder is an increase in the low-density lipoprotein (LDL), very-low-density lipoprotein (VLDL), and/or intermediate-density lipoprotein (IDL) fractions, but no change or decrease in HDL. This results in an increase in the LDL/HDL cholesterol ratio. Hyperlipidaemia is the consequence of increased synthesis of lipoproteins (such as apolipoprotein B, C-III lipoprotein (a)), as a direct consequence of a low plasma albumin. There is also a reduced clearance of the principal triglycerides bearing lipoprotein (chylomicrons and VLDL) in direct response to albuminuria.

Oedema in hypoalbuminaemia. See Chapter 12 (p. 655).

Management

General measures

- *Initial treatment* should be with dietary sodium restriction and a thiazide diuretic (e.g. bendroflumethiazide 5 mg daily). Unresponsive patients require furosemide 40–120 mg daily with the addition of amiloride (5 mg daily), with the serum potassium concentration monitored regularly. Nephrotic patients may malabsorb diuretics (as well as other drugs) owing to gut mucosal oedema, and parenteral administration is then required initially. Patients are sometimes hypovolaemic, and moderate oedema may have to be accepted in order to avoid postural hypotension.
- *Normal protein intake* is advisable. A high-protein diet (80–90 g protein daily) increases proteinuria and can be harmful in the long term. Infusion of albumin produces only a transient effect. It is only given to diuretic-resistant patients and those with oliguria and uraemia in the absence of severe glomerular damage, e.g. in minimal-change nephropathy. Albumin infusion is combined with diuretic therapy and diuresis often continues with diuretic treatment alone.
- *Hypercoagulable states* predispose to venous thrombosis. The hypercoagulable state is due to loss of clotting factors (e.g. antithrombin) in the urine and an increase in hepatic production of fibrinogen. Prolonged bed rest should therefore be avoided as thromboembolism is very common in the nephrotic syndrome. In the absence of any contraindication, long-term prophylactic anticoagulation is desirable. If renal vein thrombosis occurs, permanent anticoagulation is required.
- *Sepsis* is a major cause of death in nephrotic patients. The increased susceptibility to infection is partly due to loss of immunoglobulin in the urine. Pneumococcal infections are particularly common and pneumococcal vaccine should be given. Early detection and aggressive treatment of infections, rather than long-term antibiotic prophylaxis, is the best approach.
- *Lipid abnormalities* are responsible for an increase in the risk of cardiovascular disease in patients with proteinuria. Treatment of hypercholesterolaemia starts with an HMG-CoA reductase inhibitor.
- *ACE inhibitors and/or angiotensin II receptor antagonists* (AIIRA) are used for their antiproteinuric properties in all types of GN. These groups of drugs reduce proteinuria by lowering glomerular capillary filtration pressure; the blood pressure and renal function should be monitored regularly.

Specific measures

The aim is to reverse the abnormal urinary protein leak. These measures are discussed in detail below.

Table 11.4 shows the glomerular lesions commonly associated with the nephrotic syndrome. These are divided into diseases with or without RBC casts (active or bland urine sediments). Each of these entities occurs as a primary renal lesion or as a secondary component of a systemic disease.

Nephrotic syndrome with 'bland' urine sediments

Minimal-change glomerular lesion (minimal-change nephropathy)

In this condition the glomeruli appear normal on light microscopy (Fig. 11.13). The only abnormality seen on electron

Table 11.4 Glomerulopathies associated with the nephrotic syndrome

Nephrotic syndrome with 'bland' urine sediments
Primary glomerular disease
Minimal-change glomerular lesion
Congenital nephrotic syndrome
Focal segmental glomerular sclerosis
Membranous nephropathy
Secondary glomerular disease
Amyloidosis
Diabetic nephropathy
Nephrotic syndrome with 'active' urine sediments (mixed nephrotic/nephritic)
Primary glomerular disease
Mesangiocapillary glomerulonephritis
Mesangial proliferative glomerulonephritis
Secondary glomerular disease
Systemic lupus erythematosus
Cryoglobulinaemic disease
Henoch–Schönlein syndrome
Idiopathic fibrillary glomerulopathy
Immunotactoid glomerulopathy
Fibronectin glomerulonephropathy

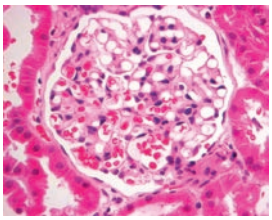


Fig. 11.13 **Normal glomerulus on light microscopy in minimal change disease.**

microscopy is fusion of the foot processes of epithelial cells (podocytes) (Fig. 11.12b). This is a non-specific finding and is seen in many conditions associated with proteinuria. Neither immune complexes nor anti-GBM antibody can be demonstrated by immunofluorescence. However, the immunological pathogenesis of this condition is suggested by three factors:

- its response to steroids and immunosuppressive drugs
- its occurrence in Hodgkin's lymphoma, with remission following successful treatment
- patients with the condition and their family members have a high incidence of asthma and eczema; remission of the nephrotic syndrome following desensitization or antigen avoidance has been described.

A suggested explanation for the proteinuria is that immature differentiating CD34 stem cells rather than mature T lymphocytes are responsible for the pathogenesis of minimal change nephropathy.

Clinical features

Minimal-change nephropathy is most common in children, particularly males, accounting for the large majority of cases of nephrotic syndrome (proteinuria is usually highly selective) in childhood. Oedema is present and in children this may be facial. The condition accounts for 20–25% of cases of adult nephrotic syndrome. It is often regarded as a condition that does not lead to chronic kidney disease (but see 'Focal segmental glomerulosclerosis' below).

Management

High-dose corticosteroid therapy with prednisolone 60 mg/m^2 daily (up to a maximum of 80 mg/day) for a maximum of 4–6 weeks followed by 40 mg/m^2 every other day for a further 4–6 weeks corrects the urinary protein leak in more than 95% of children. Response rates in adults are significantly lower and response may occur only after many months (12 weeks with daily steroid therapy and 12 weeks of maintenance with alternate-day therapy). Spontaneous remission also occurs and steroid therapy should, in general, be withheld if urinary protein loss is insufficient to cause hypoalbuminaemia or oedema.

In children, two-thirds subsequently relapse and further courses of corticosteroids are required. One-third of these children regularly relapse on steroid withdrawal, so that cyclophosphamide should be added after repeat induction with steroids. A course of cyclophosphamide 1.5–2.0 mg/kg daily is given for 8–12 weeks with concomitant prednisolone 7.5–15 mg/day. This increases the likelihood of long-term remission. Steroid unresponsive patients may also respond to cyclophosphamide. No more than two courses of cyclophosphamide should be prescribed in children because of the risk of side-effects, which include azoospermia.

In both children and adults, if remission lasts for 4 years after steroid therapy, further relapse is very rare.

An alternative to cyclophosphamide is ciclosporin 3–5 mg/kg/day, which is effective but must be continued long term to prevent relapse on stopping treatment. Excretory function and ciclosporin blood levels (recommended trough levels 80–150 ng/mL) must be monitored regularly, as ciclosporin is potentially nephrotoxic. In corticosteroid-dependent children, the anthelminthic agent levamisole 2.5 mg/kg to a maximum of 150 mg on alternate days is useful in maintenance of remission but its mode of action is unexplained. Most of the controlled studies have been conducted in the paediatric age group, making recommendations on treatment in adults difficult.

Congenital nephrotic syndrome

Congenital nephrotic syndrome (Finnish type) is an autosomal recessively inherited disorder due to mutations in the gene coding for a transmembrane protein, nephrin, that occurs with a frequency of 1 per 8200 live births in Finland. Its loss of function results in massive proteinuria shortly after birth; these patients usually have an enlarged placenta. This disorder can be diagnosed *in utero*; increased α-fetoprotein in amniotic fluid is a common feature. The microscopic features of the kidney are varied. Some glomeruli are small and infantile, whereas others are enlarged, more mature and have diffuse mesangial hypercellularity. Because of the massive proteinuria, some tubules develop microcysts and are dilated. On electron microscopy, complete effacement of the foot processes of visceral epithelial cells is observed. This condition is characterized by relentless progression to end-stage renal failure.

Other inherited nephrotic syndromes involve mutations in other genes that encode podocyte proteins such as podocin, α-actinin-4 and Wilms' tumour suppressor gene.

Focal segmental glomerulosclerosis (FSGS)

Clinical features

This disease of unknown aetiology usually presents as massive proteinuria (usually non-selective), haematuria, hypertension and renal impairment. Patients with nephrotic syndrome are often resistant to steroid therapy. All age groups are affected. It usually recurs in transplanted kidneys, sometimes within days of transplantation, particularly in patients with aggressive native renal disease.

Aetiology

A circulating permeability factor (with serine protease activity) causes the increased protein leak; plasma from patients increases membrane permeability in isolated glomeruli. Kidneys transplanted into murine models of FSGS develop the lesion, but kidneys from FSGS-prone mice transplanted to a normal strain are protected. Removal of this factor by plasmapheresis results in transient amelioration of proteinuria. Upregulation of CD80 in podocytes has a major role in the co-stimulatory immune response pathway. Anti-CD80 antibodies, used in renal transplantation, are likely to be used in FSGS in the future.

Pathology

This glomerulopathy is defined primarily by its appearance on light microscopy. Segmental glomerulosclerosis is seen,

www.studentconsult.com
Ask the authors

which later progresses to global sclerosis. The deep glomeruli at the corticomedullary junction are affected first. These may be missed on transcutaneous biopsy, leading to a mistaken diagnosis of a minimal-change glomerular lesion. A pathogenetic link may exist between minimal-change nephropathy and focal glomerulosclerosis, as a proportion of cases classified as having the former condition develop progressive renal impairment, which is unusual. Immunofluorescence shows deposits of C3 and IgM in affected portions of the glomerulus. The other glomeruli are usually enlarged but may be of normal size. In some patients mesangial hypercellularity is a feature. Focal tubular atrophy and interstitial fibrosis are invariably present. Electron microscopic findings mirror light microscopic features with capillary obliteration by hyaline deposits (mesangial matrix and basement membrane material) and lipids. The other glomeruli exhibit primarily foot process effacement, occasionally in a patchy distribution.

Five histological variants of FSGS exist:

- In *classic FSGS* (Fig. 11.14a) the involved glomeruli show sclerotic segments in any location of the glomerulus.
- The *glomerular tip lesion* is characterized by segmental sclerosis, at the tubular pole of all the affected glomeruli at a very early stage (tip FSGS) (Fig. 11.14b). Capillaries contain foam cells, and overlying visceral epithelial cells are enlarged and adherent to the most proximal portion of proximal tubules. These patients have a more favourable response to steroids and run a more benign course.
- In *collapsing FSGS* (Fig. 11.14c) the visceral cells are usually enlarged and coarsely vacuolated with wrinkled and collapsed capillary walls. These features indicate a severe lesion, with a corresponding progressive clinical course of the disease. Collapsing FSGS is commonly seen in young blacks with human immunodeficiency virus (HIV) infection or disease and is known as HIV-associated nephropathy (HIVAN) (see p. 189).
- The *perihilar variant* (Fig 11.14d) consists of perihilar sclerosis and hyalinosis in more than 50% of segmentally sclerotic glomeruli. It is frequently observed with secondary FGS due to processes associated with increased glomerular capillary pressure and declining renal mass.
- The *cellular variant* (Fig 11.14e) is characterized by at least one glomerulus with segmental endocapillary hypercellularity that occludes the capillary lumen. Other glomeruli may exhibit findings consistent with classic FGS.

The tip and collapsing variants have to be excluded histologically to make a diagnosis of the cellular variant. Patients with this variant can have severe proteinuria.

Similar glomerular changes are seen as a secondary phenomenon when the number of functioning nephrons is reduced for any reasons (e.g. nephrectomy, hypertension, gross obesity, ischaemia, sickle nephropathy, reflux nephropathy, heroin abuse, chronic allograft nephropathy, IgA nephropathy and scarring following renal vasculitis), leading to the hypothesis that FSGS results from overloading (glomerular hyperfiltration) of the remaining nephrons.

Treatment

Prednisolone 0.5–2 mg/kg/day is used in most patients and continued for 6 months before the patient is considered resistant to therapy, which is common. The use of ciclosporin at doses to maintain serum trough levels at 150–300 ng/mL may be effective in reducing or stopping urinary protein excretion. Relapse after reducing or stopping ciclosporin is very common so that long-term use is required. Cyclophosphamide, chlorambucil or azathioprine are used for second-line therapy in adults. In FSGS patients with mesangial hypercellularity and tip lesion, cyclophosphamide 1–1.5 mg/kg/day with 60 mg of prednisolone for 3–6 months followed by prednisolone and azathioprine can be used as maintenance therapy. About 50% of patients progress to end-stage renal failure within 10 years of diagnosis, particularly those who are resistant to therapy.

Membranous glomerulopathy

Clinical features

This condition occurs mainly in adults, predominantly in males. Patients present with asymptomatic proteinuria or frank nephrotic syndrome. Microscopic haematuria, hypertension and/or renal impairment may accompany the nephrotic syndrome. As in all nephritides, hypertension and a greater degree of renal impairment are poor prognostic signs. In membranous GN almost half of the patients undergo

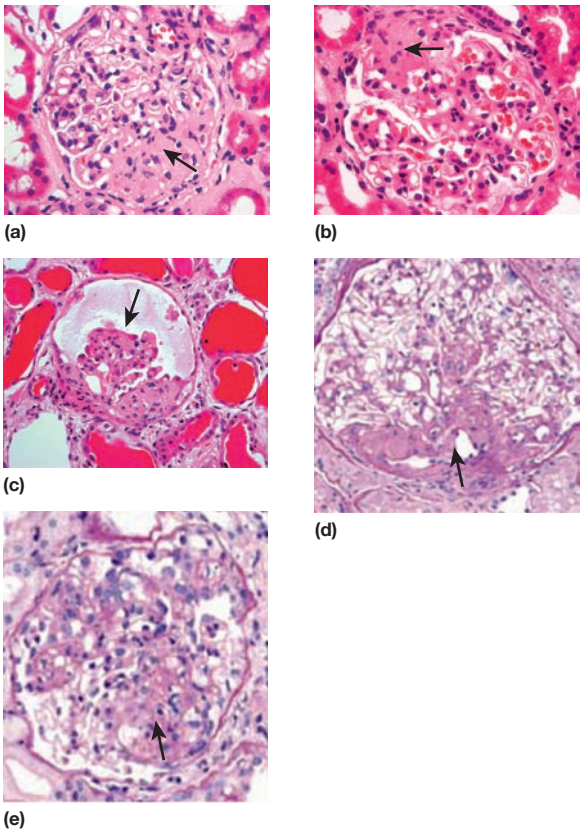


Fig. 11.14 **Focal segmental glomerulosclerosis (FSGS).**
(a) Classic FSGS showing sclerotic segments (arrow) in glomerulus.
(b) Glomerular tip lesions showing segmental sclerosis (arrow) at pole of glomerulus.
(c) Collapsing FSGS (arrow).
(d) FSGS perihilar variant.
(e) FSGS cellular variant.

spontaneous or therapy-related remission. However, eventually about 40% develop chronic kidney disease, usually in association with persistent nephrotic range proteinuria. Younger patients, females and those with asymptomatic proteinuria of modest degree at the time of presentation do best.

Aetiopathogenesis

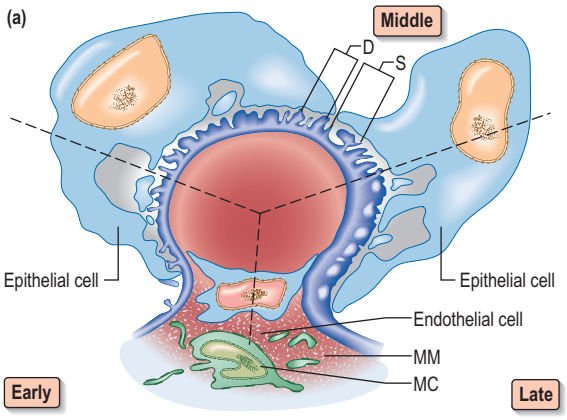
An identical glomerular histological picture is seen in the primary or idiopathic form (which comprises 75% of the cases) and also when membranous GN is secondary to drugs (e.g. penicillamine, gold, NSAIDs, probenecid, mercury, captopril), autoimmune disease (e.g. SLE, thyroiditis), infectious disease (e.g. hepatitis B, hepatitis C, schistosomiasis, *Plasmodium malariae*), neoplasia (e.g. carcinoma of lung, colon, stomach, breast and lymphoma) and other causes (e.g. sarcoidosis, kidney transplantation, sickle cell disease). At all stages, immunofluorescence shows the presence of uniform granular capillary wall deposits of IgG and complement C3. In the early stage the deposits are small and can be missed on light microscopy. Electron microscopy reveals small electron-dense deposits in the subepithelial aspects of the capillary walls. In the intermediate stage the deposits are encircled by basement membrane, which gives an appearance of spikes of basement membrane perpendicular to the basement membrane on silver staining. Late in the disease the deposits are completely surrounded by basement membrane and are undergoing resorption, which appears as uniform thickening of the capillary basement membrane on light microscopy (Fig. 11.15a, b and c).

An animal model (Heymann nephritis) in which the morphological appearances closely resemble the human condition can be induced in susceptible rats by immunization with renal autoantigens such as the brush border component of proximal tubular cells, megalin (gp330). However, the target autoantigen is unknown in humans, as megalin is absent from human podocytes. Several neutral endopeptidases located in the brush border and glomeruli are potential target autoantigens. Immune complexes are formed in situ after reacting with autoantibodies directed against these renal autoantigens.

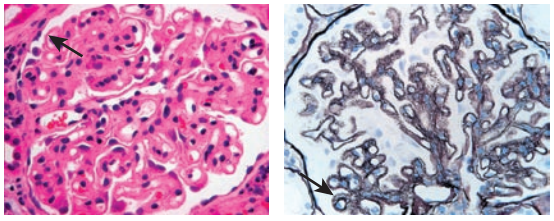
Treatment

There is no consensus on therapies for all cases but, in general, patients with moderate to heavy proteinurea are treated:

- Oral high-dose corticosteroids and azathioprine are not associated with any significant benefits.
- The alkylating agents, cyclophosphamide and chlorambucil, are both effective in the management of membranous GN but are reserved for patients with severe or prolonged nephrosis (i.e. proteinuria >6 g/day for >6 months), renal insufficiency and hypertension. There is a high likelihood of progression to end-stage renal failure.
- Chlorambucil (0.2 mg/kg/day in months 2, 4 and 6 alternating with oral prednisolone 0.4 mg/kg/day in months 1, 3 and 5) or cyclophosphamide (1.5–2.5 mg/kg/day for 6–12 months with 1 mg/kg/day of oral prednisolone on alternate days for the first 2 months) are equally effective.
- Ciclosporin and mycophenolate with oral steroids may become the agents of choice.



(a) Schematic representation of membranous nephropathy showing the early, middle and late stages. D, subepithelial deposits; MC, mesangial cell; MM, mesangial matrix; S, spikes. (After Marsh FP *Postgraduate Nephrology*. Oxford: Butterworth Heinemann 1985.)



(b) (c)

(b) Light microscopy of membranous nephropathy showing thickened basement membrane (arrow).
(c) Silver stain showing spikes (arrow) in membranous nephropathy.

Fig. 11.15 **Membranous glomerulopathy.**

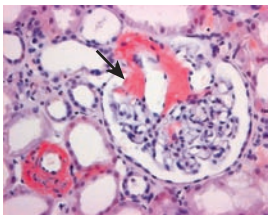
- Anti-CD20 antibodies (rituximab, which ablates B lymphocytes) have been shown to improve renal function, reduce proteinuria and increase the serum albumin; no significant adverse affects have been shown in the short term.

Amyloidosis (see p. 1072)

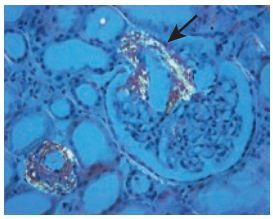
Amyloidosis is an acquired or inherited disorder of protein folding, in which normally soluble proteins or fragments are deposited extracellularly as abnormal insoluble fibrils causing progressive organ dysfunction and death.

Pathology

On light microscopy, eosinophilic deposits are seen in the mesangium, capillary loops and arteriolar walls. Staining with Congo red renders these deposits pink and they show green birefringence under polarized light (Fig. 11.16). Immunofluorescence is unhelpful, but on electron microscopy the characteristic fibrils of amyloid can be seen. Amyloid consisting of immunoglobulin light chains (AL amyloid) can be identified by immunohistochemistry in only 40% of the cases as compared to almost 100% of patients with protein found in secondary amyloid (AA amyloid). Amyloid A (AA) amyloidosis, also referred to as secondary amyloidosis, is a rare but serious complication of chronic inflammatory diseases and chronic infections.



(a)



(b)

Fig. 11.16 **Amyloid.**
(a) Light microscopy of eosinophilic amyloid deposits (arrow). **(b)** Congo red stain under polarized light showing apple green refringence.

Diagnosis and treatment (see p. 1073)

The *diagnosis* can often be made clinically when features of amyloidosis are present elsewhere. On imaging, the kidneys are often large. Scintigraphy with radiolabelled serum amyloid P (SAP), a technique for quantitatively imaging amyloid deposits in vivo, is used to detect the rate of regression or progression of amyloidosis over a period of time (p. 1073). Renal biopsy is necessary in all suspected cases of renal involvement.

Treatment. Treatments that reduce production of the amyloidogenic protein can improve organ function and survival in immunoglobulin-light-chain-related (AL) amyloidosis and hereditary transthyretin-associated (ATTR) amyloidosis (see p. 1073).

In AA amyloidosis, production of serum amyloid A can sometimes be decreased by treatment of the underlying inflammatory condition but cannot be completely suppressed. A new class of drug, eprodisate (see p. 1073), has shown modest success in patients with renal amyloidosis.

Renoprotective measures should be started (p. 632). The success of dialysis and kidney transplantation is dependent upon the extent of amyloid deposition in extrarenal sites, especially the heart.

Diabetic nephropathy

Diabetic renal disease is the leading cause of end-stage renal failure in the western world. Type 1 and type 2 diabetic patients (see p. 1053) have equivalent rates of proteinuria, azotaemia, and ultimately end-stage renal failure. Both types of diabetes show strong similarities in their rate of renal functional deterioration, and onset of co-morbid complications.

Pathology

The kidneys enlarge initially and there is glomerular hyperfiltration (GFR >150 mL/min). The major early histological lesions seen are glomerular basement membrane thickening and mesangial expansion. Moreover, progressive depletion of podocytes from the filtration barrier due to either apoptosis or detachment and resulting podocyturia appears to be a very early ultrastructural change. Later, glomerulosclerosis develops with nodules (Kimmelstiel–Wilson lesion) and hyaline deposits in the glomerular arterioles (see Fig. 11.17). It has recently been shown that the mesangial expansion and the hyalinosis are partly due to amylin (beta islet specific amyloid protein) deposits. These later changes are associated with heavy proteinuria. The lesions seen in type 1 are also seen in type 2. The pathophysiology is discussed on page 1053.

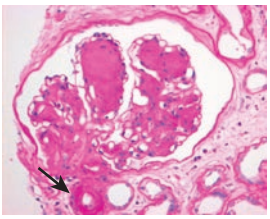


Fig. 11.17 **Diabetes mellitus**. Advanced diabetic glomerulopathy and arteriolar sclerosis (arrow).

Treatment

Lifestyle changes (cessation of smoking and increase in exercise), hypertension, poor metabolic regulation and hyperlidaemia should be addressed in every diabetic. Microalbuminuria is a reason to start treatment with ACE inhibitors or an angiotensin II receptor antagonist (AIIRA) in either type of diabetes, regardless of blood pressure elevation. Like other kidney diseases, however, nearly the entire course of renal injury in diabetes is clinically silent. The timing of medical intervention during this silent phase (see Box 11.6, p. 632) is renoprotective, as judged by slowed loss of glomerular filtration. Despite intensified metabolic control and antihypertension treatment in diabetic patients, a substantial number still go on to develop end-stage renal failure.

Nephrotic syndrome with 'active' urine sediments (mixed nephrotic/nephritic)

Mesangiocapillary (membranoproliferative) glomerulonephritis (MCGN)

This uncommon lesion has three subtypes with similar clinical presentations: the nephrotic syndrome, haematuria, hypertension and renal impairment. They also have similar microscopic findings although the pathogenesis may be different. Electron microscopy defines:

- *Type 1 MCGN*. There is mesangial cell proliferation, with mainly subendothelial immune deposition and apparent splitting of the capillary basement membrane, giving a 'tram-line' effect. It can be associated with persistently reduced plasma levels of C3 and normal levels of C4 due to activation of the complement cascade by the classical pathway. It is often idiopathic but occurs with chronic infection (abscesses, infective endocarditis, infected ventriculoperitoneal shunt) or cryoglobulinaemia secondary to hepatitis C infections (Fig. 11.18a).
- *Type 2 MCGN*. There is mesangial cell proliferation with electron-dense, linear intramembranous deposits that usually stain for C3 only (Fig. 11.18b). This type may be idiopathic or be associated with partial lipodystrophy (loss of subcutaneous fat on face and upper trunk). MCGN affects young adults. These patients have low C3 levels as in type 1 but this is due to the activation of the alternative pathway of the complement cascade; they also have autoantibodies to C3 convertase enzyme.
- *Type 3 MCGN* has features of both types 1 and 2. Complement activation appears to be via the final common pathway of the cascade.

Most patients eventually go on to develop renal failure over several years. Type 2 MCGN recurs in virtually 100% of renal transplant patients but recurrence is less common in type 1

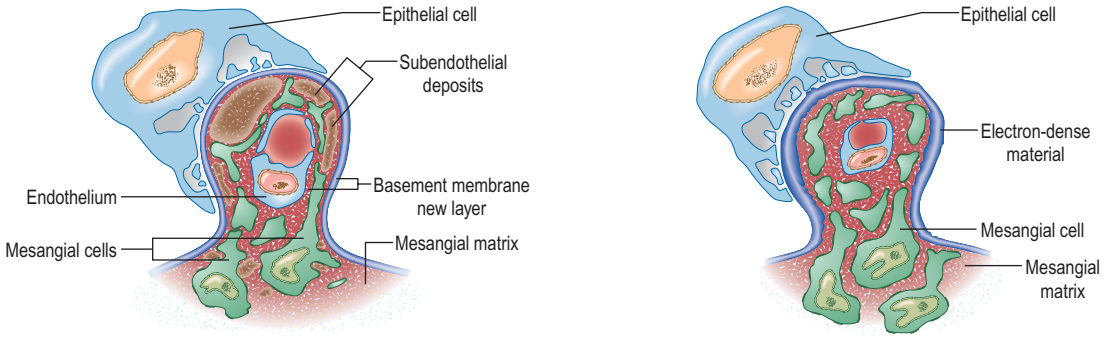


Fig. 11.18 **Mesangiocapillary glomerulonephritis (MCGN). (a)** Type 1 MCGN showing expanded mesangial matrix and mesangial cells, thickened capillary wall, large subendothelial deposits and formation of a new layer of basement membrane (tram-line effect). **(b)** Type 2 MCGN. This shows a variable glomerular appearance; very electron-dense material has replaced Bowman's capsule, tubular basement membrane and part of the capillary. There is some proliferation of mesangial cells. (After Marsh FP *Postgraduate Nephrology.* Oxford: Butterworth Heinemann 1985.)

(25%). However, recurrence does not interfere with long-term graft function.

Management

In idiopathic MCGN (*all age groups*) with normal renal function, non-nephrotic range proteinuria, no specific therapy is required. Follow-up every four months, with specific attention to blood pressure control, is required.

In *children* with the nephritic syndrome and/or impaired renal function, a trial of steroids is warranted (alternate-day prednisolone 40 mg/m^2 for a period of 6–12 months). If no benefit is seen, this treatment is discontinued. Regular follow-up with control of blood pressure, use of agents to reduce proteinuria and correction of lipid abnormalities is necessary.

In *adults* with the nephritic syndrome and/or renal impairment, aspirin (325 mg) or dipyridamole (75–100 mg) daily or a combination of the two, should be given for 6–12 months. Again, if no benefits are seen, the treatment should be stopped. Treatment to slow the rate of progression of renal failure is instituted (p. 632).

Mesangial proliferative GN (IgM nephropathy, C1q nephropathy)

In addition to minimal-change disease, there are two other disorders that usually present with heavy proteinuria with only minor changes on light microscopy.

IgM nephropathy is characterized by increased mesangial cellularity in most of the glomeruli, associated with granular immune deposits (IgM and complement) in the mesangial regions. Patients present with episodic or persistent haematuria with the nephrotic syndrome. Unlike minimal-change disease, the prognosis is not uniformly good, as steroid response is only 50% compared to 90% in minimal-change disease. Between 10% and 30% develop progressive renal insufficiency with signs of secondary FSGS (p. 587) on repeat biopsy. A trial of cyclophosphamide with prednisolone is used with persistent nephrotic syndrome, particularly with a rising plasma creatinine concentration.

C1q nephropathy is very similar to IgM nephropathy in presenting features and microscopic appearance with the exception of C1q deposits in the mesangium. Sometimes it is misdiagnosed as lupus nephritis, particularly in patients with negative serology (so-called 'seronegative lupus'). The distinguishing features are intense C1q staining and absence of tubuloreticular inclusions (attributable to high circulating alpha-interferon) on electron microscopy. Only some patients are steroid dependent. Progression to renal failure is, as in most glomerular diseases, most likely to occur in patients with heavy proteinuria and renal insufficiency.

Systemic lupus erythematosus (lupus glomerulonephritis) (see also p. 541)

Overt renal disease occurs in at least one-third of SLE patients and, of these, 25% reach end-stage renal failure within 10 years. Histologically almost all patients will have changes which were classified by the World Health Organization (1982) and have recently been reviewed by the International Society of Nephrology/Renal Pathological Society (2003). There is a progression of histological findings and clinical picture from I to VI shown in Box 11.2 with modification to include clinical details.

Box 11.2 Classification of lupus nephritis

Class I – Minimal mesangial lupus nephritis (LN), with immune deposits but normal on light microscopy. Asymptomatic.
Class II – Mesangial proliferative LN with mesangial hypercellularity and matrix expansion. Clinically, mild renal disease.
Class III – Focal LN (involving <50% of glomeruli) with subdivisions for active or chronic lesions. Subepithelial deposits seen. Clinically have haematuria and proteinuria. 10–20% of all LN.
Class IV – Diffuse LN (involving ≥50% of glomeruli) (Fig. 11.19) classified by the presence of segmental and global lesions as well as active and chronic lesions. Subendothelial deposits are present. Clinically there is progression to the nephrotic syndrome, hypertension and renal insufficiency. Most common and most severe form of LN.
Class V – Membranous LN affects 10–20% of patients. Can occur in combination with III or IV. Good prognosis.
Class VI – Advanced sclerosing LN (≥90% globally sclerosed glomeruli without residual activity). This represents the advanced stages of the above, as well as healing. Immunosuppressive therapy is unlikely to help as it is 'inactive'. Progressive renal failure.

Modified from International Society of Nephrology/Renal Pathological Society (2004)

Serial renal biopsies show that in approximately 25% of patients, histological appearances alter from one class to another during the interbiopsy interval. The prognosis is better in patients with types I, II and V. Immune deposits in the glomeruli and mesangium are characteristic of SLE (tubuloreticular structure in glomerular endothelial cells) and stain positive for IgG, IgM, IgA and the complement components C3, C1q and C4 on immunofluorescence.

Pathophysiology

SLE is now known to be an autoantigen-driven, T-cell-dependent and B-cell-mediated autoimmune disease (p. 542). Lupus nephritis typically has circulating autoantibodies to cellular antigens (particularly anti-dsDNA, anti-Ro) and complement activation which leads to reduced serum levels of C3, C4, and particularly C1q. C1q is the first component of the classical pathway of the complement cascade (see p. 56) and is involved in the activation of complement and clearance of self-antigens generated during apoptosis. Anti-C1q antibodies may help in distinguishing a renal from a non-renal relapse. However, not all autoantibodies are pathogenic to the kidney. These nephritogenic antibodies have specific physicochemical characteristics and correlate well with the pattern of renal injury. DNA was thought to be the inciting autoantigen, but now nucleosomes (structures comprising DNA and histone, generated during apoptosis) are the most likely autoantigen. Nucleosome-specific T cells, antinucleosome antibodies and nephritogenic immune complexes are generated. Positively charged histone components of the nucleosome bind to the negatively charged heparan sulphate (within the glomerular basement membrane) inciting an inflammatory reaction and resulting in mesangial cell proliferation, mesangial matrix expansion and inflammatory leucocytes. Other pathogenic mechanisms include infarction of glomerular segments, thrombotic microangiopathy, vasculitis and glomerular sclerosis.

The extraglomerular features of lupus nephritis include tubulointerstitial nephritis (75% of patients), renal vein thrombosis and renal artery stenosis. Thrombotic manifestations are associated with autoantibodies to phospholipids (anticardiolipin or lupus anticoagulant) (p. 544).

Management

Initial treatment depends on the clinical presentation but hypertension and oedema should be treated. A definite histopathological diagnosis is required. *Type I* requires no treatment. *Type II* usually runs a benign course but some patients are treated with steroids.

There have been a number of clinical trials with immunosuppressive agents in *types III, IV and V.* At present, steroids and cyclophosphamide or mycophenolate mofetil are usually used for induction, with azathioprine and mycophenolate mofetil for maintenance therapy.

B cell depletion with rituximab (anti-CD20) has been used in some patients with favourable results over the short term; controlled trials are awaited.

Prognosis

Treatment leading to the normalization of proteinuria, hypertension and renal dysfunction indicates a good prognosis. Glomerulosclerosis (type VI) usually predicts end-stage renal disease (p. 625).

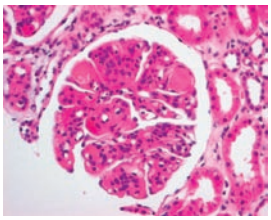


Fig. 11.19 **Lupus nephritis type IV – a diffuse proliferative nephritis.** There is proliferation of endothelial and mesangial cells.

Cryoglobulinaemic renal disease

Cryoglobulins (CG) are immunoglobulins and complement components, which precipitate reversibly in the cold. Three types are recognized:

Type I: the cryoprecipitable immunoglobulin is a single monoclonal type, as is found in multiple myeloma and lymphoproliferative disorders.

Types II and III cryoglobulinaemias are mixed types. In each, a polyclonal IgG antigen is bound to an antiglobulin. In type II, the antiglobulin component, which is usually of the IgM or IgA class with rheumatoid factor activity, is monoclonal, while in type III it is polyclonal. Type II CGs account for 40–60% cases, whilst 40–50% of all CG cases are of type III.

Glomerular disease is more common in type II than in type III cryoglobulinaemia. In approximately 30% of these 'mixed' cryoglobulinaemias, no underlying or associated disease is found (essential cryoglobulinaemia). Recognized associations include viral infections (hepatitis B and C, HIV, cytomegalovirus, Epstein–Barr infection), fungal and spirochaetal infections, malaria and infective endocarditis and autoimmune diseases (SLE, rheumatoid arthritis and Sjögren's syndrome). Glomerular pathological changes resemble MCGN (Fig. 11.18).

Presentation is usually in the fourth or fifth decades of life, and women are more frequently affected than men. Systemic features include purpura, arthralgia, leg ulcers, Raynaud's phenomenon, evidence of systemic vasculitis, a polyneuropathy and hepatic involvement. The glomerular disease presents typically as asymptomatic proteinuria, microscopic haematuria or both, but presentation with an acute nephritic and nephrotic syndrome (commonest presentation) or features of renal impairment also occurs.

A reduction in concentration of early complement components with an elevation of later components, detection of CGs, monoclonal gammopathy, rheumatoid factor, autoantibodies and antiviral antibodies or mRNA of hepatitis C, depending on the associated disorder, is seen.

Spontaneous remission occurs in about one-third of cases and approximately one-third pursue an indolent course. Corticosteroid and/or immunosuppressive therapy with cyclophosphamide may be of benefit, but evaluation of treatment is difficult owing to the rarity of the disease and the occurrence of spontaneous remissions. Intensive plasma exchange or cryofiltration has been used in selected cases. Interferon reduces the viraemia in hepatitis C but does not influence the cryoglobulinaemia. Uncontrolled studies of the anti-CD20 chimeric monoclonal antibody rituximab, which depletes B cells, appear promising, reporting improvement in general manifestations as well as glomerulonephritis.

Henoch–Schönlein syndrome (purpura)

This clinical syndrome comprises a characteristic skin rash, abdominal colic, joint pain and glomerulonephritis. Approximately 30–70% have clinical evidence of renal disease with haematuria and/or proteinuria. The renal disease is usually mild but the nephrotic syndrome and acute kidney injury can occur. The renal lesion is a focal segmental proliferative glomerulonephritis, sometimes with mesangial hypercellularity. In the more severe cases, epithelial crescents may be present. Immunoglobulin deposition is mainly IgA in the glomerular mesangium distribution, similar to IgA nephropathy. There is no treatment of proven benefit. Recently in a placebo-controlled study steroid therapy was found to be ineffective. Treatment is usually supportive but with crescentic GN aggressive immunosuppression has been tried with variable outcome.

Idiopathic fibrillary glomerulopathy

In this rare condition, characteristic microfibrillary structures are seen in the mesangium and glomerular capillary wall on electron microscopy that are clearly different from those seen in amyloidosis; the fibrils are larger than those in amyloidosis (20–30 versus 10 nm diameter) and do not stain with Congo red. The median age at presentation is approximately 45 years (range 10–80 years). Patients present with proteinuria, mostly in the nephrotic range (60%), and microscopic haematuria (70%), hypertension and renal impairment (50%) that may progress rapidly: 40–50% of patients develop end-stage renal failure within 2–6 years.

No treatment is known to be of benefit, although isolated instances of an apparent response to corticosteroid and immunosuppressive therapy have been reported.

Immunotactoid glomerulopathy

In this disorder, microtubules which are much larger (30–40 nm diameter) than the fibrils in fibrillary glomerulopathy are seen on electron microscopy. The majority of patients have circulating paraproteins, or monoclonal immunoglobulin deposition is seen in the glomeruli on immunofluorescence microscopy. A lymphoproliferative disease is the underlying cause in over 50% of cases. The clinical presentation and course are similar to fibrillary glomerulopathy, although therapy can be effective in patients with lymphoproliferative disease and/or paraproteinaemia. Complete or partial remission of the nephrotic syndrome can be achieved by various chemotherapeutic agents in over 80% of patients.

Fibronectin glomerulopathy

This is also a form of glomerulonephritis due to fibrillar deposits which, unlike amyloidosis but like fibrillary glomerulonephritis and immunotactoid glomerulopathy, is Congo red staining negative. It is inherited as an autosomal dominant disorder and is associated with the massive deposition of fibronectin, a large dimeric glycoprotein consisting of two similar subunits (approximately 250 kDa in weight). The possible genetic abnormality in this disorder is a loss of function mutation in uteroglobin.

Fibronectin glomerulopathy is extremely rare and was described only in Caucasians of European descent. An Asian family with this disease has since been reported. There are as yet no known cases in blacks or Hispanic patients. It presents with varying degrees of proteinuria seen first between the ages of 20 and 40. This is followed by hypertension, microscopic haematuria, and slow progression to end-stage renal disease in most patients. Diagnosis is confirmed by renal biopsy which demonstrates enlarged glomeruli with minimal proliferation and massive Congo red negative fibrillary deposits in the capillary walls. Treatment includes non-specific measures such as adequate blood pressure control by blockade of the renin angiotensin system. End stage renal disease (ESRD or chronic kidney disease) patients are treated with dialysis, and recurrence of disease following renal transplant has been noted.

ACUTE GLOMERULONEPHRITIS (ACUTE NEPHRITIC SYNDROME) (Table 11.5)

This comprises:

- haematuria (macroscopic or microscopic) – red-cell casts are typically seen on urine microscopy
- proteinuria
- hypertension
- oedema (periorbital, leg or sacral)
- temporarily oliguria and uraemia.

The histological pattern is characterized by cellular proliferation (mesangial and endothelial) and inflammatory cell infiltration (neutrophils, macrophages).

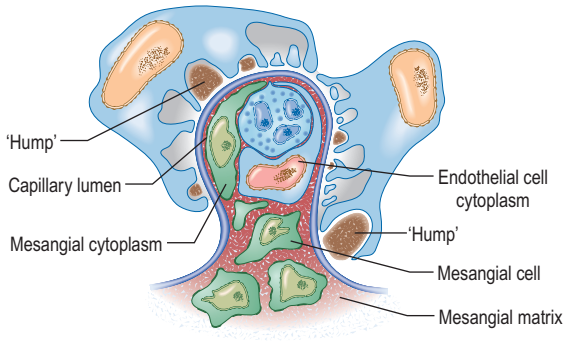
Post-streptococcal glomerulonephritis (PSGN)

The patient, usually a child, suffers a streptococcal infection 1–3 weeks before the onset of the acute nephritic syndrome. Streptococcal throat infection, otitis media or cellulitis can all be responsible. The infecting organism is a Lancefield group A β-haemolytic streptococcus of a nephritogenic type. The latent interval between the infection and development of symptoms and signs of renal involvement reflects the time taken for immune complex formation and deposition and glomerular injury to occur. PSGN is now rare in developed countries. Renal biopsy shows diffuse, florid, acute inflammation in the glomerulus (without necrosis but occasionally cellular crescents), with neutrophils and deposition of immunoglobulin (IgG) and complement (Fig. 11.20a, b). Ultrastructural findings are those of electron-dense deposits, characteristically but not solely in the subepithelial aspects of the capillary walls. Endothelial cells often are swollen. Similar biopsy findings may be seen in *non-streptococcal post-infectious glomerulonephritis* (Table 11.5).

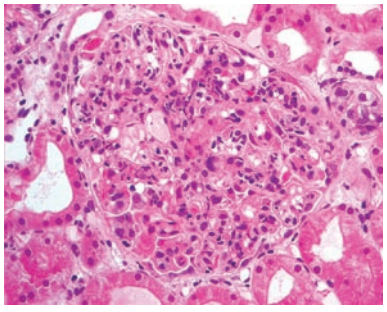
Table 11.5 Diseases commonly associated with the acute nephritic syndrome

Post-streptococcal glomerulonephritis
Non-streptococcal post-infectious glomerulonephritis, e.g. *Staphylococcus*, pneumococcus, *Legionella*, syphilis, mumps, varicella, hepatitis B and C, echovirus, Epstein–Barr virus, toxoplasmosis, malaria, schistosomiasis, trichinosis
Infective endocarditis
Shunt nephritis
Visceral abscess
Systemic lupus erythematosus (see p. 541)
Henoch–Schönlein syndrome (see p. 592)
Cryoglobulinaemia (see p. 591)

Glomerulopathies
Acute glomerulonephritis (acute nephritic syndrome)
Asymptomatic urinary abnormalities



(a)



(b)

Fig. 11.20 **Post-streptococcal glomerulonephritis.** **(a)** Diagram showing large aggregates of immune material (humps) in the extracapillary area. There is an increase in mesangial matrix and mesangial cells with occlusion of the capillary lumen by endothelial cell cytoplasm, leucocytes and mesangial cell cytoplasm. **(b)** Light microscopy showing acute inflammation of the glomerulus with neutrophils. After Marsh FP *Postgraduate Nephrology.* Oxford: Butterworth Heinemann 1985.

Management

The acute phase should be treated with antihypertensives, diuretics, salt restriction and dialysis as necessary. If recovery is slow, corticosteroids may be helpful. The prognosis is usually good in children. A small number of adults develop hypertension and/or renal impairment later in life. Therefore in older patients, an annual blood pressure check and measurement of serum creatinine are required. Evidence in support of long-term penicillin prophylaxis after the development of glomerulonephritis is lacking. In non-streptococcal post-infectious glomerulonephritis, prognosis is equally good if the underlying infection is eradicated.

Glomerulonephritis with infective endocarditis

GN occurs rarely in patients with infective endocarditis (usually i.v. drug abusers). It usually manifests itself as the acute nephritic syndrome. A similar presentation is in patients with infected ventriculoperitoneal shunt (*shunt nephritis*). Microscopic appearances resemble post-infectious GN, but lesions are usually focal and segmental. Crescentic GN with acute renal failure has been described, particularly with *Staphylococcus aureus* infection. Appropriate antibiotic therapy or surgical eradication of infection in fulminant cases (embolic infarct of the kidney) usually results in a return of normal renal function.

Glomerulonephritis associated with visceral abscesses (mainly pulmonary)

Clinical and microscopic features are indistinguishable from post-infectious GN, but usually complement levels are normal and immune deposits are absent on biopsy. Antibiotic therapy and surgical drainage of the abscess result in complete recovery of renal function in approximately 50% of patients.

ASYMPTOMATIC URINARY ABNORMALITIES

A variety of renal lesions may present as either isolated proteinuria or haematuria, alone or with proteinuria.

Isolated proteinuria without haematuria in asymptomatic patients is usually an incidental finding. It is usually in the sub-nephrotic range without an active urine sediment and there is normal renal function. Over 50% of these patients have postural proteinuria. The outcome of isolated proteinuria (postural or non-postural) is excellent in the majority of patients, with a gradual decline in proteinuria. Occasionally, it may be an early sign of a serious glomerular lesion such as membranous GN, IgA nephropathy, FSGS, diabetic nephropathy or amyloidosis. Moreover, mild proteinuria may accompany a febrile illness, congestive heart failure or infectious diseases with no clinical renal significance.

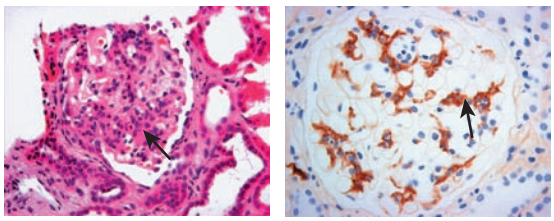
Haematuria with or without sub-nephrotic range proteinuria in an asymptomatic patient may lead to early discovery of potentially serious glomerular disease such as SLE, Henoch–Schönlein purpura, post-infectious GN or idiopathic hypercalciuria in children. Asymptomatic haematuria is also the primary presenting manifestation of a number of specific glomerular diseases discussed below.

IgA nephropathy (Fig. 11.21a, b)

This disease has replaced post-streptococcal glomerulonephritis as the commonest form of glomerulonephritis world-wide.

Pathogenesis

There is a focal and segmental proliferative glomerulonephritis with mesangial deposits of polymeric IgA_1. In some cases IgG, IgM and C3 are also seen in the glomerular mesangium. The disease may be a result of an exaggerated bone marrow and tonsillar IgA_1 immune response to viral or other antigens and is associated with an abnormality in O-linked galactosylation in the hinge region of the IgA_1 molecule. Functional abnormalities of two IgA receptors – CD89 expressed on blood myeloid cells and the transferring receptor (CD71) on mesangial cells – are seen. Abnormal IgA_1 induces the release



(a) (b)

Fig. 11.21 **IgA nephropathy. (a)** Light microscopy. Showing mesangial cell proliferation (arrow) and increased matrix. **(b)** IgA deposits on immunoperoxidase staining.

More online
www.studentconsult.com

of soluble CD89, which is responsible for the formation of circulating IgA complexes. These complexes are then trapped by CD71, which is overexpressed on mesangial cells in IgA nephropathy patients, allowing IgA complex formation in the mesangium. Up to 50% of patients exhibit elevated serum IgA (polyclonal) concentration. Superimposed crescent formation is frequent, particularly following macroscopic haematuria due to upper respiratory tract infection.

Several diseases are associated with IgA deposits, including Henoch–Schönlein purpura, chronic liver disease, malignancies (especially carcinoma of bronchus), seronegative spondyloarthritides, coeliac disease, mycosis fungoides and psoriasis.

Clinical presentation

IgA nephropathy tends to occur in children and young males. They present with asymptomatic microscopic haematuria or recurrent macroscopic haematuria sometimes following an upper respiratory or gastrointestinal viral infection. Proteinuria occurs and 5% can be nephrotic. The prognosis is usually good, especially in those with normal blood pressure, normal renal function and absence of proteinuria at presentation. Surprisingly, recurrent macroscopic haematuria is a good prognostic sign, although this may be due to 'lead-time bias' (p. 454), as patients with overt haematuria come to medical attention at an earlier stage of their illness. The risk of eventual development of end-stage renal failure is about 25% in those with proteinuria of more than 1 g per day, elevated serum creatinine, hypertension, ACE gene polymorphism (DD isoform) and tubulointerstitial fibrosis on renal biopsy.

Management

Patients with proteinuria over 1–3 g/day, mild glomerular changes only and preserved renal function should be treated with steroids. Steroids reduce proteinuria and stabilize renal function. The combination of cyclophosphamide, dipyridamole and warfarin should not be used, nor should ciclosporin. In patients with progressive disease (creatinine clearance less than 70 mL/min), fish oil or prednisolone with cyclophosphamide for 3 months followed by maintenance with prednisolone and azathioprine may be tried. A tonsillectomy can reduce proteinuria and haematuria in those patients with recurrent tonsillitis. All patients, with or without hypertension and proteinuria, should receive a combination of ACE inhibitor and angiotensin II receptor antagonist rather than each agent alone because reduction of proteinuria and preservation of renal function are better with combination therapy despite similar blood pressure control. Mesangial IgA deposits are commonly found in the allografts of transplanted patients but loss of graft function as a result is uncommon.

Alport's syndrome

Alport's syndrome is a rare condition characterized by an hereditary nephritis with haematuria, proteinuria (less than 1–2 g/day), progressive renal failure and high-frequency nerve deafness. Approximately 15% of cases may have ocular abnormalities such as bilateral anterior lenticonus and macular and perimacular retinal flecks. In about 85% of patients with Alport's syndrome there is X-linked inheritance of a mutation in the *COL4α5* gene encoding the *COL4α5* collagen chain. In female carriers, penetrance is variable and depends on the type of mutation or degree of mosaicism following hybridization of the X chromosome. Patients with autosomal recessive or dominant modes of inheritance have also been described with mutations in *COL4α3* or *COL4α4* genes. In families with stromal cell tumours there is an additional mutation in the *COL4α6* gene.

Mutations present in Alport's syndrome that produce post-translational defects in α3, α4 and α5 chains result in incorrect assembly or folding of monomers; such defective monomers are rapidly degraded. These mutations arrest the normal developmental switch and cause the persistence of embryonic α1, α1 and α2 networks in glomerular basement membrane. These networks are more susceptible to endoproteolysis and oxidative stress than the α3, α4 and α5 network. Over time, patients with Alport's syndrome probably become more sensitive to selective basement membrane proteolysis, which may explain why their glomerular membranes thicken unevenly, split and ultimately deteriorate.

The primary glomerular filtration barrier of the glomerular capillary consists of the basement membrane and the outer slit diaphragm formed between adjacent podocytes. Loss of slit function causes massive proteinuria (congenital nephrotic syndrome, p. 586) but deterioration of glomerular basement membrane produces only mild proteinuria. Proteinuria in Alport's syndrome is the result of glomerular sclerosis, rather than primary loss of slit pores. In pedigrees with a history of renal failure, disease progresses from concomitant interstitial fibrosis, macrophage and lymphocyte infiltration secondary to tubular basement disruption and transdifferentiation of epithelial mesenchymal cells to fibroblasts. This fibrogenic response destroys renal architecture. The renal histology characteristically shows split basement membrane. In some patients with Alport's syndrome and carriers, thin basement membrane, as seen in benign familial haematuria, is the only abnormality detected on histology. For this reason, the boundary between Alport's and benign familial haematuria has become increasingly vague.

Management

The disease is progressive and accounts for some 5% of cases of end-stage renal failure in childhood or adolescence. Patients with early renal failure can be treated with ACE inhibitors to attenuate proteinuria. Exciting experimental evidence suggests that mesenchymal stem cells can transdifferentiate into podocytes and repair basement abnormalities and slow the rate of progression. Anti-GBM antibody does not adhere normally to the glomerular basement membrane of affected individuals but development of crescentic glomerulonephritis in the transplanted kidney due to anti-GBM alloantibody is a well-recognized complication.

Thin glomerular basement membrane disease

This condition is inherited as an autosomal dominant and typically presents with persistent microscopic glomerular haematuria (RBC casts or dysmorphic RBCs). The diagnosis is made by renal biopsy, which shows thinning of the glomerular capillary basement membrane on electron microscopy. The condition was underdiagnosed and is much commoner than previously believed. The prognosis for renal function is usually very good but some patients develop renal insufficiency over decades. The cause of renal impairment in this condition is not known but may be due to secondary FSGS or concomitant IgA nephropathy. Misdiagnosis occurs with Alport's syndrome which shares similar histological features. No treatment is of known benefit.

RAPIDLY PROGRESSIVE GLOMERULONEPHRITIS (RPGN)

RPGN is a syndrome with glomerular haematuria (RBC casts or dysmorphic RBCs), rapidly developing acute renal failure over weeks to months and focal glomerular necrosis (Fig. 11.22) with or without glomerular crescent development on renal biopsy. The 'crescent' is an aggregate of macrophages and epithelial cells in Bowman's space (Fig. 11.22). RPGN can develop with immune deposits (anti-GBM or immune complex type, e.g. SLE) or without immune deposits (pauci-immune, e.g. anti-PR3 and or anti-MPO-ANCA positive vasculitides). It can also develop as an idiopathic primary glomerular disease or can be superimposed on secondary glomerular diseases such as IgA nephropathy, membranous GN and postinfective GN. The classification used here is based on the immunofluorescence information obtained from renal histology (Table 11.6) viz linear, granular and negative immunofluorescent patterns (Table 11.6).

Table 11.6 Types of rapidly progressive glomerulonephritis (RPGN)

Linear immunofluorescent pattern (Fig. 11.23a)
Idiopathic anti-GBM antibody-mediated RPGN
Goodpasture's syndrome

Granular immunofluorescent pattern (immune complex-mediated RPGN) (Fig. 11.23b)
Idiopathic immune complex-mediated RPGN
Associated with other primary GN
- Mesangiocapillary GN (type II > type I)
- IgA nephropathy
- Membranous glomerulopathy

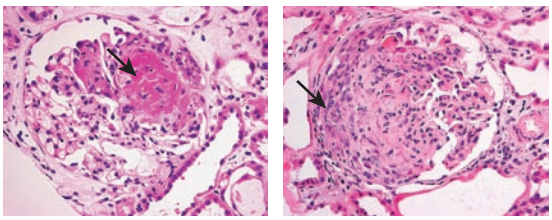
Associated with secondary GN
- Post-infectious GN
- Systemic lupus erythematosus
- Henoch–Schönlein syndrome
- Cryoglobulinaemia

Negative immunofluorescent pattern (pauci-immune RPGN)
ANCA-associated systemic vasculitides

Anti-GBM glomerulonephritis (Fig. 11.23a)

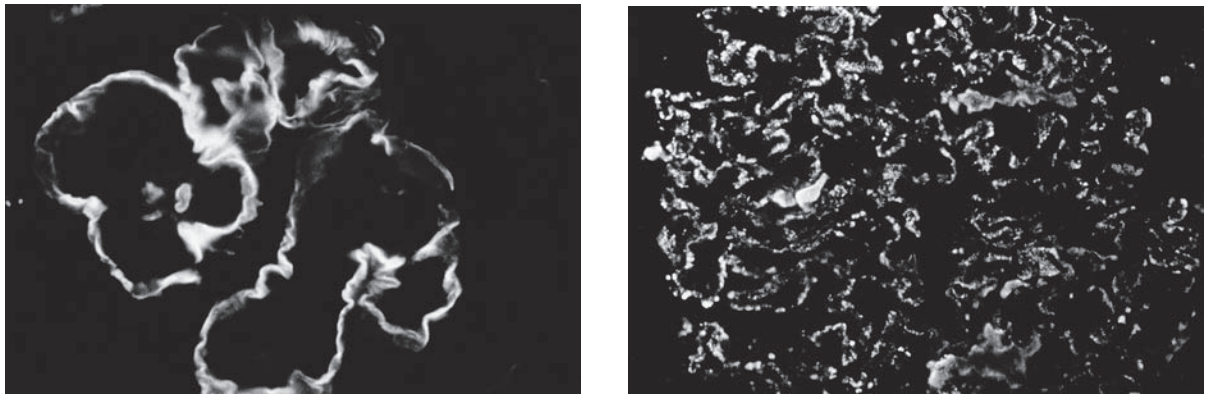
Anti-GBM glomerulonephritis, characterized by linear capillary loop staining with IgG and C3 and extensive crescent formation, accounts for 15–20% of all cases of RPGN, although overall it accounts for less than 5% of all forms of glomerulonephritis. This condition is rare, with an incidence of 1 per 2 million in the general population. About two-thirds of these patients have Goodpasture's syndrome with associated lung haemorrhage (p. 874). The remainder have a renal restricted anti-GBM RPGN, which is seen in people over 50 years and affects both genders equally.

Anti-GBM antibodies (detected by ELISA) are present in serum and are directed against the *non-collagenous (NCI) component* of α3 (IV) collagen of basement membrane. This target antigen must be present as a component of the native α3, α4, α5 (IV) network of selected basement membrane in order for pulmonary and renal disease to develop. Consequently, there are no known cases of anti-GBM glomerulonephritis in patients with Alport's syndrome (p. 594). Anti-GBM RPGN is restricted by the major histocompatibility complex; HLA-DRB1*1501 and HLA-DRB1*1502 alleles increase susceptibility, whereas HLA-DR7 and HLA-DR1 are protective. The thymus expresses α3 (IV) NCI peptides that can eliminate autoreactive $CD4^+$ helper T cells, but a few such cells escape deletion and are kept in check by circulating regulatory cells. Breakdown of this peripheral tolerance (the mechanism of which is unknown) results in these autoreactive $CD4^+$ cells producing anti-GBM antibodies. These antibodies are very specific as shown by the fact that anti-



(a) (b)

Fig. 11.22 **Rapidly progressive glomerulonephritis (RPGN).** Arrows show 'crescents' with aggregates of macrophages and epithelial cells in Bowman's space. **(a)** Focal necrotizing glomerulonephritis. **(b)** Crescentic glomerulonephritis.



(a) (b)

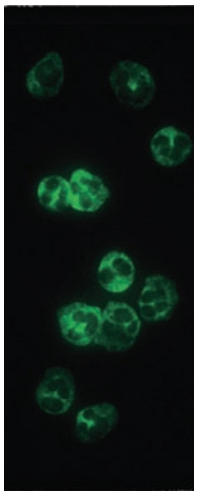
Fig. 11.23 **Immunofluorescence. (a)** Showing antiglomerular basement membrane antibody (anti-GBM) deposition in a linear pattern typical of Goodpasture's syndrome. **(b)** Showing immune complex deposition in a diffuse granular pattern.

bodies against α1, α1 and α2 NCl domains do not cause RPGN. Since the α3 (IV) NCl epitope is hidden within the α3, α4 and α5 (IV) promoter, it is presumed that an environmental factor, such as exposure to hydrocarbons or tobacco smoke, is required in order to reveal cryptic epitopes to the immune system.

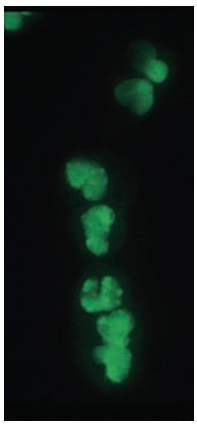
The mechanism of renal injury is complex. When anti-GBM antibody binds basement membrane it activates complement and proteases and results in disruption of the filtration barrier and Bowman's capsule, causing proteinuria and the formation of crescents. Crescent formation is facilitated by interleukin-12 and gamma-interferon which are produced by resident and infiltrating inflammatory cells.

Management

This is based on counteracting the factors involved in the pathogenesis. Thus plasma exchange is used to remove circulating antibodies, steroids to suppress inflammation from antibody already deposited in the tissue, and cyclophosphamide to suppress further antibody synthesis. The prognosis is directly related to the extent of glomerular damage (measured by percentage of crescents, serum creatinine and need for dialysis) at the initiation of treatment. When oliguria occurs or serum creatinine rises above 600–700 μmol/L, renal failure is usually irreversible. Once the active disease is treated, this condition, unlike other autoimmune diseases, does not follow a remitting/ relapsing course. Furthermore, if left untreated, autoantibodies diminish spontaneously within 3 years and autoreactive T cells cannot be detected in the convalescent patients. This is suggestive of re-establishment of peripheral tolerance which coincides with re-emergence of regulatory $CD25^+$ cells in the peripheral blood; these play a key role in inhibiting the autoimmune response. The emergence and persistence of these regulatory cells may underlie the 'single hit' nature of this condition.



(a) cANCA.



(b) pANCA.

Fig. 11.25 **ANCA positive glomerulonephritis – immunofluorescence.** From *The Shrier Atlas of Diseases of the Kidney* Vol IV with permission of Wiley-Blackwell.

ANCA-positive vasculitides (see also p. 551)

Inflammation and necrosis of the blood vessel wall occurs in many primary vasculitic disorders. *Wegener's granulomatosis, microscopic polyangiitis* and *Churg–Strauss syndrome* are described as small vessel vasculitides and are commonly associated with antineutrophil cytoplasm antibodies (ANCA). These diseases share common pathology with focal necrotizing lesions, which affect many different vessels and organs; in the lungs, a capillaritis may cause lung haemorrhage; within the glomerulus of the kidney, crescentic GN and/or focal necrotizing lesions may cause acute renal failure (Fig. 11.22a, b); in the dermis, a purpuric rash or vasculitic (Fig. 11.24) ulceration. Wegener's and Churg–Strauss syndrome may have additional granulomatous lesions.

Pathogenesis

There are two forms of ANCA (p. 504) viz PR3-ANCA (cANCA) and MPO-ANCA (pANCA). If ELISA and indirect immunofluorescence techniques (Fig. 11.25a,b) are combined, diagnostic specificity is 99%. Testing for antineutrophil cytoplasmic antibodies should be accompanied by appropriate tests of autoantibodies directed against DNA and the glomerular basement membrane antigen. The simultaneous occurrence of ANCA and anti-GMB antibody is well documented; such patients tend to follow the natural history of Goodpasture's disease. Variations in the ANCA titres have been used in the assessment of disease activity.

Fig. 11.24 **Vasculitic rash.**

- PR3-ANCA positivity is found in the large majority (>90%) of patients with active Wegener's granulomatosis and in up to 50% of patients with microscopic polyangiitis.
- Anti-MPO positivity is present in the majority of patients with idiopathic crescentic glomerulonephritis and in a variable number of cases of microscopic polyangiitis. There is some evidence to suggest that ANCA are pathogenic and not just markers of disease; for example, development of drug-induced ANCA is associated with vasculitic lesions in humans. Churg–Strauss syndrome may have either anti-MPO- or anti-PR3-ANCA.
- Positivity for both types of ANCA antibodies occurs in up to 10% of patients who have a variable clinical course but a worse renal outcome.
- *Drugs* (e.g. propylthiouracil, hydralazine, minocycline, penicillamine) may induce vasculitides associated with ANCA. Most patients reported with drug-induced ANCA-associated vasculitis have MPO-ANCA, often in very high titres. In addition to MPO-ANCA, most also have antibodies to elastase or to lactoferrin. A relatively small number have PR3-ANCA. Many cases of drug-induced ANCA-associated vasculitis present with constitutional symptoms, arthralgias/arthritis, and cutaneous vasculitis. However, the full range of clinical features associated with ANCA, including crescentic GN and lung haemorrhage, can also occur.
- *Autoimmunity.* It is unclear how and why autoimmunity causes the formation of ANCA antibodies. Patients with anti-PR3 also have autoantibodies to a peptide translated from the antisense DNA strand of PR-3 (complementary PR-3; cPR-3) or to a mimetic of this peptide. This suggests that autoimmunity can be initiated through an immune response against a peptide that is antisense or complementary to the autoantigen, which then induces anti-idiotypic antibodies (autoantibodies) that cross-react with the autoantigen.

There may be multiple factors that contribute to the initiation of an ANCA autoimmune response and the induction of injury by ANCA, such as genetic predisposition (α_1-antitrypsin deficiency; Pi-Z allele) and environmental factors (e.g. silica exposure, viral infection, *Staph. aureus* infection) can result in high local or systemic pro-inflammatory cytokines such as tumour necrosis factor (TNF).

Treatment

The sooner treatment is instituted the more chance there is of recovery of renal function.

- *Corticosteroids* and *cyclophosphamide* are of benefit: high-dose oral prednisolone (maximum 80 mg/day reducing over time to 15 mg/day by 3 months) and cyclophosphamide (2 mg/kg/day, adjusted for age, renal function and prevailing WBC count). Intravenous pulse, rather than daily oral, cyclophosphamide is associated with an equivalent response with better side-effect profile but is associated with higher relapse rate. The best indicators of prognosis are pulmonary haemorrhage and severity of renal failure at presentation.
- Patients who present with *fulminant disease* need intensification of immunosuppression with adjuvant plasma exchanges (7 × 3–4 L over 14 days) or intravenous pulse methyl prednisolone (1 g/day for 3 consecutive days). Plasma exchange appeared to have better outcome than pulse methyl prednisolone in one study.
- Once remission has been achieved, azathioprine should be substituted for cyclophosphamide. In cases of intolerance to azathioprine or cyclophosphamide, mycophenolate or methotrexate has been tried with some success.
- Colonization of the upper respiratory tract with *Staph. aureus* increases the risk of relapse, and treatment with sulfamethoxazole/trimethoprim reduces the relapse rate.
- Relapse after complete cessation of immunosuppressive therapy has been observed relatively frequently, and therefore long-term, albeit relatively low-dose, immunosuppression is necessary.
- Intravenous immunoglobulin (anti-thymocyte globulin, ATG, directed against activated T lymphocytes causes lymphopenia), lymphocyte-depleting anti-CD52 (campath-IH) antibodies, anti-CD20 antibody (Rituximab) a B lymphocyte depleting antibody and anti-TNF therapy have shown promise in the treatment of severe and drug-resistant cases as induction therapy. However, an anti-TNF agent, etanercept, has been ineffective as a sole agent for maintenance.

OTHER GLOMERULAR DISORDERS

HIV-associated nephropathy (HIVAN) (see p. 189)

A number of renal lesions have been described in association with HIV infection. These include glomerulonephritis of various histological types and the haemolytic uraemic syndrome. The most common (80–90%) histological abnormality is a focal glomerulosclerosis (FGS).

HIV-associated FGS

A characteristic 'collapsed' appearance of glomeruli is often seen on light microscopy similar to that seen in other causes of focal segmental glomerulosclerosis (see Fig. 11.14c). In HIVAN many visceral epithelial cells (podocytes) are enlarged, hyperplastic, coarsely vacuolated, contain protein absorption droplets and overlie capillaries with varying degrees of wrinkling and collapse of the walls. It is associated with loss of podocyte-specific markers such as Wilms' tumour factor and synaptopodin due to HIV-1 infection of podocytes of patients with HIVAN. HIVAN has striking predilection; over 90% of patients are black. Clinically HIVAN presents with proteinuria in the nephrotic range, oedema and a 'bland' urine. Hypertension is unusual. If untreated, patients go on to renal failure which can be rapid in progression.

IgA may be an integral feature of HIV-1 infection, as is IgA nephropathy. In this setting, HIV antigen may be a part of the glomerular immune complexes and circulating immune complexes.

Highly active antiretroviral therapy (HAART) may result in stabilization of renal function and prevention of progression to end-stage renal failure (efficacy 23%) and HIV-associated mortality in patients with end-stage renal failure. A cyclin-dependent kinase inhibitor, roscovitine, has been successfully used in the treatment of experimental HIVAN.

Fabry's disease

This is the result of deficiency of the enzyme α-galactosidase with accumulation of sphingolipids in many cells. In the kidney, accumulation of sphingolipids especially affects podocytes, which on light microscopy appear enlarged and vacuolated. Ultrastructurally, these inclusion bodies appear as zebra or myeloid bodies representing sphingolipids. Similar appearances have been described in patients taking chloroquine, hydroxychloroquine and amantidine because these can cause hyperlipidosis. These structures can also be found in endothelial, mesangial, and arterial and arteriolar smooth muscle cells. The most common renal manifestation is proteinuria and progressive renal failure.

Treatment (p. 1072).

Sickle nephropathy

Sickle disease or trait is complicated relatively commonly by papillary sclerosis or necrosis, nephrogenic diabetes insipidus and incomplete renal tubular acidosis. Glomerular lesions are rare and can sometimes be traced to hepatitis B or C infection acquired through repeated blood transfusions. Occasionally, proteinuria or nephrotic syndrome with progressive renal insufficiency is seen without prior infection. The rare glomerular lesion is that of membranous GN or membranoproliferative GN with IgG deposits. No form of effective therapy is known.

Glomerulopathy associated with pre-eclampsia

The glomerular lesion of pre-eclampsia is characterized by marked endothelial swelling and obliteration of capillary lumina. Fibrinogen–fibrin deposits may be found in the mesangium. The renal lesion may not be reversible and 30% of patients have changes for 6 months or longer. Patients who have had pre-eclampsia are more likely to develop hypertension in subsequent pregnancies. Severe proteinuria may occur during the course of pre-eclampsia and from time to time produce features of nephrotic syndrome. Ordinarily, proteinuria disappears after delivery.

In severe cases, associated with cortical necrosis, there may be microangiopathic haemolytic anaemia. Vascular endothelial growth factor (VEGF) and placental growth factor (PLGF) play a key role in the development of the placenta.

Relative deficiency of either factor can theoretically cause implantation abnormalities normally seen in pre-eclampsia. A soluble fms-like tyrosine kinase (sFlt1) receptor also called VEGF-receptor, which is an antagonist of PLGF and specifically of VEGF, is upregulated in the placenta of patients with pre-eclampsia. High circulating levels of these receptors antagonize angiopoeitic factors and cause endothelial dysfunction. Excessive free radical generation in the placenta of pre-eclamptic patients is due to upregulation of NADPH oxidase activity caused by generation of an angiotensin II receptor agonist antibody in some patients.

FURTHER READING

Dooley MA, Falk RJ. Human clinical trials in lupus nephritis. *Seminars in Nephrology* 2007; **27**(1): 115–127.

Jayne D. Challenges in the management of microscopic polyangiitis: past, present and future. *Current Opinion in Rheumatology* 2008; **20**(1): 3–9.

Jennette JC, Falk RJ. New insight into the pathogenesis of vasculitis associated with antineutrophil cytoplasmic autoantibodies. *Current Opinion in Rheumatology* 2008; **20**(1): 55–60.

Perysinaki G, Panagiotakis S, Bertsias G, Boumpas DT. Pharmacotherapy of lupus nephritis: time for a consensus? *Expert Opinion on Pharmacotherapy* 2008; **9**(12): 2099–2115.

Rodgers M et al. Diagnostic tests and algorithms used in the investigation of haematuria: systematic reviews and economic evaluation. *Health Technology Assessment* 2006; **10**(18): iii–iv, xi–259.

Vikse BE, Irgens LM, Leivestad T, Skjaerven R, Iversen BM. Preeclampsia and the risk of end-stage renal disease. *New England Journal of Medicine* 2008; **359**(8): 800–809.

RENAL INVOLVEMENT IN OTHER DISEASES

Polyarteritis nodosa (PAN)

(see also p. 550)

Classical PAN is a multisystem disorder. Aneurysmal dilatation of medium-sized arteries may be seen on renal arteriography. The condition is more common in men and in the elderly and, typically, the patient is ANCA negative. Hypertension, polyneuropathy and features indicating ischaemic infarction of various organs including the kidney are presenting features. It may be associated with drug abuse and hepatitis B infection. This form of polyangiitis is associated with slowly progressive renal failure, often accompanied by severe hypertension. Rapidly progressive renal failure is rare. Treatment with immunosuppression is less effective than it is for microscopic polyangiitis.

Haemolytic uraemic syndrome (HUS)

HUS is characterized by intravascular haemolysis with red cell fragmentation (microangiopathic haemolysis), thrombocytopenia and acute kidney injury due to thrombosis in small arteries and arterioles (Fig. 11.26). These features are also seen in disseminated intravascular coagulation, but coagulation tests are typically normal in HUS. The syndrome often

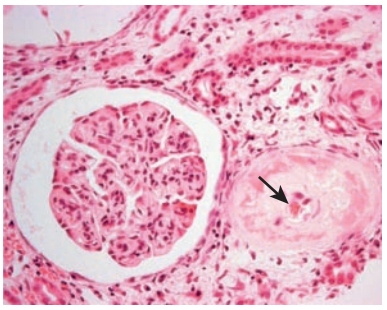


Fig. 11.26 **Typical haemolytic uraemic syndrome (HUS) renal lesion – light microscopy.** Arrow shows microthrombi.

follows a febrile illness, particularly gastroenteritis (also known as *diarrhoea-associated HUS* (*D+HUS*) often associated with *Escherichia coli*, notably strain O157. This strain of *E. coli* produces verocytotoxin (shiga toxin), which has an A unit and five B units. The A unit is pathogenic by inhibiting protein synthesis and initiating endothelial damage. The role of B units is to facilitate the entry of the A unit into the endothelial cells by binding to a receptor (Gb3) on the endothelial cell. The toxins are transported to endothelial cells from the gut on neutrophils. Most patients with D+HUS recover renal function, but supportive care including maintenance of fluid and electrolyte balance, antihypertensive medication, nutritional support and dialysis is commonly required. Plasmapheresis is not beneficial but is usually tried as a last resort. About 5% die during the acute episode, 5% develop chronic renal failure and 30% exhibit evidence of long-term damage with persistent proteinuria. Antibiotic and antimotility agents for the diarrhoea increase the risk of HUS and its complications.

Recurrent episodes of HUS have been described in the same individual, and familial forms of the disease (with both recessive and dominant inheritance) exist.

The *non-diarrhoeal-induced form of HUS (D–HUS)* may be a complement-driven illness due to a deficiency of complement factor H. Factor H is a soluble protein produced by the liver, which regulates the activity of the alternative complement activation pathway; in particular, it protects host cell surfaces from complement-mediated damage. In some families with D–HUS, a mutation has been traced to another complement regulatory protein known as membrane cofactor protein (MCP). This protein is highly expressed in the kidney and normally prevents glomerular C3 activation. A loss of function mutation in MCP results in unopposed complement activation and development of HUS. *Treatment* is often very difficult because of severe hypertension and the possibility of frequent recurrences. The course of the disease is often indolent and progressive. Plasmapheresis or plasma infusion, although unproven, is still used. Liver transplantation is potentially the only curative treatment.

Sporadic cases of D–HUS can be associated with pregnancy, SLE, scleroderma, malignant hypertension, metastatic cancer, HIV infection and various drugs including oral contraceptives, ciclosporin, tacrolimus, chemotherapeutic agents (e.g. cisplatin, mitomycin C, bleomycin) and heparin. Treatment is supportive with removal of the offending agent or specific treatment of the underlying cause. There is no

evidence in favour of plasma infusion or plasmapheresis in these sporadic cases but it is tried, usually as a last resort.

Pneumococcus-associated HUS. This rare complication of *Streptococcus pneumoniae* infection was previously associated with a high morbidity and mortality. This organism produces an enzyme (possibly neuroaminidase) which can expose an antigen (Thomsen antigen) present on RBCs, platelets and glomeruli. Antibodies to the Thomsen antigen result in an antigen–antibody reaction and can lead to HUS and anaemia. The improved outcome is due to increasing awareness of this complication, judicious use of blood products (washed blood products) and avoiding plasma infusion or plasmapheresis.

Thrombotic thrombocytopenic purpura (TTP) (see p. 437)

TTP is characterized by microangiopathic haemolysis, renal failure and evidence of neurological disturbance. Young adults are most commonly affected.

Antiphospholipid syndrome (APS) (see p. 544)

The central feature of APS is recurrent thrombosis (both venous and arterial) and fetal loss in the presence of antiphospholipid antibodies. Such antibodies may be primary or secondary to infections (HIV, hepatitis C) or autoimmune disease (SLE). Fifty per cent have renal involvement with proteinuria. Thrombotic microangiopathy is a rare but well-recognized presentation. In some cases a lupus nephritis-like (usually membranous GN) lesion is seen. The only proven treatment for APS is warfarin with an INR of 3–4. Use of steroids or plasmapheresis is reserved for patients with APS and life-threatening renal involvement with thrombotic microangiopathy. Treatment is variably successful (30–70%).

Multiple myeloma

Acute kidney injury is relatively common in myeloma, occurring in 20–30% of affected individuals at the time of diagnosis, and is mainly due to the nephrotoxic effects of the abnormal immunoglobulins. It is often irreversible. The following types of renal lesions are associated with myeloma.

1. *Light chain cast nephropathy* – intratubular deposition of light chains, particularly kappa chains facilitated by Tamm–Horsfall glycoprotein, which characteristically appear on renal histology as fractured casts with giant cell reaction (Fig. 11.27).
2. *AL amyloidosis* – deposition of amyloid fibrils of light chains (Congo red positive).
3. *Light chain deposition disease* – nodular glomerulosclerosis with granular deposits of usually lambda light chains (Congo red negative).
4. *Plasma cell infiltration* – often incidental finding at autopsy.
5. *Fanconi's syndrome* – tubular toxicity due to light chains.
6. *Hypercalcaemic nephropathy* – bone resorption causing hypercalcaemia.
7. *Hyperuricaemic nephropathy* – tumour lysis causing tubular crystallization of uric acid.

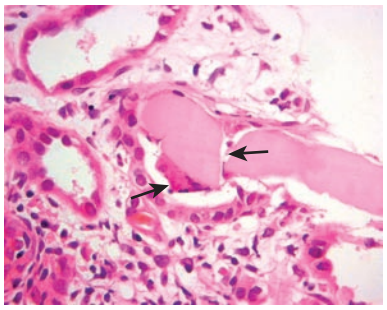


Fig. 11.27 **Cast nephropathy in a patient with multiple myeloma.** Light microscopy picture showing characteristic fractured cast and giant cell reaction (arrows).

8. *Radiocontrast nephropathy* – interaction between light chains and radiocontrast.

Treatment of underlying myeloma is indicated (p. 485).

URINARY TRACT INFECTION

Urinary tract infection (UTI) is common in women, in whom it usually occurs in an anatomically normal urinary tract. Conversely, it is uncommon in men and children, and the urinary tract is often abnormal and requires investigation. The incidence of UTI is 50 000 per million persons per year and accounts for 1–2% of patients in primary care. Recurrent infection causes considerable morbidity; if complicated, it can cause severe renal disease including end-stage renal failure. It is also a common source of life-threatening Gram-negative septicaemia.

Aetiology and pathogenesis

Infection is most often due to bacteria from the patient's own bowel flora (Table 11.7). Transfer to the urinary tract is most often via the ascending transurethral route but may be via the bloodstream, the lymphatics or by direct extension (e.g. from a vesicocolic fistula).

Symptomatic infection is related to the virulence of the organisms, which competes with the innate host defence system. However, inflammation and injury are determined by the host response and not by the bacterium.

Virulence. Ability to adhere to epithelial cells determines the degree of virulence of the organism. For *E. coli*, these adhesive factors include flagellae (for motility), aerobactin (for iron acquisition in the iron-poor environment of the urinary tract), haemolysin (for pore forming) and above all, the presence of adhesins on the bacterial fimbriae and on the cell surface.

There are two types of *E. coli*: those with *type 1 fimbriae* (with adhesin known as FimH) associated with cystitis; and those with *type P fimbriae* (with adhesin known as PapG) commonly responsible for pyelonephritis. Bacterial adhesins are necessary for attachment of bacteria to the mucous membranes of the perineum and urothelium. There are several molecular forms of adhesins. The most studied is the PapG adhesin, which is located on the tip of P fimbriae. This lectin (one of the P blood group antigens) structure recognizes binding sites consisting of oligosaccharide sequences present on the mucosal surface.

FURTHER READING

Tarr PL, Gordon CA, Chandler WL. Shiga-toxin-producing *E. coli* and haemolytic uraemic syndrome. *Lancet* 2005; **365**: 1073–1086.

Tsai HM. Advances in the pathogenesis, diagnosis and the treatment of thrombotic thrombocytopenic purpura. *Journal of the American Society of Nephrology* 2003; **14**: 1072–1081.

Table 11.7 Organisms causing urinary tract infection in domiciliary practice

Organism	Approximate frequency (%)
Escherichia coli and other 'coliforms'	68+
Proteus mirabilis	12
*Klebsiella aerogenes**	4
*Enterococcus faecalis**	6
Staphylococcus saprophyticus or *epidermidis*†	10

*More common in hospital practice
†More common in young women (20–30%)

Innate host defence. The following host defence mechanisms are necessary to prevent UTI:

- *Neutrophils* – adhesins activate receptors, e.g. Toll receptor 4, on the mucosal surface, resulting in IL-8 production and expression of its receptor CXCR1 on neutrophil surfaces. Activation of neutrophils is essential for bacterial killing. Defective IL-8 production or reduced expression of CXCR1 results in impaired function of neutrophils predisposing an individual to severe UTI.
- *Urine osmolality and pH* – urinary osmolality >800 mOsm/kg and low or high pH reduce bacterial survival.
- *Complement* – complement activation with IgA production by uroepithelium (acquired immunity) also plays a major role in defence against UTI.
- *Commensal organisms* – such as lactobacilli, corynebacteria, streptococci and bacteroides are part of the normal host defence. Eradication of these commensal organisms by spermicidal jelly or disruption by certain antibiotics results in overgrowth of *E. coli*.
- *Urine flow* – urine flow and normal micturition wash out bacteria. Urine stasis promotes UTI.
- *Uroepithelium* – mannosylated proteins such as Tamm–Horsfall proteins (THP), which are present in the mucus and glycocalyx covering uroepithelium, have antibacterial properties. These proteins interfere with bacterial binding to uroepithelium. Disruption of this uroepithelium by trauma (e.g. sexual intercourse or catheterization) predisposes to UTI. Cranberry juice and blueberry juice contain a large-molecular-weight factor (proanthrocyanidins) that prevents binding of *E. coli* to the uroepithelium (see p. 603).
- *Blood group antigens* – women who are non-secretors of ABH blood group antigens are three to four times more likely to have recurrent UTIs.

Natural history

UTI is commonly an isolated, rather than a repeated, event (Fig. 11.28).

Complicated versus uncomplicated infection (Fig. 11.29)

It is necessary to distinguish between UTI occurring in patients with functionally normal urinary tracts and in those with abnormal tracts.

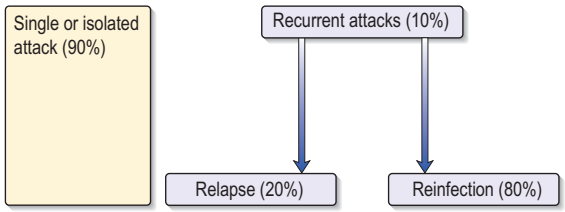


Fig. 11.28 The natural history of urinary tract infection.

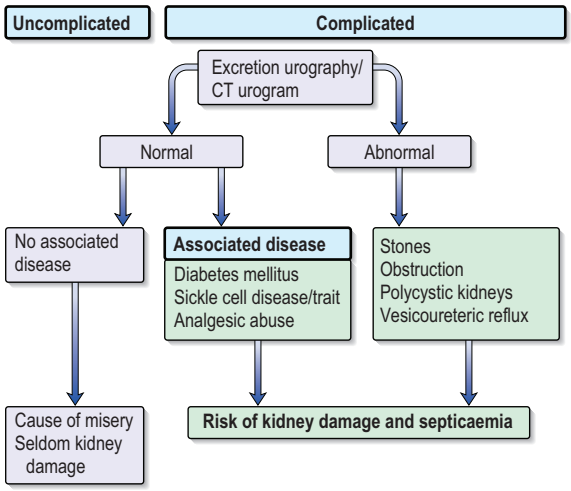


Fig. 11.29 Complicated versus uncomplicated urinary tract infection.

Functionally normal urinary tracts (with normal renal imaging). Here, persistent or recurrent infection seldom results in serious kidney damage (uncomplicated UTI).

Abnormal urinary tracts. Tracts with stones, or associated diseases such as diabetes mellitus which themselves cause kidney damage, may be made worse with infection (complicated UTI). UTI, particularly with *Proteus*, may predispose to stone formation. The combination of infection and obstruction results in severe, sometimes rapid, kidney damage (obstructive pyonephrosis) and is a major cause of Gram-negative septicaemia.

Acute pyelonephritis

The combination of fever, loin pain with tenderness and significant bacteriuria usually implies infection of the kidney (acute pyelonephritis). Small renal cortical abscesses and streaks of pus in the renal medulla are often present. Histologically there is focal infiltration by polymorphonuclear leucocytes and many polymorphs in tubular lumina.

Although, with antibiotics, significant permanent kidney damage in adults with normal urinary tracts is rare, CT scanning can show wedge-shaped areas of inflammation in the renal cortex (Fig. 11.30) and hence damage to renal function.

Reflux nephropathy

This was called chronic pyelonephritis or atrophic pyelonephritis, and it results from a combination of:

- vesicoureteric reflux, and
- infection acquired in infancy or early childhood.

Fig. 11.30 **CT scan showing a wedge-shaped area of renal cortical loss (arrow) following acute pyelonephritis.**

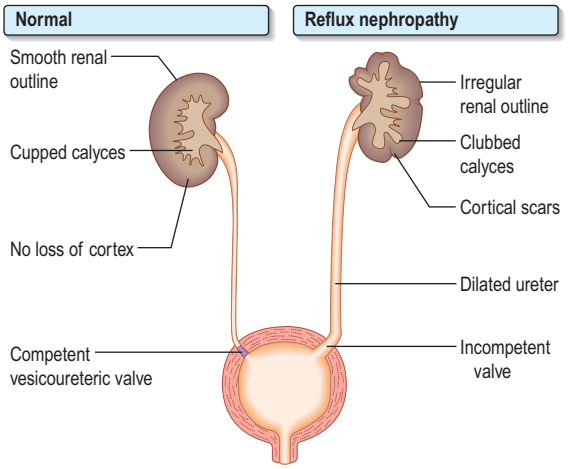


Fig. 11.31 **Findings in reflux nephropathy compared with normal.**

Normally the vesicoureteric junction acts as a one-way valve (Fig. 11.31), urine entering the bladder from above; the ureter is shut off during bladder contraction, thus preventing reflux of urine. In some infants and children – possibly even in utero – this valve mechanism is incompetent, bladder voiding being associated with variable reflux of a jet of urine up the ureter. A secondary consequence is incomplete bladder emptying, as refluxed urine returns to the bladder after voiding. This latter event predisposes to infection, and the reflux of infected urine leads to kidney damage.

Typically there is papillary damage, tubulointerstitial nephritis and cortical scarring in areas adjacent to 'clubbed calyces'.

Diagnosis is based on CT scan of the kidneys, which shows irregular renal outlines, clubbed calyces and a variable reduction in renal size. The condition may be unilateral or bilateral and affect all or part of the kidney.

Reflux usually ceases around puberty with growth of the bladder base. Damage already done persists and progressive renal fibrosis and further loss of function occur in severe cases even though there is no further infection.

Reflux nephropathy cannot occur in the absence of reflux and therefore does not begin in adult life. Consequently, adult females with bacteriuria and a normal urogram can be reassured that kidney damage will not develop.

Chronic reflux nephropathy acquired in infancy predisposes to hypertension in later life and, if severe, is a relatively common cause of end-stage renal failure in childhood or adult life. Meticulous early detection and control of infection, with or without ureteral reimplantation to create a competent valve, can prevent further scarring and allow normal growth of the kidneys. No proof exists, however, that reimplantation surgery confers long-term benefit.

Reinfection versus relapsing infection

When UTI is recurrent it is necessary to distinguish between relapse and reinfection.

Relapse is diagnosed by recurrence of bacteriuria with the same organism within 7 days of completion of antibacterial treatment and implies failure to eradicate infection (Fig. 11.32) usually in conditions such as stones, scarred kidneys, polycystic disease or bacterial prostatitis.

Reinfection is when bacteriuria is absent after treatment for at least 14 days, usually longer, followed by recurrence of infection with the same or different organisms. This is not due to failure to eradicate infection, but is the result of reinvasion of a susceptible tract with new organisms. Approximately 80% of recurrent infections are due to reinfection.

Symptoms and signs of UTI

The most typical symptoms of UTI are:

- frequency of micturition by day and night
- painful voiding (dysuria)
- suprapubic pain and tenderness
- haematuria
- smelly urine.

These symptoms relate to bladder and urethral inflammation, commonly called 'cystitis', and suggest lower urinary tract infection. Loin pain and tenderness, with fever and systemic upset, suggest extension of the infection to the pelvis and kidney, known as pyelitis or pyelonephritis. However, localization of the site of infection on the basis of symptoms alone is unreliable.

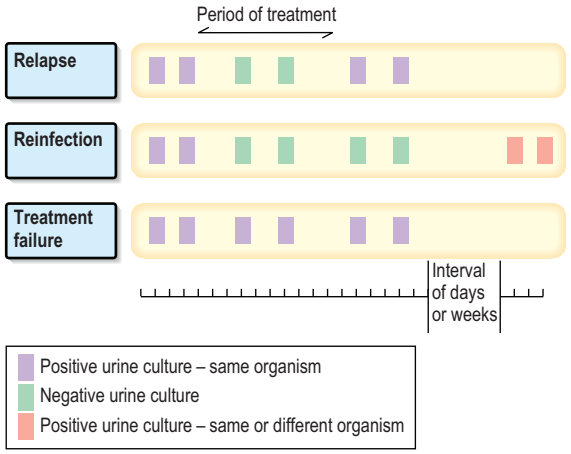


Fig. 11.32 **A comparison of reinfection, relapse and treatment failure in urinary tract infection.**

Ask the authors
www.studentconsult.com

UTI also presents with minimal or no symptoms or may be associated with atypical symptoms such as abdominal pain, fever or haematuria in the absence of frequency or dysuria.

In small children, who cannot complain of dysuria, symptoms are often 'atypical'. The possibility of UTI must always be considered in the fretful, febrile sick child who fails to thrive.

Diagnosis

This is based on quantitative culture of a clean-catch midstream specimen of urine and the presence or absence of pyuria. The criteria for the diagnosis of UTI, particularly in symptomatic women, are shown in Table 11.8. Most Gram-negative organisms reduce nitrates to nitrites and produce a red colour in the reagent square. False-negative results are common. Dipsticks that detect significant pyuria depend on the release of esterases from leucocytes. Dipstick tests positive for both nitrite and leucocyte esterase are highly predictive of acute infection (sensitivity of 75% and specificity of 82%).

Abacteriuric frequency or dysuria ('urethral syndrome')

Causes of truly abacteriuric frequency/dysuria include postcoital bladder trauma, vaginitis, atrophic vaginitis or urethritis in the elderly, and interstitial cystitis (Hunner's ulcer). In symptomatic young women with 'sterile pyuria', *Chlamydia* infection and tuberculosis must be excluded.

Interstitial cystitis is an uncommon but distressing complaint, most often affecting women over the age of 40 years. It presents with frequency, dysuria and often severe suprapubic pain. Urine cultures are sterile. Cystoscopy shows typical inflammatory changes with ulceration of the bladder base. It is commonly thought to be an autoimmune disorder. Various treatments are advocated with variable success. These include oral prednisolone therapy, bladder instillation of sodium cromoglicate or dimethyl sulphoxide and bladder stretching under anaesthesia.

Predominant frequency and passage of small volumes of urine ('irritable bladder') is possibly consequent on previous UTI or conditioned by psychosexual factors. Such patients must be distinguished from those with frequency due to polyuria. Repeated courses of antibiotics in patients with genuine abacteriuric frequency or dysuria are quite inappropriate and detract from identifying the true nature of the problem.

Table 11.8 Criteria for diagnosis of bacteriuria

Symptomatic young women

$\geq 10^2$ coliform organisms/mL urine plus pyuria (>10 WCC/mm^3)

OR

$\geq 10^5$ any pathogenic organism/mL urine

OR

any growth of pathogenic organisms in urine by suprapubic aspiration

Symptomatic men

$\geq 10^3$ pathogenic organisms/mL urine

Asymptomatic patients

$\geq 10^5$ pathogenic organisms/mL urine on two occasions

Special investigations

Uncomplicated UTI usually does not require radiological evaluation unless it is recurrent or affecting males and children or there are unusually severe symptoms. Patients with predisposing conditions such as diabetes mellitus or immunocompromised states benefit from early imaging.

- *Ultrasound* is used in the assessment of patients with suspected pyelonephritis that requires drainage. This allows the detection of calculi, obstruction and also incomplete emptying.
- *CT* is a more sensitive modality for diagnosis and follow-up of complicated renal tract infection. Contrast-enhanced CT allows different phases of excretion to be studied and can define the extent of disease and identify significant complications or obstruction.
- *MRI* is particularly useful in those with iodinated contrast allergies, offering an ionizing radiation-free alternative in the diagnosis of both medical and surgical diseases of the kidney.
- *Nuclear medicine* has a limited role in the evaluation of UTI in adults. Its main role is in the assessment of renal function and detection of scars by DMSA scan, often prior to surgery.

Treatment

Single isolated attack

Pretreatment urine culture is desirable.

- *Antibiotics* for 3–5 days with amoxicillin (250 mg three times daily), nitrofurantoin (50 mg three times daily), trimethoprim (200 mg twice daily) or an oral cephalosporin. The treatment regimen is modified in light of the result of urine culture and sensitivity testing, and/or the clinical response.
- For *resistant organisms* the alternative drugs are co-amoxiclav or ciprofloxacin.
- *Single-shot treatment* with 3 g of amoxicillin or 1.92 g of co-trimoxazole is used for patients with bladder symptoms of less than 36 hours' duration who have no previous history of UTI.
- A high (2 L daily) fluid intake should be encouraged during treatment and for some subsequent weeks. Urine culture should be repeated 5 days after treatment.
- If the patient is acutely ill with high fever, loin pain and tenderness (acute pyelonephritis), antibiotics are given intravenously, e.g. aztreonam, cefuroxime, ciprofloxacin or gentamicin (2–5 mg/kg daily in divided doses), switching to a further 7 days' treatment with oral therapy as symptoms improve. Intravenous fluids may be required to achieve a good urine output.
- In patients presenting for the first time with high fever, loin pain and tenderness, urgent renal ultrasound examination is required to exclude an obstructed pyonephrosis. If this is present it should be drained by percutaneous nephrostomy (p. 580).

Recurrent infection

Pre-treatment and post-treatment urine cultures are necessary to confirm the diagnosis and identify whether recurrent infection is due to relapse or reinfection.

Relapse a search should be made for a cause (e.g. stones or scarred kidneys), and this should be eradicated. Intense or prolonged treatment – intravenous or intramuscular

aminoglycoside for 7 days or oral antibiotics for 4–6 weeks – is required. If this fails, long-term antibiotics are required.

Reinfection implies that the patient has a predisposition to periurethral colonization or poor bladder defence mechanisms. Contraceptive practice should be reviewed and the use of a diaphragm and spermicidal jelly discouraged. Atrophic vaginitis should be identified in postmenopausal women, who should be treated (see below). All patients must undertake prophylactic measures:

- a 2 L daily fluid intake
- voiding at 2- to 3-hour intervals with double micturition if reflux is present
- voiding before bedtime and after intercourse
- avoidance of spermicidal jellies and bubble baths and other chemicals in bathwater
- avoidance of constipation, which may impair bladder emptying.

Evidence of impaired bladder emptying on excretion urography/ultrasound requires urological assessment. If UTI continues to recur, treatment for 6–12 months with low-dose prophylaxis (trimethoprim 100 mg, co-trimoxazole 480 mg, cefalexin 125 mg at night, or macrocrystalline nitrofurantoin) is required; it should be taken last thing at night when urine flow is low. Intravaginal oestrogen therapy has been shown to produce a reduction in the number of episodes of UTI in postmenopausal women. Cranberry juice is said to reduce the risk of symptoms and reinfection by 12–20% but studies are limited.

Urinary infections in the presence of an indwelling catheter

Colonization of the bladder by a urinary pathogen is common after a urinary catheter has been present for more than a few days, partly due to organisms forming biofilms. So long as the bladder catheter is in situ, antibiotic treatment is likely to be ineffective and will encourage the development of resistant organisms. Treatment with antibiotics is indicated only if the patient has symptoms or evidence of infection, and should be accompanied by replacement of the catheter.

Infection by *Candida* is a frequent complication of prolonged bladder catheterization. Treatment should be reserved for patients with evidence of invasive infection or those who are immunosuppressed, and should consist of removal or replacement of the catheter. In severe infections continuous bladder irrigation with amphotericin 50 μg/mL is used.

Bacteriuria in pregnancy

The urine of all pregnant women must be cultured as 2–6% have asymptomatic bacteriuria. Whilst asymptomatic bacteriuria in the non-pregnant female seldom leads to acute pyelonephritis and often does not require treatment, acute pyelonephritis frequently occurs in pregnancy under these circumstances. Failure to treat may thus result in severe symptomatic pyelonephritis later in pregnancy, with the possibility of premature labour. Asymptomatic bacteriuria, in the presence of previous renal disease, may predispose to pre-eclamptic toxaemia, anaemia of pregnancy, and small or premature babies. Therefore bacteriuria must always be treated and be shown to be eradicated. Reinfection may require prophylactic therapy. Tetracycline, trimethoprim, sulphonamides and 4-quinolones must be avoided in pregnancy. Amoxicillin and ampicillin, nitrofurantoin and oral cephalosporins may safely be used in pregnancy.

Bacterial prostatitis

Bacterial prostatitis is a relapsing infection which is difficult to treat. It presents as perineal pain, recurrent epididymo-orchitis and prostatic tenderness, with pus in expressed prostatic secretion. Treatment is for 4–6 weeks with drugs that penetrate into the prostate, such as trimethoprim or ciprofloxacin. Long-term low-dose treatment may be required. Prostadynia (prostatic pain in the absence of active infection) may be a very persistent sequel to bacterial prostatitis. Amitriptyline and carbamazepine may alleviate the symptoms.

Renal carbuncle

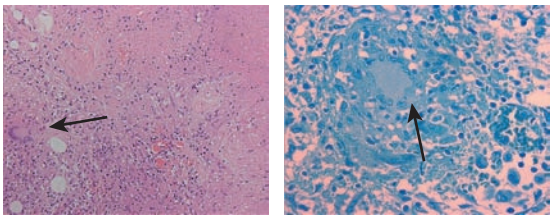
Renal carbuncle is an abscess in the renal cortex caused by a blood-borne *Staphylococcus*, usually from a boil or carbuncle of the skin. It presents with a high swinging fever, loin pain and tenderness, and fullness in the loin. The urine shows no abnormality as the abscess does not communicate with the renal pelvis, more often extending into the perirenal tissue. Staphylococcal septicaemia is common. Diagnosis is by ultrasound or CT scanning. Treatment involves antibacterial therapy with flucloxacillin and surgical drainage.

Tuberculosis of the urinary tract

Tuberculous infection is on the increase world-wide, partly due to the reservoir of infection in susceptible HIV-infected individuals and the emergence of drug-resistant strains. Tuberculosis of the urinary tract presents with frequency, dysuria or haematuria. In the UK it is mainly seen in the Asian immigrant population. Cortical lesions result from haematogenous spread in the primary phase of infection. Most heal, but in some, infection persists and spreads to the papillae, with the formation of cavitating lesions and the discharge of mycobacteria into the urine. Infection of the ureters and bladder commonly follows, with the potential for the development of ureteral stricture and a contracted bladder. Rarely, cold abscesses may form in the loin. In males the disease may present with testicular or epididymal discomfort and thickening.

Diagnosis depends on constant awareness, especially in patients with sterile pyuria. Imaging may show cavitating lesions in the renal papillary areas, commonly with calcification. There may also be evidence of ureteral obstruction with hydronephrosis. Diagnosis of active infection depends on culture of mycobacteria from early-morning urine samples. Imaging may be normal in diffuse interstitial renal tuberculosis when diagnosis is made by renal biopsy demonstrating caseating granuloma with mutinucleate giant cells and acid-fast bacilli on Ziehl–Neelsen staining (Fig. 11.33). Some patients present with small unobstructed kidneys, when the diagnosis is easy to miss.

Treatment. The treatment is as for pulmonary tuberculosis (see p. 866). Renal ultrasonography and/or CT scanning should be carried out 2–3 months after initiation of treatment as ureteric strictures may first develop in the healing phase.



(a) (b)

Fig 11.33 **Renal tuberculosis. (a)** Caseating granulomatous interstitial nephritis showing multinuclear giant cell (arrow). **(b)** Ziehl–Neelsen staining showing myobacteria (arrow).

Xanthogranulomatous pyelonephritis

This is an uncommon chronic interstitial infection of the kidney, most often due to *Proteus*, in which there is fever, weight loss, loin pain and a palpable enlarged kidney. It is usually unilateral and associated with staghorn calculi. CT scanning shows up intrarenal abscesses as lucent areas within the kidney. Nephrectomy is the treatment of choice; antibacterial treatment rarely, if ever, eradicates the infection.

Malakoplakia

This is a rare condition in which plaques of abnormal inflammatory tissue grow within the urinary tract in the presence of urinary infection. The histological appearances are characteristic. It is thought that the condition is caused by an acquired inability of macrophages to kill phagocytosed bacteria. Muscarinic agonists and ascorbic acid may improve macrophage function; ciprofloxacin penetrates the macrophage well and is the antibiotic of choice. Prolonged treatment is often needed.

FURTHER READING

Drekonja DM, Johnson JR. Urinary tract infections. *Primary Care*. 2008; **35**(2): 345–367.

Foster RT Sr. Uncomplicated urinary tract infections in women. *Obstetrics and Gynecology Clinics of North America*. 2008; **35**(2): 235–248.

Table 11.9 Common causes of acute tubulointerstitial nephritis

Drugs (70%)

Antibiotics
Cephalosporins
Ciprofloxacin
Erythromycin
Penicillin
Rifampicin
Sulphonamides

Analgesics
Non-steroidal anti-inflammatory drugs

Diuretics
Furosemide
Thiazides

Miscellaneous
Cimetidine
Phenytoin
Valproate
Carbamazepine
Allopurinol

Infection (15%)
Viruses, e.g. hantavirus
Bacteria, e.g. streptococci

Idiopathic (8%)

Tubulointerstitial nephritis with uveitis (TINU) (5%)

Systemic inflammatory disorders, e.g. systemic lupus erythematosus (SLE)

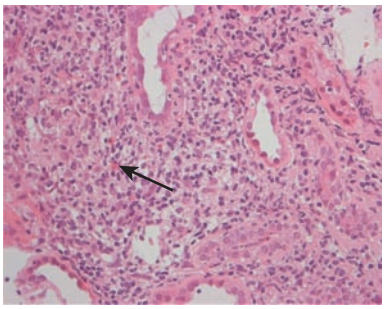


Fig. 11.34 **Tubulointerstitial nephritis showing diffuse interstitial infiltrate with red staining (arrow).**

TUBULOINTERSTITIAL NEPHRITIS (TIN)

Diseases of the kidney primarily affect the glomeruli, vasculature, or the remainder of the renal parenchyma that consists of the tubules and interstitium. Although the tubules and the interstitium are distinct functional entities, they are intimately related. Injury involving one of them invariably results in damage to the other.

Acute tubulointerstitial nephritis (TIN)

In approximately 70% of the cases, acute TIN is due to a hypersensitivity reaction to drugs (Table 11.9), most commonly drugs of the penicillin family and non-steroidal anti-inflammatory drugs (NSAIDs).

Drug induced acute TIN. Patients present with fever, arthralgia, skin rashes and acute oliguric or non-oliguric renal failure. Many have eosinophilia and eosinophiluria. Renal histology shows an intense interstitial cellular infiltrate, often including eosinophils, with variable tubular necrosis (Fig. 11.34). Rarely, NSAIDs can cause a glomerular minimal-change lesion in addition to TIN and present as the nephrotic syndrome. Treatment involves withdrawal of offending drugs. High-dose steroid therapy (prednisolone 60 mg daily) is commonly given but its efficacy has not been proven. Patients may require dialysis for management of the acute kidney injury. Most patients make a good recovery in kidney function, but some may be left with significant interstitial fibrosis and a persistent high serum creatinine.

Infection causing acute TIN. Acute pyelonephritis leads to inflammation of the tubules, producing a neutrophilic cellular infiltrate. TIN can complicate systemic infections with viruses (hantavirus, Epstein–Barr virus, HIV, measles, adenovirus), bacteria (*Legionella, Leptospira*, streptococci, *Mycoplasma, Brucella, Chlamydia*) and others (*Leishmania, Toxoplasma*). Hantavirus causes haemorrhagic fever with TIN and can be fatal. Epstein–Barr virus DNA has been found in renal biopsy tissue of cases of idiopathic TIN. In immunocompromised patients such as post-renal transplantation, CMV, polyoma, and HSV can cause acute TIN in the renal graft. Treatment involves eradication of infection by appropriate antibiotics or antiviral agents.

Acute TIN as part of multisystem inflammatory diseases. Several non-infectious inflammatory disorders such as Sjögren's syndrome, SLE and Wegener's granulomatosis can cause acute or chronic TIN rather than glomerulonephritis. Sjögren's syndrome may additionally present as renal tubular acidosis. Sarcoidosis presents as granulomatous TIN in up to 20% of patients. Associated hypercalcaemia causes acute kidney injury. These heterogeneous conditions with TIN generally respond to steroids.

TINU syndrome. In this syndrome, uveitis generally coincides with acute TIN. It is common in childhood, but has been reported in adulthood. Among adults it is more common in females, but its cause remains unknown. Patients present with weight loss, anaemia and raised ESR. A prolonged course of steroids leads to improvement in both renal function and uveitis.

Chronic tubulointerstitial nephritis

The major causes of chronic tubulointerstitial nephritis are given in Table 11.10. It is characterized by generalized chronic inflammatory cellular infiltation of the interstitium with tubular atrophy and generalized interstitial oedema or fibrosis. In many cases no cause is found. Chronic TIN changes evolve into progressive primary glomerular or vascular disease of the kidney, where its severity is a better predictor of long-term renal survival than the primary site of insult.

The patient usually presents with either polyuria and nocturia, or is found to have proteinuria or uraemia. Proteinuria is usually slight (less than 1 g daily). Papillary necrosis with ischaemic damage to the papillae occurs in a number of tubulointerstitial nephritides, for example in analgesic abuse, diabetes mellitus, sickle cell disease or trait. The papillae can separate and be passed in the urine. Microscopic or overt haematuria or sterile pyuria also occurs, and occasionally a sloughed papilla may cause ureteric colic or produce acute ureteric obstruction. The radiological appearances must be distinguished from those of reflux nephropathy (Fig. 11.35) or obstructive uropathy which is usually accompanied by tubular dilatation and atrophy and intense interstitial fibrosis with patchy inflammatory cellular infiltrate in the scarred areas.

Tubular damage to the medullary area of the kidney leads to defects in urine concentration and sodium conservation with polyuria and salt wasting. Fibrosis progressing into the cortex leads to loss of excretory function and uraemia.

Table 11.10 Causes of chronic tubulointerstitial nephritis

Drugs and toxins
e.g. All causes of ATN†, e.g. analgesics
Ciclosporin
Cadmium, lead, titanium
Irradiation
Diseases
e.g. Diabetes mellitus
Sickle cell disease or trait
SLE/vasculitis
Sarcoidosis
Metabolic
e.g. Hyperuricaemia
Nephrocalcinosis
Hyperoxaluria
Infection
e.g. HIV
EBV*
Miscellaneous
Hypertension
Balkan nephropathy
Herbal nephropathy
Alport's syndrome

*Epstein–Barr virus
†Acute tubulointerstitial nephritis

Analgesic nephropathy

The chronic consumption of large amounts of analgesics (especially those containing phenacetin) and NSAIDs leads to chronic tubulointerstitial nephritis and papillary necrosis. Analgesic nephropathy is twice as common in women as in men and presents typically in middle age. Patients are often depressed or neurotic. Presentation may be with anaemia, chronic kidney disease, UTIs, haematuria or urinary tract obstruction (owing to sloughing of a renal papilla). Salt and water-wasting renal disease may occur. Chronic analgesic abuse also predisposes to the development of uroepithelial tumours. Diagnosis is usually made on clinical grounds combined with the non-pathognomonic appearance on imaging (such as ultrasonography or CT scan), which demonstrates smallish irregularly outlined kidneys.

The consumption of the above analgesics should be discouraged. If necessary, dihydrocodeine or paracetamol is a reasonable alternative. This may result in the arrest of the

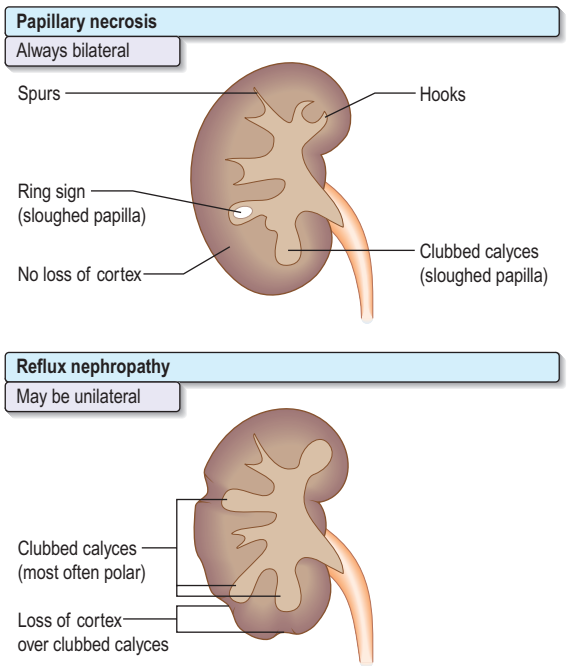


Fig. 11.35 **A comparison of radiological appearances of papillary necrosis and reflux nephropathy.**

disease and even in improvement in function. UTI, hypertension (if present) and saline depletion will require appropriate management. The development of flank pain or an unexpectedly rapid deterioration in renal function should prompt ultrasonography to screen for urinary tract obstruction due to a sloughed papilla.

Chinese herb nephropathy

Chinese herbal medicines have been increasingly used in the West, e.g. for slimming, and have caused a nephropathy. The renal histology is similar to Balkan nephropathy but the clinical course is very aggressive. The causal agent has been identified as aristolochic acid produced as a result of fungal contamination of the herbal medicine. It is characterized by relentless progression to end-stage renal failure. There is a high incidence of uroepithelial tumours.

Balkan nephropathy (BN)

This is a chronic TIN endemic in areas along the tributaries of the River Danube. Inhabitants of the low-lying plains, which are subjected to frequent flooding and where the water supply comes from shallow wells, are affected, whereas the disease does not occur in hillside villages where surface water provides the water supply. Its cause was essentially unknown. However, new research suggests that chronic dietary poisoning by aristolochic acid (AA) is responsible for BN and its associated urothelial cancer. AA-DNA adduct is found in tissue biopsies both from BN and associated urothelial cancers. This research suggests that AA is the environmental agent responsible for BN and its attendant transitional-cell cancer. The disease is insidious in onset, with mild proteinuria progressing to renal failure in 3 months to 10 years. There is no treatment.

FURTHER READING

Appel GB. The treatment of acute interstitial nephritis: More data at last. *Kidney International* 2008; **73**(8): 905–907.

Other forms of chronic tubulointerstitial nephritis

These are rare (see Table 11.10). Diagnosis of all forms depends on a history of drug ingestion or industrial exposure to nephrotoxins. In patients with unexplained renal impairment with normal-sized kidneys, renal biopsy must always be undertaken to exclude a treatable tubulointerstitial nephritis such as granulomatous TIN due to *renal sarcoidosis* (Fig. 11.36), which may be the first presentation of sarcoidosis (see p. 868). Renal sarcoidosis generally responds rapidly to steroids.

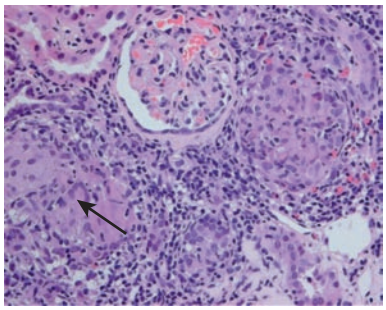


Fig. 11.36 **Granulomatous tubulointerstitial nephritis**. The granuloma consists of both giant cells (arrow) and epithelioid cells. These findings are characteristic of *renal sarcoid* but can be seen with any cause (drug or infection) of interstitial nephritis.

HYPERURICAEMIC NEPHROPATHY

(see p. 536)

Acute hyperuricaemic nephropathy is a well-recognized cause of acute kidney injury in patients with marked hyperuricaemia that is usually due to lymphoproliferative or myeloproliferative disorders. It may occur prior to treatment but most often follows commencement of treatment, when there is rapid lysis of malignant cells, release of large amounts of nucleoprotein and increased uric acid production. Renal failure is due to intrarenal and extrarenal obstruction caused by deposition of uric acid crystals in the collecting ducts, pelvis and ureters. The condition is manifest as oliguria or anuria with increasing uraemia. There may be flank pain or colic. Plasma urate levels are above 0.75 mmol/L and may be as high as 4.5 mmol/L. *Diagnosis* is based on the hyperuricaemia and the clinical setting. Ultrasound may demonstrate extrarenal obstruction due to stones, but a negative scan does not exclude this where there is coexistent intrarenal obstruction.

Allopurinol 100–200 mg three times daily for 5 days is given prior to and throughout treatment with radiotherapy or cytotoxic drugs. A high rate of urine flow must be maintained by oral or parenteral fluid and the urine kept alkaline by the administration of sodium bicarbonate 600 mg four times daily and acetazolamide 250 mg three times daily, since uric acid is more soluble in an alkaline than in an acid medium. *Rasburicase*, a recombinant urate oxidase, (p. 458) is occasionally used. In severely oliguric or anuric patients, dialysis is required to lower the plasma urate.

There is no convincing evidence for *chronic hyperuricaemia nephropathy*.

HYPERTENSION AND THE KIDNEY

Hypertension can be the cause or the result of renal disease. It is often difficult to differentiate between the two on clinical grounds. Routine tests as described on page 802 should be performed on all hypertensive patients, but renal imaging is usually unnecessary. A guide to how patients should be fully investigated is given on page 1022.

The *mechanisms* responsible for the normal regulation of arterial blood pressure and the development of essential primary hypertension are unclear (p. 798). One basic concept is that the long-term regulation of arterial pressure is closely linked to the ability of the kidneys to excrete sufficient salt to maintain normal sodium balance, extracellular fluid volume, and normal blood volume at normotensive arterial pressures. Several studies point to the kidney being involved in the genesis of essential primary hypertension. Cross-transplantation experiments suggest that hypertension travels with the kidney, in that hypertension will develop in a normotensive recipient of a kidney genetically programmed for hypertension. Likewise patients with end-stage renal failure due to hypertension become normotensive after receiving a renal allograft from normotensive donors, provided the new kidney works.

One of many renal factors involved in the genesis of hypertension is the total number of nephrons in the kidney. There is circumstantial as well direct evidence to suggest that patients with hypertension and normal renal function have a significantly reduced number of nephrons in each

kidney alongside enlargement of the remaining glomeruli due to glomerular hyperfiltration. Moreover, certain races (blacks, Hispanics) with predilection for hypertension have increased glomerular volume, a surrogate marker for a reduced number of nephrons.

Whether the reduced number of nephrons is caused by genetic or environmental factors is unclear. Changes in the intrauterine environment may lead to retarded renal growth before birth and low birth weight and hypertension during adult life. In humans, an association has been found between low birth weight and reduced renal volume, possibly indicating reduced numbers of nephrons.

Essential hypertension

Pathophysiology

In benign essential hypertension, arteriosclerosis of major renal arteries and changes in the intrarenal vasculature (nephrosclerosis) occur as follows:

- In small vessels and arterioles, intimal thickening with reduplication of the internal elastic lamina occurs and the vessel wall becomes hyalinized.
- In large vessels, concentric reduplication of the internal elastic lamina and endothelial proliferation produce an 'onion skin' appearance.
- Reduction in size of both kidneys occurs; this may be asymmetrical if one major renal artery is more affected than the other.
- The proportion of sclerotic glomeruli is increased compared with age-matched controls.

Deterioration in excretory function accompanies these changes, but severe renal failure is unusual in whites (1 in 10 000). In black Africans, by contrast, hypertension much more often results in the development of renal failure with a fourfold higher incidence of end-stage renal failure in blacks compared to whites. This racial difference in incidence of hypertensive renal disease may be due to overestimation of diagnosis on clinical grounds, poor compliance with medication, higher incidence of hypertension, which is usually a salt-sensitive type, and reduced number of nephrons.

In *accelerated, or malignant-phase hypertension*:

- Arteriolar fibrinoid necrosis occurs, probably as a result of plasma entering the media of the vessel through splits in the intima. It is prominent in afferent glomerular arterioles.
- Fibrin deposition within small vessels is often associated with thrombocytopenia and red-cell fragmentation seen in the peripheral blood film (microangiopathic haemolytic anaemia).

Microscopic haematuria, proteinuria, usually of modest degree (1–3 g daily), and progressive uraemia occur. If untreated, fewer than 10% of patients survive 2 years.

Management

The management of benign essential and malignant hypertension is described on page 802.

If treatment is begun before renal impairment has occurred, the prognosis for renal function is good. Stabilization or improvement in renal function with healing of intrarenal arteriolar lesions and resolution of microangiopathic haemolysis occur with effective treatment of malignant-phase hypertension. Lifelong follow-up of the patient is mandatory. In blacks with hypertensive nephrosclerosis and renal impairment, a blood pressure target of <140/85 mmHg should be achieved. ACE inhibitors appear to be more effective than beta-blockers, and, if possible, calcium-channel blockers of the dihydropyridine group (e.g. amlodipine) should be avoided as they have a negative intropic effect and can cause heart failure.

Renal hypertension

Hypertension commonly complicates bilateral renal disease such as chronic glomerulonephritis, bilateral reflux nephropathy, polycystic disease and analgesic nephropathy. Two main mechanisms are responsible:

- activation of the renin–angiotensin–aldosterone system
- retention of salt and water owing to impairment in excretory function, leading to an increase in blood volume and hence blood pressure.

The second of these assumes greater significance as renal function deteriorates.

Hypertension occurs earlier, is more common and tends to be more severe in patients with renal cortical disorders, such as glomerulonephritis, than in those with disorders affecting primarily the renal interstitium, such as reflux or analgesic nephropathy.

Management is described on page 802. Meticulous control of the blood pressure is necessary to prevent further deterioration of renal function secondary to vascular changes produced by the hypertension itself. There is good evidence that ACE-inhibitor drug treatment confers an additional renoprotective effect for a given degree of blood pressure control than other hypotensive drugs.

Renovascular disease

Mechanism of hypertension

Renal ischaemia results in a reduction in the pressure in afferent glomerular arterioles. This leads to an increase in the production and release of renin from the juxtaglomerular apparatus (see p. 572) with a consequent increase in angiotensin II, a very potent vasoconstrictor. Angiotensin II also causes hypertension by upregulating NADPH oxidase enzyme with excessive superoxide generation. Superoxide chelates nitric oxide (a potent vasodilator) resulting in reduced vasodilator activity and also hypertension.

Physiological changes in renal artery stenosis

In renal artery stenosis, renal perfusion pressure is reduced and nephron transit time is prolonged on the side of the stenosis; salt and water reabsorption is therefore increased. As a result, urine from the ischaemic kidney is more concentrated and has a lower sodium concentration than urine from the contralateral kidney. Creatinine clearance is decreased on the ischaemic side.

Pathology

Narrowing of the renal arteries (renal artery stenosis) is caused by one of two pathological entities: *fibromuscular disease* or *atherosclerotic renovascular disease (ARVD).*

Fibromuscular disease of the renal arteries

Fibromuscular disease accounts for 20–40% of renal vascular disease and encompasses *four distinct types*: **(1)** medial fibroplasia (65–85%); **(2)** perimedial fibroplasia (10–15%); **(3)** intimal fibroplasia (5–10%); and **(4)** medial hyperplasia (5%). Medial fibroplasia usually follows a benign course and never follows a progressive course after the age of 40 years. The other two types of fibroplasia follow a progressive course and may lead to total occlusion.

Medial hyperplasia is a distinct but rare entity, which accounts for only 1% of renovascular disease. It commonly affects young females, who exhibit elevated blood pressures but with well-preserved renal function.

MR angiography (gadolinium enhanced) reveals a characteristic string of beads appearance in fibroplasia. Angioplasty (occasionally stent insertion) is usually successful in improving or even curing hypertension in affected individuals, and overall prognosis is very good.

Atherosclerotic renovascular disease (ARVD)

This is a common cause of hypertension and chronic kidney disease due to ischaemic nephropathy. Its incidence increases with age, rising from 5% under 60 years to 16% in those over 60 years old. In most patients the atherosclerotic lesion is ostial (within 1 cm of the origin of the renal artery) and usually associated with symptomatic atherosclerotic vascular disease elsewhere. Patients with peripheral vascular disease (39%), coronary artery disease (10–29%), congestive cardiac failure (34%) and aortic aneurysm (38%) are at high risk of developing significant renal artery stenosis.

Many patients are asymptomatic and are discovered incidentally during investigation for other conditions. Aortography experience from the USA shows 11% of asymptomatic patients have significant unilateral stenosis and 4% have bilateral disease. The renal consequences of ARVD are functional, such as hypertension (present in 50%), sodium retention (ankle and flash pulmonary oedema), proteinuria (usually sub-nephrotic range) and decreased GFR. The morphological features of the affected kidneys include vascular sclerosis, tubular atrophy, interstitial fibrosis with inflammatory cellular infiltrate, atubular glomeruli, cholesterol emboli, and secondary focal segmental glomerulosclerosis (FSGS) changes. Baseline renal function is related to the extent of renal parenchymal injury rather than to the degree of stenosis, as is the response (improvement in hypertension and renal function) to revascularization.

Renovascular disease should be looked for in patients with hypertension and/or chronic kidney disease; patients with abdominal audible bruits, as well as bruits over carotid arteries suggestive of generalized arterial disease; doppler ultrasonography showing >1.5 cm renal asymmetry; recurrent flash pulmonary oedema without cardiopulmonary disease; progressive chronic renal failure in patients with evidence of generalized atherosclerosis.

Treatment

The *aim of treatment* is to correct hypertension and improve renal perfusion and excretory function. Renal artery stenosis can progress to occlusion, particularly in patients with stenosis >75% as shown by serial angiography, necessitating revascularization in ARVD.

- The *options in renal artery stenosis* include transluminal angioplasty to dilate the stenotic region, insertion of stents across the stenosis (sometimes the only endoscopic option when the stenosis occurs close to the origin of the renal artery from the aorta, rendering angioplasty technically difficult or impossible), reconstructive vascular surgery and nephrectomy.
- *Indications for revascularization*. Vessels with stenosis >75% and recurrent flash pulmonary oedema, drug resistant severe hypertension, ARVD affecting solitary functioning kidney, patients with cardiac failure needing ACE inhibitors, unexplained progressive renal failure and dialysis-dependent renal failure. With good selection of patients, hypertension is cured or improved by intervention in more than 50%. Occasional dramatic improvements in renal function ensue but results are generally disappointing.
- *Medication*. All patients with ARVD should be treated with a combination of aspirin, statins and optimal control of blood pressure as prophylaxis against progression of atherosclerosis.

Prognosis. Mortality is high because of other associated comorbidities, and ARVD patients have generalized endothelial dysfunction. ARVD patients with end-stage renal failure have higher rates than those with good renal function. Five-year survival is only 18% in patients with end-stage renal failure due to ARVD.

Screening for renovascular disease

Radionuclide studies (see p. 581). These can demonstrate decreased renal perfusion on the affected side. In unilateral renal artery stenosis, a disproportionate fall in uptake of isotope on the affected side following administration of captopril or aspirin is suggestive of the presence of significant renal artery stenosis. A completely normal result renders the diagnosis unlikely.

Doppler ultrasound. This method is very sensitive but highly operator-dependent and time-consuming. It generates data about intrarenal vascular resistance, which can be valuable in predicting the success of revascularization procedures. A resistive index of ≥80 is a predictor of poor response following intervention.

Magnetic resonance angiography. MRA can be used to visualize the renal arteries and there is a good – though not perfect – correlation between MRA findings and those of renal arteriography.

Helical ('spiral') CT scanning. This permits non-invasive imaging of the renal arteries. It is much less expensive than MRA but does expose the patient to ionizing radiation and to contrast injection and is less reliable than MRA.

Renal arteriography (see p. 581) remains the 'gold standard' investigation in the diagnosis of renal arterial disease.

Systemic sclerosis (see also p. 545)

Interlobular renal arteries are affected with intimal thickening, and fibrinoid changes occur in afferent glomerular arterioles similar to renal changes in malignant hypertension. Glomerular changes are non-specific. The pathogenesis is unknown and neither steroid nor immunosuppressive therapy is of

value. Serum ANCAs are not present. Treatment with ACE-inhibitor drugs is of immense benefit, reducing proteinuria and often halting or even partially reversing decline in renal function. 'Scleroderma renal crisis' is a term applied to rapid loss of renal function owing to rapid progression of renal microvascular disease in this condition. Early aggressive therapy with ACE inhibitors/blockers and continuous intravenous infusion of prostacyclin may help.

Prognosis. Fifty per cent will require renal dialysis which may be just temporary. Function can improve over 2 years. ACE inhibitors should be continued. Renal transplantation is sometimes required.

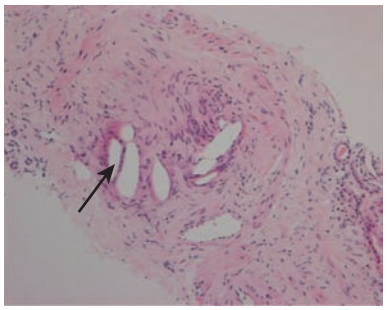


Fig 11.37 **Renal cholesterol emboli showing the characteristic cholesterol biconvex clefts.**

OTHER VASCULAR DISORDERS OF THE KIDNEY

Renal artery occlusion

This occurs from thrombosis in situ, usually in a severely damaged atherosclerotic vessel, or more commonly from embolization, e.g. in atrial fibrillation. Both lead to renal infarction, resulting in a wide spectrum of clinical manifestations depending on the size of the artery involved. Occlusion of a small branch artery may produce no effect, but occlusion of larger vessels results in dull flank pain and varying degrees of renal failure.

Intra-arterial thrombolytic therapy has been tried with mixed results.

Cholesterol embolization or atheroembolic renal disease (AERD)

Showers of cholesterol-rich atheromatous material from ulcerated plaques reach the kidney from the aorta and/or renal arteries, particularly after catheterization of the abdominal aorta or attempts at renal artery angioplasty. Anticoagulants and thrombolytic agents also precipitate cholesterol embolization. Renal failure from cholesterol emboli may be acute or slowly progressive. Clinical features include fever, eosinophilia, back and abdominal pain, and evidence of embolization elsewhere, for example to the retina or digits. The diagnosis can be confirmed by renal biopsy (Fig. 11.37). It is more common in males, the elderly (>70 years) and patients with cardiovascular disease. Over 80% have abnormal renal function at baseline. AERD occurs spontaneously in 25% of the cases. Two-year mortality is 30% and a similar percentage of patients develop chronic kidney disease. Baseline co-morbidities, i.e. reduced renal function, presence of diabetes, history of heart failure, acute/subacute presentation, and gastrointestinal tract involvement, are significant predictors of event occurrence. The risk of dialysis and death is 50% lower among those receiving statins.

Renal vein thrombosis

This is usually of insidious onset, occurring in patients with the nephrotic syndrome, with a renal cell carcinoma, and in thrombophilia (p. 437) with an increased risk of venous thrombosis. Anticoagulation is indicated.

RENAL CALCULI AND NEPHROCALCINOSIS

Renal and vesical calculi

Renal stones are very common world-wide with a lifetime risk of about 10%. Prevalence of stone disease is much higher in the Middle East. Most stones occur in the upper urinary tract.

Most stones are composed of calcium oxalate and phosphate; these are more common in men (Table 11.11). Mixed infective stones, which account for about 15% of all calculi, are twice as common in women as in men. The overall male to female ratio of stone disease is 2:1.

Stone disease is frequently a recurrent problem. More than 50% of patients with a history of nephrolithiasis will develop a recurrence within 10 years. The risk of recurrence increases if a metabolic or other abnormality predisposing to stone formation is present and is not modified by treatment.

Aetiology

Inhibitors of crystal formation are present in normal urine preventing the formation of stones, as the concentrations of stone-forming substances in many cases exceed their maximum solubility in water. Many stone-formers have no detectable metabolic defect, although microscopy of warm, freshly passed urine reveals both more and larger calcium oxalate crystals than are found in normal subjects. Factors predisposing to stone formation in these so-called 'idiopathic stone-formers' are:

Table 11.11 Type and frequency of renal stones in the UK

Type of renal stone	Percentage of stones
Calcium oxalate usually with calcium phosphate	65
Calcium phosphate alone	15
Magnesium ammonium phosphate (struvite)	10–15
Uric acid	3–5
Cystine	1–2

FURTHER READING

Balk E, Raman G, Chung M et al. Effectiveness of management strategies for renal artery stenosis: a systematic review. *Annals of Internal Medicine* 2006; **145**(12): 901–912.

Ikee R, Kobyashi S, Hemmi N et al. Correlation between the resistive index by Doppler ultrasound and kidney function and histology. *American Journal of Kidney Diseases* 2005; **46**(4): 603–609.

Tullus K, Brennan E, Hamilton G. Renovascular hypertension in children. *Lancet* 2008; **371**: 1453–1463.

Williams GJ, Macaskill P, Chan SF et al. Comparative accuracy of renal duplex sonographic parameters in the diagnosis of renal artery stenosis: paired and unpaired analysis. *American Journal of Roentgenology* 2007; **188**(3): 798–811.

- Chemical composition of urine that favours stone crystallization.
- Production of a concentrated urine as a consequence of dehydration associated with life in a hot climate or work in a hot environment.
- Impairment of inhibitors that prevent crystallization in normal urine. Postulated inhibitors include inorganic magnesium, pyrophosphate and citrate. Organic inhibitors include glycosaminoglycans and nephrocalcin (an acidic protein of tubular origin). Tamm–Horsfall protein may have a dual role in both inhibiting and promoting stone formation.

Recognized causes of stone formation are listed in Table 11.12.

Hypercalcaemia

If the GFR is normal, hypercalcaemia almost invariably leads to hypercalciuria. The common causes of hypercalcaemia leading to stone formation are:

- primary hyperparathyroidism
- vitamin D ingestion
- sarcoidosis.

Of these, primary hyperparathyroidism (see p. 1019) is the most common cause of stones.

Hypercalciuria

This is by far the most common metabolic abnormality detected in calcium stone-formers.

Approximately 8% of men excrete in excess of 7.5 mmol of calcium in 24 hours. Calcium stone formation is more common in this group, but as the majority of even these individuals do not form stones the definition of 'pathological' hypercalciuria is arbitrary. A reasonable definition is 24-hour calcium excretion of more than 7.5 mmol in male stone-formers and more than 6.25 mmol in female stone-formers.

The kidney is the major site for plasma calcium regulation. Approximately 90% of the ionized calcium filtered by the kidney is reabsorbed. Renal tubular reabsorption is controlled largely by parathyroid hormone (PTH).

Approximately 65% of the filtered calcium is absorbed in the proximal convoluted tubule, 20% by the thick ascending limb of the loop of Henle, and 15% by the distal convoluted tubule and collecting ducts.

Causes of hypercalciuria are:

Table 11.12 Causes of urinary tract stone formation

Dehydration
Hypercalcaemia
Hypercalciuria
Hyperoxaluria
Hyperuricaemia and hyperuricosuria
Infection
Cystinuria
Renal tubular acidosis
Primary renal disease (polycystic kidneys, medullary sponge kidneys)
Drugs (see text)

- hypercalcaemia
- an excessive dietary intake of calcium
- excessive resorption of calcium from the skeleton, such as occurs with prolonged immobilization or weightlessness
- idiopathic hypercalciuria.

Idiopathic hypercalciuria is a common risk factor for the formation of stones, and uncontrolled hypercalciuria is a cause of recurrences. The majority of patients with idiopathic hypercalciuria have increased absorption of calcium from the gut. Moreover, studies have shown that animal protein and salt also have a considerable influence on calcium excretion.

Hyperoxaluria

There are two inborn errors of glyoxalate metabolism that cause increased endogenous oxalate biosynthesis, and are inherited in an autosomal recessive manner. In both types, calcium oxalate stone formation occurs.

The prognosis is poor owing to widespread calcium oxalate crystal deposition in the kidneys. Renal failure typically develops in the late teens or early twenties. Successful liver transplantation has been shown to cure the metabolic defect.

Much more common causes of mild hyperoxaluria are:

- excess ingestion of foodstuffs high in oxalate, such as spinach, rhubarb and tea
- dietary calcium restriction, with compensatory increased absorption of oxalate
- gastrointestinal disease (e.g. Crohn's), usually with an intestinal resection, associated with increased absorption of oxalate from the colon.

Dehydration secondary to fluid loss from the gut also plays a part in stone formation.

Hyperuricaemia and hyperuricosuria

Uric acid stones account for 3–5% of all stones in the UK, but in Israel the proportion is as high as 40%. Uric acid is the end-point of purine metabolism. Hyperuricaemia (see p. 537) can occur as a primary defect in idiopathic gout, and as a secondary consequence of increased cell turnover, for example in myeloproliferative disorders. Increased uric acid excretion occurs in these conditions, and stones will develop in some patients. Some uric acid stone-formers have hyperuricosuria (>4 mmol per 24 hours on a low-purine diet) without hyperuricaemia.

Dehydration alone may also cause uric acid stones to form. Patients with ileostomies are at particular risk both from dehydration and from the fact that loss of bicarbonate from gastrointestinal secretions results in the production of an acid urine (uric acid is more soluble in an alkaline than in an acid medium).

Some patients with calcium stones also have hyperuricaemia and/or hyperuricosuria; it is believed the calcium salts precipitate upon an initial nidus of uric acid in such patients.

Urinary tract infection

Mixed infective stones are composed of magnesium ammonium phosphate together with variable amounts of calcium. Such struvite stones are often large, forming a cast of the

collecting system (staghorn calculus). These stones are usually due to UTI with organisms such as *Proteus mirabilis* that hydrolyse urea, with formation of the strong base ammonium hydroxide. The availability of ammonium ions and the alkalinity of the urine favour stone formation. An increased production of mucoprotein from infection also creates an organic matrix on which stone formation can occur.

Cystinuria (see also p. 1070)

Cystinuria results in the formation of cystine stones. About 1–2% of all stones are composed of cystine.

Primary renal diseases

- *Polycystic renal disease* (see p. 642) shows a high prevalence of stone disease.
- *Medullary sponge kidney* is also associated with stones. There is dilatation of the collecting ducts with associated stasis and calcification (Fig. 11.38). Approximately 20% of these patients have hypercalciuria and a similar proportion have a renal tubular acidification defect.
- *Renal tubular acidoses,* both inherited and acquired, are associated with nephrocalcinosis and stone formation, owing, in part, to the production of a persistently alkaline urine and reduced urinary citrate excretion.

Drugs

Some drugs promote calcium stone formation (e.g. loop diuretics, antacids, glucocorticoids, theophylline, vitamins D and C, acetazolamide); some promote uric acid stones (e.g. thiazides, salicylates,); and some precipitate into stones (e.g. indinavir, triamterene, sulphadiazine).

Aetiology of bladder stones

Bladder stones are endemic in some developing countries but the incidence is declining. The cause of this is unknown but dietary factors probably play a role. Bladder stones may be the result of:

- bladder outflow obstruction (e.g. urethral stricture, neuropathic bladder, prostatic obstruction)
- the presence of a foreign body (e.g. catheters, non-absorbable sutures).

Significant bacteriuria is usually found in patients with bladder stones. Some stones found in the bladder have been passed down from the upper urinary tract.

Pathology

Stones may be single or multiple and vary enormously in size from minute, sand-like particles to staghorn calculi or large stone concretions in the bladder. They may be located within the renal parenchyma or within the collecting system. Pressure necrosis from a large calculus can cause direct damage to the renal parenchyma, and stones regularly cause obstruction, leading to hydronephrosis. They may ulcerate through the wall of the collecting system, including the ureter. A combination of obstruction and infection accelerates damage to the kidney.

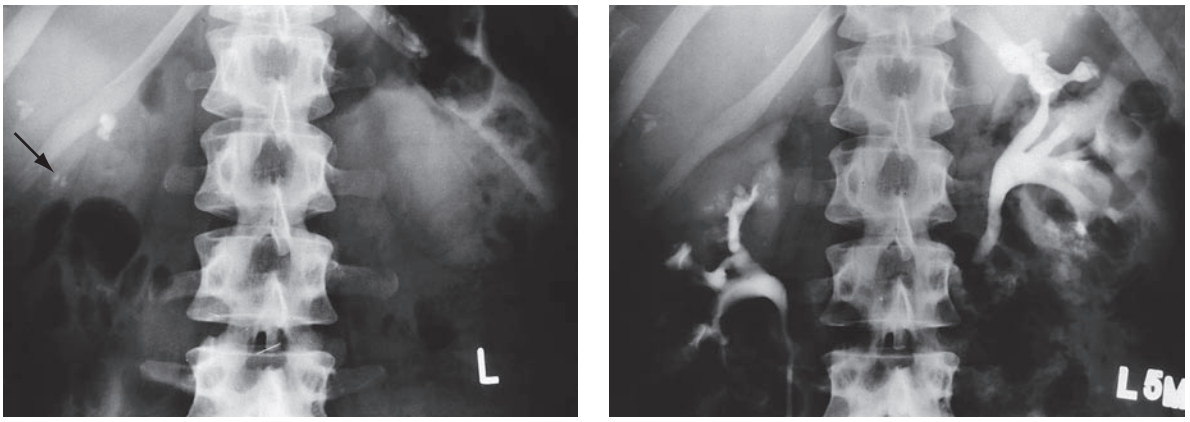
Clinical features

Most people with urinary tract calculi are asymptomatic. Pain is the most common symptom and may be sharp or dull, constant, intermittent or colicky (Table 11.13).

When urinary tract obstruction is present, measures that increase urine volume, such as copious fluid intake or diuretics, including alcohol, make the pain worse. Physical exertion may cause mobile calculi to move, precipitating pain and, occasionally, haematuria. Ureteric colic occurs when a stone enters the ureter and either obstructs it or causes spasm during its passage down the ureter. This is one of the most severe pains known. Radiation from the flank to the iliac fossa and testis or labium in the distribution of the first lumbar nerve root is common. Pallor, sweating and vomiting often occur and the patient is restless, tending to assume a variety of positions in an unsuccessful attempt to obtain relief

Table 11.13 Clinical features of urinary tract stones

Asymptomatic
Pain: renal colic
Haematuria
Urinary tract infection
Urinary tract obstruction



(a) (b)

Fig. 11.38 **Medullary sponge kidney. (a)** Plain film showing 'spotty' calcification in the renal areas (arrow). **(b)** After injection of contrast, the calcification is shown to be small calculi in the papillary zones.

from the pain. Haematuria often occurs. Untreated, the pain of ureteric colic typically subsides after a few hours.

When urinary tract obstruction and infection are present, the features of acute pyelonephritis or of a Gram-negative septicaemia may dominate the clinical picture.

Vesical calculi associated with bladder bacteriuria present with frequency, dysuria and haematuria; severe introital or perineal pain may occur if trigonitis is present. A calculus at the bladder neck or an obstruction in the urethra may cause bladder outflow obstruction, resulting in anuria and painful bladder distension.

A history of possible aetiological factors should be obtained, including:

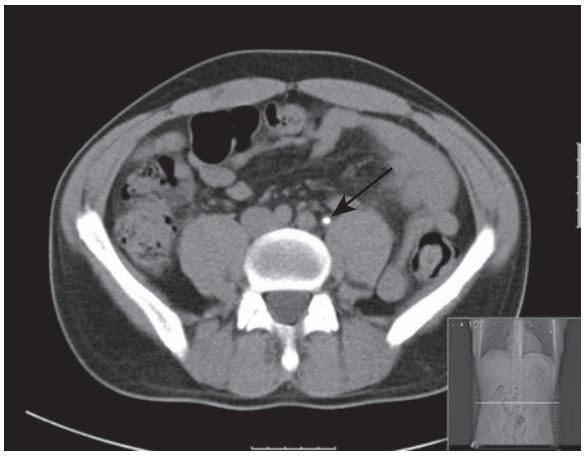
- occupation and residence in hot countries likely to be associated with dehydration
- a history of vitamin D consumption
- gouty arthritis.

Calcified papillae may mimic ordinary calculi, so that causes of papillary necrosis such as analgesic abuse should be considered.

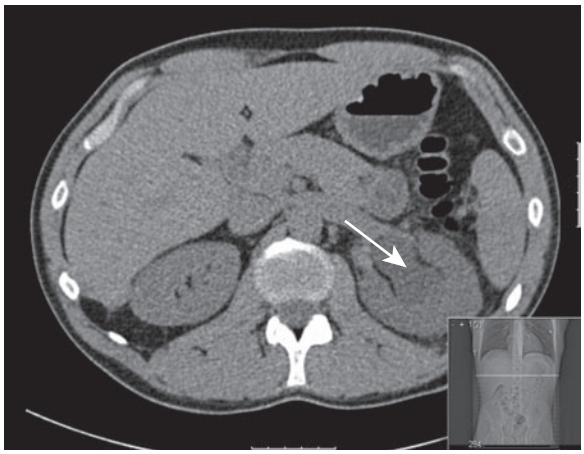
Physical examination should include a search for corneal or conjunctival calcification, gouty tophi and arthritis and features of sarcoidosis.

Investigations

- A *mid-stream specimen of urine* for culture
- Serum urea, electrolyte, creatinine and calcium levels.
- *Plain abdominal X-ray*.
- Unenhanced helical (spiral) CT is the best diagnostic test available. Ureteric stones can be missed by ultrasound.
- CT-KUB (CT of kidney, ureter and bladder) is carried out during the episode of pain; a normal CT excludes the diagnosis of pain due to calculous disease. The CT-KUB appearances in a patient with acute left ureteric obstruction are shown in Figure 11.39 a and b.



(a) Left ureteric calculus.



(b) A dilated renal pelvis (arrow) proximal to the ureteric stone in (a).

Fig. 11.39 **CT-KUB in ureteric stone obstruction.**

Pure uric acid stones are radiolucent and show as a filling defect after injection of contrast medium if excretion urography is performed. Such stones are readily seen on CT scanning (Fig. 11.40). Mixed infective stones in which organic matrix predominates are barely radiopaque.

The urine of the patient should be passed through a sieve to trap any calculi for chemical analysis.

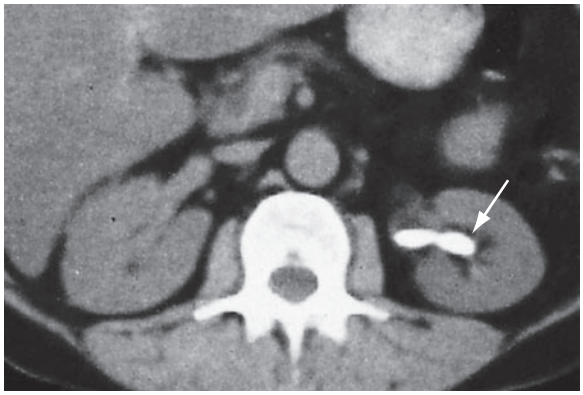


Fig. 11.40 **CT scan, showing a uric acid stone, which appears as a bright lesion in the left kidney (arrow).**

Management

Adequate analgesia should be given. An NSAID, e.g. diclofenac 75 mg by i.v. infusion, compares favourably with pethidine and does not cause nausea. Stones less than 0.5 cm diameter usually pass spontaneously. Stones greater than 1 cm diameter usually require urological or radiological intervention. Extracorporeal shock wave lithotripsy (ESWL) will fragment most stones, which then pass spontaneously. Ureteroscopy with a Yag laser can be used for larger stones. Percutaneous nephrolithotomy is also used. Open surgery is rarely needed.

Investigating the cause of stone formation

In an elderly patient who has had a single episode with one stone, only limited investigation is required. Younger patients and those with recurrent stone formation require detailed investigation.

- **Renal imaging** is necessary to define the presence of a primary renal disease predisposing to stone formation.
- **Significant bacteriuria** may indicate mixed infective stone formation, but relapsing bacteriuria may be a consequence of stone formation rather than the original cause.

- **Chemical analysis** of any stone passed is of great value and all that is required in the diagnosis of cystinuria or uric acid stone formation.
- **Serum calcium concentration** should be estimated and corrected for serum albumin concentration (see p. 593). Hypercalcaemia, if present, should be investigated further (see p. 1021).
- **Serum urate concentration** is often, but not invariably, elevated in uric acid stone-formers.
- A **screening test for cystinuria** should be carried out by adding sodium nitroprusside to a random unacidified urine sample; a purple colour indicates that cystinuria may be present. Urine chromatography is required to define the diagnosis precisely.
- **Urinary calcium, oxalate and uric acid output** should be measured in two consecutive carefully collected 24-hour urine samples. After withdrawing aliquots for estimation of uric acid, it is necessary to add acid to the urine in order to prevent crystallization of calcium salts upon the walls of the collection vessel, which would give falsely low results for urinary calcium and oxalate.
- **Plasma bicarbonate** is low in renal tubular acidosis. The finding of a urine pH that does not fall below 5.5 in the face of metabolic acidosis is diagnostic of this condition (see p. 675).

Prophylaxis

The age of the patient and the severity of the problem affect both the need for and the type of prophylaxis.

Idiopathic stone-formers

Where no metabolic abnormality is present, the mainstay of prevention is maintenance of a high intake of fluid throughout the day and night. The aim should be to ensure a daily urine volume of 2–2.5 L, which requires a fluid intake in excess of this, substantially so in the case of those who live in hot countries or work in a hot environment.

Idiopathic hypercalciuria

Severe dietary calcium restriction is inappropriate (see p. 610) and patients should be encouraged to consume a normal-calcium (30 mmol/day) diet. Dietary calcium restriction results in hyperabsorption of oxalate, and so foods containing large amounts of oxalate should also be limited. A high fluid intake should be advised as for idiopathic stone-formers. Patients who live in a hard-water area may benefit from drinking softened water.

If hypercalciuria persists and stone formation continues, a thiazide is used (e.g. bendroflumethiazide 2.5 or 5 mg each morning). Thiazides reduce urinary calcium excretion by a direct effect on the renal tubule. They may precipitate diabetes mellitus or gout and worsen hypercholesterolaemia. Reduction of animal proteins to 50 g/day and sodium intake to 50 mmol/day is also advisable.

Mixed infective stones

Recurrent stones should be prevented by maintenance of a high fluid intake and meticulous control of bacteriuria. This will require long-term follow-up and often the use of long-term low-dose prophylactic antibacterial agents.

Uric acid stones

Dietary measures are probably of little value and are difficult to implement. Effective prevention can be achieved by the long-term use of allopurinol to maintain the serum urate and urinary uric acid excretion in the normal range. A high fluid intake should also be maintained. Uric acid is more soluble at alkaline pH, and long-term sodium bicarbonate supplementation to maintain an alkaline urine is an alternative approach in those few patients unable to take allopurinol. However, alkalinization of the urine facilitates precipitation of calcium oxalate and phosphate.

Cystine stones

These can be prevented and indeed will dissolve slowly with a high fluid intake. Five litres of water is drunk each 24 hours, and the patient must wake twice during the night to ingest 500 mL or more of water. Many patients cannot tolerate this regimen. Alkalinization to a pH of 7 requires high doses of potassium citrate or bicarbonate. An alternative option is the long-term use of the chelating agent penicillamine; this causes cystine to be converted to the more soluble penicillamine–cysteine complex. Side-effects include drug rashes, blood dyscrasias and immune complex-mediated glomerulonephritis. However, it is especially effective in promoting dissolution of cystine stones already present. Other cystine-binding drugs include tiopronin.

Mild hyperoxaluria with calcium oxalate stones

A high fluid intake and dietary oxalate restriction are required. Dietary advice as in hypercalciuria is also advisable.

Nephrocalcinosis

The term ‘nephrocalcinosis’ means diffuse renal parenchymal calcification that is detectable radiologically (Fig. 11.41). The condition is typically painless. Hypertension and renal impairment commonly occur. The main causes of nephrocalcinosis are listed in Table 11.14.

Dystrophic calcification occurs following renal cortical necrosis. In *hypercalcaemia and hyperoxaluria*, deposition of calcium oxalate results from the high concentration of calcium and oxalate within the kidney.

In *renal tubular acidosis* (see p. 676) failure of urinary acidification and a reduction in urinary citrate excretion both favour calcium phosphate and oxalate precipitation, since

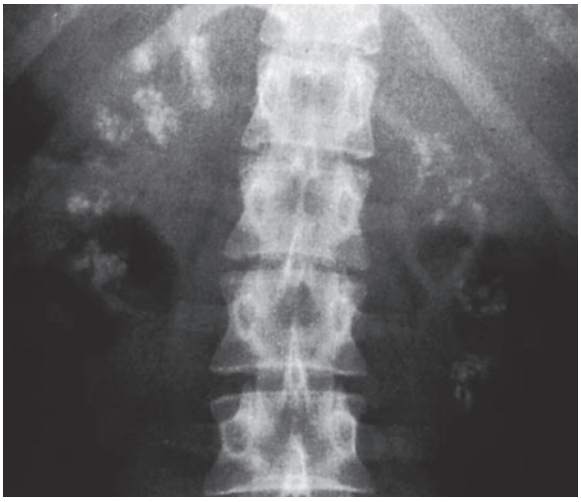


Fig. 11.41 **X-ray of nephrocalcinosis.**

FURTHER READING

Moe OW. Kidney stones: pathophysiology and medical management. *Lancet* 2006; **367**: 333–344.

Thongboonkerd V. Proteomics and kidney stone disease. *Contributions to Nephrology* 2008; **160**: 142–158.

Table 11.14 Causes of nephrocalcinosis

Mainly cortical (rare)

Renal cortical necrosis (tram-line calcification)

Mainly medullary

Hypercalcaemia (primary hyperparathyroidism, hypervitaminosis D, sarcoidosis)

Renal tubular acidosis (inherited and acquired)

Primary hyperoxaluria

Medullary sponge kidney

Tuberculosis

precipitation occurs more readily in an alkaline medium and the calcium-chelating action of urinary citrate is reduced.

Treatment and prevention of nephrocalcinosis consists of treatment of the cause.

URINARY TRACT OBSTRUCTION

The urinary tract can be obstructed at any point between the kidney and the urethral meatus. This results in dilatation of the tract above the obstruction. Dilatation of the renal pelvis is known as hydronephrosis.

Aetiology

Obstructing lesions may lie within the lumen, or in the wall of the urinary tract, or outside the wall, causing obstruction by external pressure. The major causes of obstruction are shown in Table 11.15. Overall, the frequency is the same in men and women. However, in the elderly, urinary tract obstruction is more common in men owing to the frequency of bladder outflow obstruction from prostatic disease.

Table 11.15 Causes of urinary tract obstruction

Within the lumen

Calculus

Blood clot

Sloughed papilla (diabetes; analgesia abuse; sickle cell disease or trait)

Tumour of renal pelvis or ureter

Bladder tumour

Within the wall

Pelviureteric neuromuscular dysfunction (congenital, 10% bilateral)

Ureteric stricture (tuberculosis, especially after treatment; calculus; after surgery)

Ureterovesical stricture (congenital; ureterocele; calculus; schistosomiasis)

Congenital megaureter

Congenital bladder neck obstruction

Neuropathic bladder

Urethral stricture (calculus; gonococcal; after instrumentation)

Congenital urethral valve

Pin-hole meatus

Pressure from outside

Pelviureteric compression (bands; aberrant vessels)

Tumours (e.g. retroperitoneal tumour or glands; carcinoma of colon; tumours in pelvis, e.g. carcinoma of cervix)

Diverticulitis

Aortic aneurysm

Retroperitoneal fibrosis (periaortitis)

Accidental ligation of ureter

Retrocaval ureter (right-sided obstruction)

Prostatic obstruction

Phimosis

Pathophysiology

Obstruction with continuing urine formation results in:

- progressive rise in intraluminal pressure
- dilatation proximal to the site of obstruction
- compression and thinning of the renal parenchyma, eventually reducing it to a thin rim and resulting in a decrease in the size of the kidney.

Acute obstruction is followed by transient renal arterial vasodilatation succeeded by vasoconstriction, probably mediated mainly by angiotensin II and thromboxane A_2. Ischaemic interstitial damage mediated by free oxygen radicals and inflammatory cytokines compounds the damage induced by compression of the renal substance.

Clinical features

Symptoms of upper tract obstruction

Loin pain occurs which can be dull or sharp, constant or intermittent. It is often provoked by measures that increase urine volume and hence distension of the collecting system, such as a high fluid intake or diuretics, including alcohol. Complete anuria is strongly suggestive of complete bilateral obstruction or complete obstruction of a single kidney.

Conversely, polyuria may occur in partial obstruction owing to impairment of renal tubular concentrating capacity. Intermittent anuria and polyuria indicates intermittent complete obstruction.

Infection complicating the obstruction may give rise to malaise, fever and septicaemia.

Symptoms of bladder outflow obstruction

Symptoms may be minimal. Hesitancy, narrowing and diminished force of the urinary stream, terminal dribbling and a sense of incomplete bladder emptying are typical features. The frequent passage of small volumes of urine occurs if a large volume of residual urine remains in the bladder after urination. Incontinence of such small volumes of urine is known as 'overflow incontinence' or 'retention with overflow'.

Infection commonly occurs, causing increased frequency, urgency, urge incontinence, dysuria and the passage of cloudy smelly urine. It may precipitate acute retention.

Signs

Loin tenderness may be present. An enlarged hydronephrotic kidney is often palpable. In acute or chronic retention the enlarged bladder can be felt or percussed. Examination of the genitalia, rectum and vagina is essential, since prostatic obstruction and pelvic malignancy are common causes of urinary tract obstruction. However, the apparent size of the

prostate on digital examination is a poor guide to the presence of prostatic obstruction.

Investigations

- **Routine blood and biochemical investigations** show a raised serum urea or creatinine, hyperkalaemia, anaemia of chronic disease or blood in the urine.
- **Plain abdominal X-ray** may detect radiolucent stones/calcification but can miss stones lying over the bone.
- **Helical (spiral) CT scanning** has a high sensitivity and can visualize uric acid (radiolucent) stones as small as 1 mm, as well as details of the obstruction.
- **Ultrasonography** (see p. 579) can rule out upper urinary tract dilatation. Ultrasound cannot distinguish a baggy, low-pressure unobstructed system from a tense, high-pressure obstructed one, so that false-positive scans are seen. Stones in the ureter can be missed.
- **Excretion urography**. In recent unilateral obstruction, the affected kidney is enlarged and smooth in outline. The nephrogram is delayed on the affected side, owing to a reduction in the GFR. With time, the nephrogram on the affected side becomes denser than normal, owing to the prolonged nephron transit time, and the site of obstruction with proximal dilatation is seen (Fig. 11.42). A full-length film should be taken after an attempt at bladder emptying by the patient. Complete emptying indicates either that no obstruction to bladder outflow exists or that intravesical pressure can be raised sufficiently to overcome it. Vesicoureteric reflux can result in contrast medium returning to the bladder from above, giving the appearance of a partially full bladder.
- **Radionuclide studies** (see p. 581). These have no place in the initial investigation of acute obstruction. Their main role is in possible long-standing obstruction to differentiate true obstructive nephropathy from retention of tracer in a baggy, low-pressure, unobstructed pelvicalyceal system.
- **Antegrade pyelography and ureterography** (see p. 580) defines the site and cause of obstruction. It can be combined with drainage of the collecting system by percutaneous needle nephrostomy.
- **Retrograde ureterography** (see p. 581) is indicated if antegrade examination cannot be carried out or if there is the possibility of dealing with ureteric obstruction from below at the time of examination. The technique carries the risk of introducing infection into an obstructed urinary tract. In obstruction due to neuromuscular dysfunction at the pelviureteric junction or retroperitoneal fibrosis, the collecting system may fill normally from below.
- **Cystoscopy, urethroscopy and urethrography** can visualize obstructing lesions within the bladder and urethra directly. Urethrography involves introducing contrast medium into the bladder by catheterization or suprapubic bladder puncture, and taking X-ray films during voiding to show obstructing lesions in the urethra. It is of particular value in the diagnosis of urethral valves and strictures.

Fig. 11.42 **Intravenous urographic X-ray taken 24 hours after injection of contrast, showing a delayed nephrogram and pyelogram on the left side and dilatation of the system to the level of the block.** By this time contrast medium has disappeared from the normal right side.

Treatment

The aim is to relieve symptoms and preserve renal function by:

- relieving the obstruction
- treating the underlying cause
- preventing and treating infection.

Temporary external drainage of urine by nephrostomy may be valuable, as this allows time for further investigation when the site and nature of the obstructing lesion are uncertain, doubt exists as to the viability of the obstructed kidney, or when immediate definitive surgery would be hazardous.

Recent, complete upper urinary tract obstruction demands urgent relief to preserve kidney function, particularly if infection is present.

In contrast, with partial urinary tract obstruction, particularly if spontaneous relief is expected – such as by passage of a calculus – there is no immediate urgency.

Surgical management depends on the cause of the obstruction (see below) and local expertise. Dialysis may be required in the ill patient prior to surgery.

Diuresis usually follows relief of obstruction at any site in the urinary tract. Massive diuresis may occur following relief of bilateral obstruction owing to previous sodium and water overload and the osmotic effect of retained solutes combined with a defective renal tubular reabsorptive capacity (as in the diuretic phase of recovering acute tubular necrosis). This diuresis is associated with increased blood volume and high levels of atrial natriuretic peptide (ANP). Defective renal tubular reabsorptive capacity cannot be the sole mechanism of severe diuresis since this phenomenon is not observed

Ask the authors
www.studentconsult.com

following relief of unilateral obstruction. The diuresis is usually self-limiting, but a minority of patients will develop severe sodium, water and potassium depletion requiring appropriate intravenous replacement. In milder cases, oral salt and potassium supplements together with a high water intake are sufficient.

Specific causes of obstruction

Calculi

These are discussed on page 609.

Pelviureteric junction obstruction

(Fig. 11.43)

This results from a functional disturbance in peristalsis of the collecting system in the absence of mechanical obstruction. Surgical attempts at correction of the obstruction by open or percutaneous pyeloplasty are indicated in patients with recurrent loin pain and those in whom serial scans or measurements of GFR indicate progressive kidney damage. Nephrectomy to remove the risk of developing pyonephrosis and septicaemia is indicated if long-standing obstruction has destroyed kidney function.

Obstructive megaureter

This childhood condition may become evident only in adult life. It results from the presence of a region of defective peristalsis at the lower end of the ureter adjacent to the ureterovesical junction. The condition is more common in males. It presents with UTI, flank pain or haematuria. The diagnosis is made on imaging with ultrasound, CT or, if necessary, ascending ureterography.

Excision of the abnormal portion of ureter with reimplantation into the bladder is always indicated in children, and in adults when the condition is associated with evidence of progressive deterioration in renal function, bacteriuria that cannot be controlled by medical means, or recurrent stone formation.

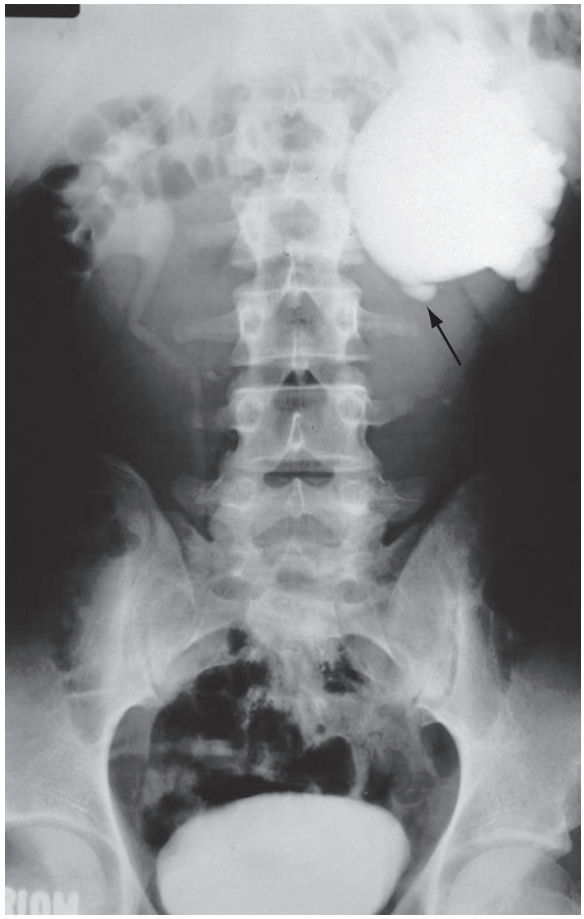


Fig. 11.43 **X-ray, showing left pelviureteric junction obstruction (arrow).**

Retroperitoneal fibrosis (RPF; chronic periaortitis)

The incidence and prevalence are 0.1 per 100 000 and 1.38 per 100 000, respectively. It is three times more common in men than in women.

The ureters become embedded in dense retroperitoneal fibrous tissue with resultant unilateral or bilateral obstruction. Obstruction is usually due to loss of peristalsis rather than occlusion. The condition may extend from the level of the second lumbar vertebra to the pelvic brim. In up to 15% of patients, the fibrotic process can extend outside the retroperitoneum, consistent with it being a systemic condition. Mediastinal fibrosis, Riedel fibrosing thyroiditis, sclerosing cholangitis, fibrotic orbital pseudotumor, fibrotic arthropathy, pleural, pericardial and lung fibrosis have been reported with increasing frequency.

Aetiology is either an autoallergic response to leakage of material, probably ceroid, from atheromatous plaques producing an inflammatory reaction, or a systemic autoimmune disease. There is an association with HLA-DRB1*03, an allele linked to various autoimmune diseases. RPF possibly is initiated as a vasa vasorum vasculitis in the aortic wall which is often seen in chronic periaortitis. This inflammatory process can cause medial wall thinning and promote atherosclerosis, and also extends into the surrounding retroperitoneum with a fibro-inflammatory reaction typical of chronic periaortitis. The autoimmune reaction to plaque antigens could be an epiphenomenon of this immune-mediated process. Activating antibodies against fibroblasts (detectable in one-third of patients) have also been implicated in the pathogenesis, as has the presence of IgG_4-bearing plasma cells; the latter is a common finding in sclerosing pancreatitis, a disorder sometimes associated with idiopathic retroperitoneal fibrosis. In addition, several infiltrating B cells show clonal or oligoclonal immunoglobulin heavy chain rearrangement. These findings raise the possibility of RPF being a primary B-cell disorder.

Pathology. The hallmark of idiopathic RPF is background sclerosis with myofibroblasts associated with a diffuse and perivascular infiltrate mainly consisting of T and B lymphocytes and IgG_4 isotype bearing plasma cells. Small vessel vasculitis may be found in approximately 50% of the patients.

Causes are many but 66% are idiopathic. *Secondary causes* include drugs (methysergide, lysergic acid, ergot derived dopamine receptor agonists (cabergoline), bromocriptine, pergolide, ergotamine, methyldopa, hydralazine, beta-blockers), malignant diseases (carcinomas of the colon, prostate, breast, stomach, carcinoid, Hodgkin's and non-

Hodgkin's lymphomas, sarcomas), infections (tuberculosis, syphilis, histoplasmosis, actinomycosis and fungal infections), and surgery/radiotherapy (lymph node resection, colectomy, hysterectomy, aortic aneurysm repair). Recognized associations include untreated abdominal aortic aneurysm, smoking and asbestosis.

Clinical. Malaise, low back pain, weight loss, testicular pain, claudication and haematuria occur.

Laboratory tests show normochromic anaemia, uraemia and a raised erythrocyte sedimentation rate (ESR) and CRP.

Imaging with ultrasound will show a poorly circumscribed periaortic mass. The test of choice is a contrast-enhanced CT, which will show the mass, lymph nodes and tumour (Fig. 11.44). MRI will show similar findings but does not require contrast. Intravenous urography is now rarely performed. Positron emission tomography (PET) is helpful (see below).

Treatment. Obstruction is relieved surgically by ureterolysis. A biopsy is performed to exclude an underlying lymphoma or carcinoma. Corticosteroids are of benefit, and in bilateral obstruction in frail patients it may be best to free only one ureter and to rely upon steroid therapy to induce regression of fibrous tissue on the contralateral side, since bilateral ureterolysis is a major operation. An alternative approach is placement of a ureteric stent or stents and corticosteroid therapy, but regular (usually 6-monthly) changes of the stent or stents are required if the periaortic mass does not regress.

Management. Response to treatment and disease activity are assessed by serial measurements of ESR and eGFR supplemented by isotopic and imaging techniques including CT scanning. Fluorodeoxyglucose-PET (FDG-PET), a functional imaging modality, assesses the metabolic activity of the retroperitoneal mass. FDG-PET also allows whole-body imaging and can detect occult malignant or infectious foci, particularly in secondary retroperitoneal fibrosis. In idiopathic RPF, FDG-PET can be used to monitor the residual inflammatory component following medical therapy Relapse after withdrawal of steroid therapy may occur and treatment may need to be continued for years. Mycophenolate or tamoxifen is also effective. Long-term follow-up is mandatory.

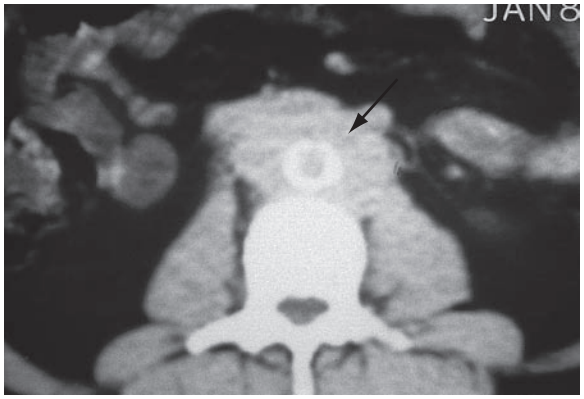


Fig. 11.44 **Retroperitoneal fibrosis (periaortitis).** Note the large mass surrounding the abdominal aorta on this CT scan (arrow).

Benign prostatic hypertrophy

Benign prostatic hypertrophy is a common cause of urinary tract obstruction. It is described on page 645.

Prognosis of urinary tract obstruction

The prognosis depends upon the cause and the stage at which obstruction is relieved. In obstruction, four factors influence the rate at which kidney damage occurs, its extent and the degree and rapidity of recovery of renal function after relief of obstruction. These are:

- whether obstruction is partial or complete
- the duration of obstruction
- whether or not infection occurs
- the site of obstruction.

Complete obstruction for several weeks will lead to irreversible or only partially reversible kidney damage. If the duration of complete obstruction is several months, total irreversible destruction of the affected kidney will result. Partial obstruction carries a better prognosis, depending upon its severity. Bacterial infection coincident with obstruction rapidly increases kidney damage. Obstruction at or below the bladder neck may induce hypertrophy and trabeculation of the bladder without a rise in pressure within the upper urinary tract, in which case the kidneys are protected from the effects of back-pressure.

DRUGS AND THE KIDNEY

Drug-induced impairment of renal function

Prerenal

Impaired perfusion of the kidneys can result from drugs that cause:

- hypovolaemia – for example:
  (a) potent loop diuretics such as furosemide, especially in elderly patients
  (b) renal salt and water loss, such as from hypercalcaemia induced by vitamin D therapy (since hypercalcaemia adversely affects renal tubular salt and water conservation)
- decrease in cardiac output, which impairs renal perfusion (e.g. beta-blockers)
- decreased renal blood flow (e.g. ACE inhibitors particularly in the presence of renovascular disease).

Renal

Several mechanisms of drug-induced renal damage exist and may coexist.

- *Acute tubular necrosis produced by direct nephrotoxicity.* Examples include prolonged or excessive treatment with aminoglycosides (e.g. gentamicin, streptomycin), amphotericin B, heavy metals or carbon tetrachloride. The combination of aminoglycosides with furosemide is particularly nephrotoxic.

FURTHER READING

Yaqoob M, Junaid I. Urinary tract obstruction. In: Warrell DA, Cox TM, Firth JD (eds) *Oxford Textbook of Medicine*, 4th edn. Oxford: Oxford University Press, 2003.

Vagilo A, Palmisano A, Corradi D et al. Retroperitoneal fibrosis: evolving concepts. *Rheumatic Disease Clinics of North America* 2007; **33**: 803–817.

- *Acute tubulointerstitial nephritis* (see p. 604) with interstitial oedema and inflammatory cell infiltration. This cell-mediated hypersensitivity nephritis occurs with many drugs, including penicillins, sulphonamides and NSAIDs (which have many other effects on the kidney; Box 11.3).
- *Chronic tubulointerstitial nephritis* due to drugs, see page 605.
- *Membranous glomerulonephritis*, e.g. penicillamine, gold, anti-TNF (p. 587).

Postrenal

Retroperitoneal fibrosis with urinary tract obstruction can result from the use of drugs (p. 616).

Use of drugs in patients with impaired renal function (Box 11.4)

Many aspects of drug handling are altered in patients with renal impairment.

Absorption

This may be unpredictable in uraemia as nausea and vomiting are frequently present.

Metabolism

Oxidative metabolism of drugs by the liver is altered in uraemia, although this is rarely of clinical significance.

The rate of drug metabolism by the kidney is reduced as a result of two factors:

- *Reduced drug catabolism*. Insulin, for example, is in part catabolized by the normal kidney. In renal disease, insulin catabolism is reduced. The insulin requirements of diabetics decline as renal function deteriorates, for this reason.
- *Reduced conversion of a precursor to a more active metabolite*, such as the conversion of 25-hydroxycholecalciferol to the more active 1,25-$(OH)_2D_3$. The 1α-hydroxylase enzyme responsible for this conversion is located in the kidney. In renal disease, production of the enzyme declines and deficiency of 1,25-$(OH)_2D_3$ results.

Protein binding

Reduced protein binding of a drug potentiates its activity and increases the potential for toxic side-effects. Measurement of the total plasma concentration of such a drug can give misleading results. For example, the serum concentration of phenytoin required to produce an antiepileptic effect is much higher in normal individuals than in those with renal failure, since in the latter proportionately more drug is present in the free form.

Some patients with renal disease are hypoproteinaemic and reduced drug-binding to protein results. This is not the sole mechanism of reduced drug-binding in such patients. For example, hydrogen ions, which are retained in renal failure, bind to receptors for acidic drugs such as sulphonamides, penicillin and salicylates, thus enhancing their potential for causing toxicity.

Volume of distribution

Salt and water overload or depletion may occur in patients with renal disease. This affects the concentration of drug obtained from a given dose.

End-organ sensitivity

The renal response to drug treatment may be reduced in renal disease. For example, mild thiazide diuretics have little diuretic effect in patients with severe renal impairment.

Box 11.3 Non-steroidal anti-inflammatory drugs and the kidney

Problem	Cause
Sodium and water retention	Reduction of prostaglandin production
Acute tubulointerstitial nephritis	Hypersensitivity reaction
Nephrotic syndrome	Membranous glomerulopathy
Analgesic nephropathy	Papillary necrosis after chronic use
Acute kidney injury	Acute tubular necrosis
Hyperkalaemia	Decreased renal excretion of K^+

Box 11.4 Safe prescribing in renal disease

Safe prescribing in renal failure demands knowledge of the clinical pharmacology of the drug and its metabolites in normal individuals and in uraemia. The clinician should ask the following questions when prescribing and discuss with the patient:

1 *Is treatment mandatory?* Unless it is, it should be withheld.
2 *Can the drug reach its site of action?* For example, there is little point in prescribing the urinary antiseptic, nitrofurantoin, in renal failure since bacteriostatic concentrations will not be attained in the urine.
3 *Is the drug's metabolism altered in uraemia?*
4 *Will accumulation of the drug or metabolites occur?* Even if accumulation is a potential problem owing to the drug or its metabolites being excreted by the kidneys, it is not necessarily an indication to change the drug given. The size of the loading dose will depend upon the size of the patient and is unrelated to renal function.
Avoidance of toxic levels of drug in blood and tissues subsequently requires the administration of normal doses of the drug at longer time intervals than usual or smaller doses at the usual time intervals.
5 *Is the drug toxic to the kidney?*
6 Are the *effective concentrations* of the drug in biological tissues similar to the toxic concentrations? Should blood levels of the drug be measured?
7 *Will the drug worsen the uraemic state* by means other than nephrotoxicity, e.g. steroids, tetracycline?
8 *Is the drug a sodium or potassium salt?* These are potentially hazardous in uraemia.

Not surprisingly, adverse drug reactions are more than twice as common in renal failure as in normal individuals. Elderly patients, in whom unsuspected renal impairment is common, are particularly at risk. Attention to the above and titration of the dose of drugs employed should reduce the problem.

The dose may be titrated by:

- observation of its clinical effect, e.g. hypotensive agents
- early detection of toxic effects
- measurement of drug levels in the blood, e.g. gentamicin levels.

Renal elimination

The major problem in the use of drugs in renal failure concerns the reduced elimination of many drugs normally excreted by the kidneys.

Water-soluble drugs such as gentamicin that are poorly absorbed from the gut, typically given by injection and not metabolized by the liver, give rise to far more problems than lipid-soluble drugs such as propranolol, which are well absorbed and principally metabolized by the liver. Metabolites of lipid-soluble drugs, however, may themselves be water-soluble and potentially toxic.

Drugs causing uraemia by effects upon protein anabolism and catabolism

Tetracyclines, with the exception of doxycycline, have a catabolic effect and as a result the concentration of nitrogenous waste products is increased. They may also cause impairment of GFR by a direct effect. Corticosteroids have a catabolic effect and so also increase the production of nitrogenous wastes. A patient with moderate impairment of renal function may therefore become severely uraemic if given tetracyclines or corticosteroid therapy.

Drugs and toxic agents causing specific renal tubular syndromes include mercury, lead, cadmium and vitamin D.

Problem patients

Particular problems are presented by patients in whom renal function is altering rapidly, such as those with recovering acute tubular necrosis. In addition, drugs may be removed by dialysis and haemofiltration, which will affect the dosage required.

ACUTE RENAL FAILURE (ARF)/ ACUTE KIDNEY INJURY (AKI)

The term 'renal failure' means failure of renal excretory function due to depression of the glomerular filtration rate. This is accompanied to a variable extent by failure of erythropoietin production (see p. 576), vitamin D hydroxylation (p. 577), regulation of acid–base balance (p. 674) and regulation of salt and water balance and blood pressure (p. 576).

Definition

Acute kidney injury has variably been defined as an abrupt deterioration in parenchymal renal function, which is usually, but not invariably, reversible over a period of days or weeks. In clinical practice, such deterioration in renal function is sufficiently severe to result in uraemia. Oliguria is usually, but not invariably, a feature. Acute kidney injury may cause sudden, life-threatening biochemical disturbances and is a medical emergency. The distinction between acute and chronic renal failure, or even acute-on-chronic renal failure, cannot be readily apparent in a patient presenting with uraemia. In view of these difficulties, the Acute Dialysis Quality Initiative group proposed the RIFLE (Risk, Injury, Failure, Loss, End-stage renal disease; Box 11.5) criteria utilizing either increases in serum creatinine or decreases in urine output. It characterizes three levels of renal dysfunction (R, I, F) and two outcome measures (L, E). These criteria indicate an increasing degree of renal damage and have a predictive value for mortality.

Box 11.5 RIFLE classification for acute kidney injury

Grade	GFR criteria	UO criteria
Risk	SCr × 1.5	UO <0.5 mL/kg/h × 6 h
Injury	SCr × 2	UO <0.5 mL/kg/h × 12 h
Failure	SCr × 3 or SCr >350 µmol/L with an acute rise >40 µmol/L	UO <0.3 mL/kg/h × 24 h
Loss	Persistent ARF >4 weeks	
ESKD	Persistent renal failure >3 months	

SCr, serum creatinine; UO, urinary output; ESKD, end-stage kidney disease
RIFLE, Risk, Injury, Failure, Loss, End-stage renal disease

Epidemiology

The observed incidence and outcomes of AKI are highly dependent upon the populations studied and the definition of AKI employed. The incidence of hospital-acquired AKI has increased. The incidence of community-acquired AKI on admission to hospital is approximately 1% in the UK, with AKI superimposed on chronic kidney disease accounting for 50% of these patients. The incidence of severe AKI (creatinine >500 µmol/L) is about 130–140/million population and that of less severe AKI (patients with creatinine levels of up to 177 µmol/L or a 50% rise from baseline) is about 200/million/year. Uncomplicated AKI can usually be managed outside the intensive care unit (ITU) setting and carries a good prognosis, with mortality rates less than 5–10%. In contrast, AKI complicating non-renal organ system failure (in the ITU setting) is associated with mortality rates of 50–70%, which have not changed for several decades. Moreover, sepsis-related AKI has a significantly worse prognosis than AKI in the absence of sepsis.

Classification

Renal failure results in reduced excretion of nitrogenous waste products, of which urea is the most commonly measured. A raised serum urea concentration (uraemia) is classified as: (i) prerenal, (ii) renal or (iii) postrenal. More than one category may be present in an individual patient. Other causes of altered serum urea and creatinine concentration are shown in Table 11.16.

Prerenal uraemia

In prerenal uraemia, there is impaired perfusion of the kidneys with blood. This results either from hypovolaemia, hypotension, impaired cardiac pump efficiency or vascular disease limiting renal blood flow, or combinations of these factors. Usually the kidney is able to maintain glomerular filtration close to normal despite wide variations in renal perfusion pressure and volume status – so-called 'autoregulation'. Further depression of renal perfusion leads to a drop in glomerular filtration and development of prerenal uraemia. Drugs which impair renal autoregulation, such as ACE inhibitors and NSAIDs, increase the tendency to develop prerenal uraemia.

Table 11.16 Causes of altered serum urea and creatinine concentration other than altered renal function

	Decreased concentration	Increased concentration
Urea	Low protein intake Liver failure Sodium valproate treatment	Corticosteroid treatment Tetracycline treatment Gastrointestinal bleeding
Creatinine	Low muscle mass	High muscle mass Red meat ingestion Muscle damage (rhabdomyolysis) Decreased tubular secretion (e.g. cimetidine, trimethoprim therapy)

All causes of prerenal uraemia may lead to established parenchymal kidney damage and the development of acute renal failure. By definition, excretory function in prerenal uraemia improves once normal renal perfusion has been restored.

A number of criteria have been proposed to differentiate between prerenal and intrinsic renal causes of uraemia (Table 11.17).

- *Urine specific gravity* and *urine osmolality* are easily obtained measures of concentrating ability but are unreliable in the presence of glycosuria or other osmotically active substances in the urine.
- *Urine sodium* is low if there is avid tubular reabsorption, but may be increased by diuretics or dopamine.
- *Fractional excretion of sodium (FENa)*, the ratio of sodium clearance to creatinine clearance, increases the reliability of this index but may remain low in some 'intrinsic' renal diseases, including contrast nephropathy and myoglobinuria.

Laboratory tests, however, are no substitute for clinical assessment. A history of blood or fluid loss, sepsis potentially leading to vasodilatation, or of cardiac disease may be helpful. Hypotension (especially postural), a weak rapid pulse and a low jugular venous pressure will suggest that the uraemia is prerenal. In doubtful cases, measurement of central venous pressure is often invaluable, particularly with fluid challenge (see p. 901).

Management

If the prerenal uraemia is a result of hypovolaemia and hypotension, prompt replacement with appropriate fluid is essential to correct the problem and prevent development of ischaemic renal injury and acute kidney disease (p. 904). Since prerenal and renal uraemia may coexist, and fluid challenge in the latter situation may lead to volume overload with pulmonary oedema, careful clinical monitoring is vital. Blood pressure should be checked regularly and signs of elevated jugular venous pressure and of pulmonary oedema sought frequently. Central venous pressure monitoring is usually advisable (see p. 901). If the problem relates to cardiac pump insufficiency or occlusion of the renal vasculature, appropriate measures – albeit often unsuccessful – need to be taken.

Table 11.17 Criteria for distinction between prerenal and intrinsic causes of renal dysfunction

	Prerenal	Intrinsic
Urine specific gravity	>1.020	<1.010
Urine osmolality (mOsm/kg)	>500	<350
Urine sodium (mmol/L)	<20	>40
$FE_{Na} = \frac{U_{Na}}{P_{Na}} \div \frac{U_{Cr}}{P_{Cr}} \times 100$	<1%	>1%

FE, fractional excretion; P, plasma; U, urine; Cr, creatinine; Na, sodium

Table 11.18 Some causes of acute tubular necrosis

- Haemorrhage
- Burns
- Diarrhoea and vomiting, fluid loss from fistulae
- Pancreatitis
- Diuretics
- Myocardial infarction
- Congestive cardiac failure
- Endotoxic shock
- Snake bite
- Myoglobinaemia
- Haemoglobinaemia (due to haemolysis, e.g. in falciparum malaria, 'blackwater fever')
- Hepatorenal syndrome
- Radiological contrast agents (see p. 624)
- Drugs, e.g. aminoglycosides, NSAIDs, ACE inhibitors, platinum derivatives
- Abruptio placentae
- Pre-eclampsia and eclampsia

Postrenal uraemia

Here, uraemia results from obstruction of the urinary tract at any point from the calyces to the external urethral orifice. The causes and presentation of urinary tract obstruction are dealt with on page 614. Screening for urinary tract obstruction is by renal ultrasonography. Urinary tract obstruction may present in an acute fashion (if obstruction of a single functioning kidney by, for example, a calculus occurs) but typically is of insidious onset.

Acute uraemia due to renal parenchymal disease

Causes

This is most commonly due to acute renal tubular necrosis (Table 11.18). Other causes include disease affecting the intrarenal arteries and arterioles as well as glomerular capillaries, such as a vasculitis (p. 548), accelerated hypertension, cholesterol embolism, haemolytic uraemic syndrome, thrombotic thrombocytopenic purpura (TTP), pre-eclampsia and crescentic glomerulonephritis. Acute tubulointerstitial nephritis (p. 604) may also cause acute renal failure. This also

occurs when renal tubules are acutely obstructed by crystals, for example following sulphonamide therapy in a dehydrated patient (sulphonamide crystalluria) or after rapid lysis of certain malignant tumours following chemotherapy (acute hyperuricaemic nephropathy). Acute bilateral suppurative pyelonephritis or pyelonephritis of a single kidney can cause acute uraemia.

Acute tubular necrosis

Causes

Acute tubular necrosis (ATN) is common, particularly in hospital practice. It results most often from renal ischaemia but can also be caused by direct renal toxins including drugs such as the aminoglycosides, lithium and platinum derivatives (Table 11.18).

Kidneys are particularly vulnerable to ischaemic injury when cholestatic jaundice is present, and more than one ischaemic factor appears to be present in some situations. For example, disseminated intravascular coagulation complicating Gram-negative septicaemia and complications of pregnancy such as placental rupture, pre-eclampsia and eclampsia may result in occlusion or partial occlusion of intrarenal vessels, exacerbating the ischaemic insult resulting from hypotension associated with the underlying condition.

Myoglobinaemia and haemoglobinaemia consequent upon muscle injury (rhabdomyolysis) complicating trauma, pressure necrosis or heroin use predispose to ATN, perhaps in part owing to occlusion of renal tubules by myoglobin and haemoglobin casts. In liver failure, acute renal failure appears to result from rapidly reversible vasomotor abnormalities within the kidney. A kidney removed from a patient with hepatic cirrhosis and liver failure dying with oliguric renal failure may function normally immediately after transplantation into a normal individual. Efferent glomerular arteriolar dilatation resulting from ACE-inhibitor drug therapy, with consequent lowering of glomerular filtration pressure, may cause acute deterioration in excretory function if renal arterial disease is also present. The effect is compounded by concomitant use of non-steroidal anti-inflammatory agents which reduce prostaglandin production, opposing this effect.

Pathogenesis

Factors postulated to be involved in the development of ATN include:

- *Intrarenal microvascular vasoconstriction:*
  (a) Vasoconstriction is increased in response to endothelin, adenosine, thromboxane A_2, leukotrienes and sympathetic nerve activity. However, endothelin antagonists failed to show any beneficial effect in the clinical setting.
  (b) Vasodilatation is impaired due to reduced sensitivity in response to:
    (i) Nitric oxide, prostaglandins (PGE_2), acetylcholine and bradykinin.
    (ii) Increased endothelial and vascular smooth muscle cell structural damage.
  (c) Increased leucocyte–endothelial adhesion, vascular congestion and obstruction, leucocyte activation and inflammation. After success in the prevention of ARF in animal models, anti-ICAM (intercellular adhesion molecule) in the clinical setting failed to live up to its initial promise.
- *Tubular cell injury*. Ischaemic injury results in rapid depletion of intracellular ATP stores resulting in cell death either by necrosis or apoptosis, due to the following:
  (i) entry of calcium into cells with an increase in cytosolic cell calcium concentration
  (ii) induction by hypoxia of inducible nitric oxide synthases with increased production of nitric oxide causing cell death
  (iii) increased production of intracellular proteases such as calpain, which cause proteolysis of cytoskeletal proteins and cell wall collapse
  (iv) activation of phospholipase A_2 with increased production of free fatty acids, particularly arachidonic acid, due to its action on the lipid layer of cell membranes
  (v) cell injury resulting from reperfusion with blood after initial ischaemia causing excessive free radical generation
  (vi) tubular obstruction by desquamated viable or necrotic cells and casts
  (vii) loss of cell polarity, i.e. integrins located on the basolateral side of the cell are translocated to the apical surface, which when combined with other desquamated cells forms casts, with tubular obstruction and back leak of tubular fluid.
- *Tubular cellular recovery*. Tubular cells have the capacity to regenerate rapidly and to reform the disrupted tubular basement membrane, which explains the reversibility of ATN. Multiple growth factors, including insulin-like growth factor 1, epidermal growth factor and hepatocyte growth factor, and their receptors are upregulated during the regenerative process after injury. Clinical utility of growth factors in the treatment of ARF is unproven.

In established ATN renal blood flow is much reduced, particularly blood flow to the renal cortex. Ischaemic tubular damage contributes to a reduction in glomerular filtration by a number of interrelated mechanisms:

- *Glomerular contraction* reducing the surface area available for filtration, due to reflex afferent arteriolar spasm mediated by increased solute delivery to the macula densa. Increased solute delivery is due to impaired sodium absorption in the proximal tubular cells because of loss of cell polarity with mislocalization of the Na^+/K^+-ATPase and impaired tight junction integrity, resulting in decreased apical-to-basal transcellular sodium absorption.
- *'Back leak' of filtrate* in the proximal tubule owing to loss of function of the tubular cells.
- *Obstruction of the tubule* by debris shed from ischaemic tubular cells; these appear on renal biopsy as flat rather than the normal tall appearance (Fig. 11.45).

Course

The clinical course of acute renal failure associated with ATN is variable depending on the severity and duration of the renal insult. Oliguria is common in the early stages: non-oliguric renal failure is usually a result of a less severe renal insult. Recovery of renal function typically occurs after 7–21 days, although recovery is delayed by continuing sepsis. In the recovery phase, GFR may remain low while urine output increases, sometimes to many litres a day owing to defective

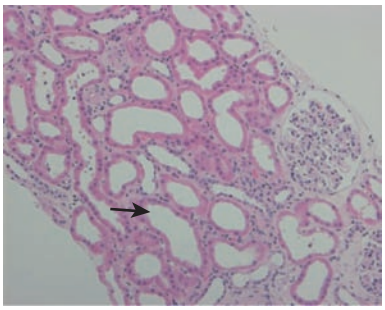


Fig. 11.45 **Acute tubular necrosis** showing effacement and loss of the proximal tubule brush border, patchy loss of tubular cells and focal areas of proximal tubule dilatation.

tubular reabsorption of filtrate. The clinical course is variable and ATN may last for up to 6 weeks, even after a relatively short-lived initial insult. Eventually renal function usually returns to almost normal or to normal, although exceptions exist (e.g. in renal cortical necrosis – see below).

No treatment is, as yet, known which will reduce the duration of acute renal tubular necrosis once it has occurred. The use of intravenous mannitol, furosemide or 'renal-dose' dopamine is not supported by controlled trial evidence, and none of these treatments is without risk. Trials with the synthetic analogue of atrial natriuretic factor, anaritide, and the use of exogenous IGF-1 have failed to show any beneficial effect in clinical ARF so far.

Whether a state of 'incipient' ATN exists in some patients with prerenal uraemia, and whether ATN can be prevented by administration of mannitol, furosemide or dopamine, also remain uncertain. Many nephrologists will administer one or more of these agents if correction of prerenal factors does not initiate a diuresis, but proof of benefit is lacking.

Clinical and biochemical features

These are the features of the causal condition together with features of rapidly progressive uraemia. The rate at which serum urea and creatinine concentrations increase is dependent upon the rate of tissue breakdown in the individual patient. This is increased in the presence of trauma, sepsis and following surgery. *Hyperkalaemia* is common, particularly following trauma to muscle and in haemolytic states. *Metabolic acidosis* is usual unless hydrogen ion loss by vomiting or aspiration of gastric contents is a feature. *Hyponatraemia* may be present owing to water overload if patients have continued to drink in the face of oliguria, or if overenthusiastic fluid replacement with 5% dextrose has been carried out. Pulmonary oedema owing to salt and water retention is not uncommon, particularly after inappropriate attempts to initiate a diuresis by infusion of 0.9% saline without adequate monitoring of the patient's volume status. Hypocalcaemia due to reduced renal production of 1,25-dihydroxycholecalciferol and hyperphosphataemia due to phosphate retention are common.

Symptoms of uraemia such as anorexia, nausea, vomiting and pruritus develop, followed by intellectual clouding, drowsiness, fits, coma and haemorrhagic episodes. Epistaxes and gastrointestinal haemorrhage are relatively common. Severe infection may have initiated the acute kidney disease or have complicated it owing to the impaired immune defences of the uraemic patient or ill-considered management, such as the insertion and retention of an unnecessary bladder catheter with complicating urinary tract infection and bacteraemia.

Investigation of the uraemic emergency

Investigations are aimed at defining whether the patient has acute or chronic uraemia, whether uraemia results from prerenal, renal or postrenal factors, and establishing the cause.

Acute or chronic uraemia?

The distinction between acute and chronic uraemia depends in part on the history, duration of symptoms and previous urinalysis or measurements of renal function.

A rapid rate of change of serum urea and creatinine with time suggests an acute process. A normochromic, normocytic anaemia suggests chronic disease, but anaemia may complicate many of the diseases that cause acute renal failure, owing to a combination of haemolysis, haemorrhage and deficient erythropoietin production.

Ultrasound assessment of renal echogenicity and size is helpful. Small kidneys of increased echogenicity are diagnostic of a chronic process, although the reverse is not true; the kidney may remain normal in size in diabetes and amyloidosis, for instance.

Evidence of renal osteodystrophy (for example, digital subperiosteal erosions due to hyperparathyroid bone disease) is indicative of chronic disease.

Measurement of carbamylated haemoglobin (a product of non-enzymatic reaction between urea and haemoglobin, cf. glycosylated haemoglobin) is not widely employed.

Prerenal, renal or postrenal uraemia?

Bladder outflow obstruction is ruled out by insertion of a urethral catheter or flushing of an existing catheter, which should then be removed unless a large volume of urine is obtained. Absence of upper tract dilatation on renal ultrasonography will, with very rare exceptions, rule out urinary tract obstruction.

The distinction between prerenal and renal uraemia may be difficult (see p. 619). Assessment of the patient's volume status is essential and central venous pressure measurement is extremely helpful. If volume status is low, appropriate corrective measures are indicated. If no diuresis ensues, acute intrinsic renal failure is present.

Other investigations

Urinalysis, urine microscopy, particularly for red cells and red-cell casts (indicative of glomerulonephritis), and urine culture. Urine should be tested for free haemoglobin and myoglobin, where appropriate.

Blood tests include measurement of serum urea, electrolytes, creatinine, calcium, phosphate, albumin, alkaline phosphatase and urate concentrations, as well as full blood count and examination of the peripheral blood film where necessary. Coagulation studies, blood cultures and measurements of nephrotoxic drug blood levels should be carried out.

Management

The aim of management of acute renal tubular necrosis is to keep the patient alive until spontaneous recovery of renal function occurs. Ideally patients should be managed by a

nephrologist or intensivist with access to facilities for blood purification and fluid removal (see below). Early specialist referral is advisable. Poor initial management and late referral result in the arrival in the specialist centre of a patient who is severely uraemic, acidotic and hyperkalaemic, with pulmonary oedema following overenthusiastic intravenous fluid administration and with a Gram-negative septicaemia complicating the presence of an unnecessary indwelling bladder catheter.

General measures

Good nursing and physiotherapy are vital. Regular oral toilet, chest physiotherapy and consistent documentation of fluid intake and output, and where possible measurement of daily bodyweight to assess fluid balance changes, all have a role. The patient should be confined to bed only if essential.

Emergency measures

Hyperkalaemia

This is a life-threatening complication owing to the risk of cardiac dysrhythmias, particularly ventricular fibrillation. Treatment is outlined in Emergency box 12.1. Correction of acidosis with intravenous sodium bicarbonate will also reduce serum potassium concentration, but administration of sodium may be inappropriate if the patient is salt and water overloaded. Rapid correction of acidosis in a hypocalcaemic patient may also trigger tetany, since hydrogen ions displace calcium from albumin-binding sites, thus increasing the physiologically active calcium concentration in blood. Ion exchange resins are used to prevent subsequent hyperkalaemia rather than to deal with the acute emergency. In many patients, hyperkalaemia will be controlled only by dialysis or haemofiltration.

Pulmonary oedema

Unless a diuresis can be induced with intravenous furosemide, dialysis or haemofiltration will be required.

Sepsis

Infections, when detected, should be treated promptly, bearing in mind the need to avoid nephrotoxic drugs and to use drugs with appropriate monitoring and drug levels (e.g. gentamicin, vancomycin). Prophylactic antibiotics or barrier nursing is not recommended in all cases.

Use of drugs

Great care must be exercised in the use of drugs (see p. 618).

Fluid and electrolyte balance

Twice-daily clinical assessment is needed. In general, once the patient is euvolaemic, daily fluid intake should equal urine output plus losses from fistulae and from vomiting, plus an allowance of 500 mL daily for insensible loss. Febrile patients will require an additional allowance. Sodium and potassium intake should be minimized. If abnormal losses of fluid occur, for example in diarrhoea, additional fluid and electrolytes will be required. The development of signs of salt and water overload (peripheral oedema, basal crackles, elevation of jugular venous pressure) or of hypovolaemia should prompt reappraisal of fluid intake. Large changes in daily weight reflecting change in fluid balance status should also prompt a reappraisal of the situation.

Diet

With rare exceptions, sodium and potassium restriction are appropriate. The place of dietary protein restriction is controversial. If it is hoped to avoid dialysis or haemofiltration, protein intake is sometimes restricted to approximately 40 g daily. This poses the risk of a negative nitrogen balance despite attempts to reduce endogenous protein catabolism by maintenance of a high energy intake in the form of carbohydrate and fat. Patients treated by blood purification techniques are more appropriately managed by providing 70 g protein daily or more. Hypercatabolic patients will require an even higher nitrogen intake to prevent negative nitrogen balance.

Routes of intake are, in preferred order, enteral by mouth, enteral by nasogastric tube, and parenteral. The last of these is, however, only necessary if vomiting or bowel dysfunction render the enteral route inappropriate.

Vitamin supplements are usually supplied. Vitamin D analogue therapy and pharmacological doses of erythropoietin are not employed routinely.

Dialysis and haemofiltration

The main indications for blood purification and/or excess fluid removal by these techniques are:

- symptoms of uraemia
- complications of uraemia, such as pericarditis
- severe biochemical derangement in the absence of symptoms (especially if a rising trend is observed in an oliguric patient and in hypercatabolic patients)
- hyperkalaemia not controlled by conservative measures
- pulmonary oedema
- severe acidosis
- for removal of drugs causing the acute renal failure, e.g. gentamicin, lithium, severe aspirin overdose.

The main options are peritoneal dialysis, intermittent haemodialysis (HD) combined with ultrafiltration, if necessary, intermittent haemofiltration, continuous arteriovenous or venovenous haemofiltration, and haemodiafiltration. For reasons that are incompletely understood, adverse cardiovascular effects are much less during haemofiltration than during haemodialysis. Continuous treatments are superior to intermittent ones in this respect.

Continuous renal replacement treatments (CRRT)

Blood flow is achieved either by using the patient's own blood pressure to generate arterial blood flow through a filter or by the use of a blood pump to draw blood from the lumen of a dual-lumen catheter placed in the jugular, subclavian or femoral vein.

Continuous arteriovenous or venovenous haemofiltration (CAVH, CVVH) refers to the continuous removal of ultrafiltrate from the patient, usually at rates of up to 1000 mL/h, combined with simultaneous infusion of replacement solution. For instance, in a fluid-overloaded patient one might remove filtrate at 1000 mL/h and replace at a rate of 900 mL/h, achieving a net fluid removal of 100 mL/h.

Continuous haemodiafiltration (CAVHDF, CVVHDF) is a combination of haemofiltration and haemodialysis, involving both the net removal of ultrafiltrate from the blood and its

replacement with a replacement solution, together with the countercurrent passage of dialysate (which may be identical to the replacement solution). Both the ultrafiltrate and the spent dialysate appear as 'waste'.

Comparisons of dialysis modalities

Peritoneal dialysis (PD) is used infrequently in the management of ARF, with decreasing utilization over the past 5–10 years. Drawbacks to the use of PD in ARF are:

(i) low efficiency in fluid and solute removal compared to CRRT or intermittent HD;
(ii) ARF complicating intra-abdominal pathology is unsuitable for PD;
(iii) increasing intra-abdominal pressure can compromise lung function;
(iv) use of dialysis fluids with a high dextrose content may produce hyperglycaemia and other metabolic derangements. Data suggest that PD is significantly less effective than CRRT in the management of ARF and should be reserved for situations where other modalities of therapy are not available.

Several randomized trials comparing intermittent HD to CRRT in ARF failed to demonstrate any difference in morbidity or mortality between the two modalities. There are thus insufficient data to favour either HD or CRRT as a superior mode of therapy in ARF. However, there is consensus that in haemodynamically unstable patients, CRRT is better tolerated.

Membrane biocompatibility

The concept of membrane 'biocompatibility' relates to the activation of cellular (neutrophils, platelets) and humoral (complement system and coagulation cascade) components upon contact between blood and dialysis membranes. As a general rule, unsubstituted cellulose membranes (cuprophan) are the least biocompatible, with biocompatibility improving with substitution of free hydroxyl groups by tertiary amino groups (hemophan), acetate (cellulose acetate, diacetate, triacetate) or the use of synthetic polymers (e.g. polysulphone, polyamide, polyacrilonitrile and polymethylmethacrylate).

Synthetic membranes appear to confer a significant survival advantage over unsubstituted cellulose (cuprophan)-based membranes but no benefit on recovery of renal function.

Acute kidney injury in the intensive care unit (ITU)

Increasing numbers of patients with acute renal failure are managed in the setting of an intensive care unit. Many such patients have multiorgan failure, sepsis or both, with associated cardiovascular instability. Continuous methods of blood purification and control of fluid balance, such as venovenous haemofiltration, are preferable to intermittent haemodialysis or peritoneal dialysis in such patients. Advantages include:

- much less disturbance of cardiovascular stability
- the ability to generate as much 'space' for fluid administration as is required, which can be adjusted flexibly to the needs of the patient (many patients require large volumes of fluid to be administered for nutritional and other reasons)
- the removal of potentially harmful substances such as inflammatory cytokines via the more porous membrane employed in haemofiltration.

Acute respiratory distress syndrome (ARDS) is not uncommon in patients with multiorgan failure, including acute kidney injury, requiring intensive therapy. In such patients the wish to remove as much fluid from the patient as possible to reduce pulmonary congestion must be balanced against the need of organs, including the kidneys, for an adequate blood flow, if recovery is to occur. Anaemia is relatively common in patients admitted to ITU and is managed by blood transfusions. In a study, erythropoietin (EPO) given weekly did not reduce the blood transfusion requirements but unexpectedly increased survival in trauma patients. This also showed that EPO has non-haemopoietic pleiotrophic effects by which it reduces the risk of acute ischaemia and reperfusion injury in multiple organs.

Management of the recovery phase

Usually, after 1–3 weeks, renal function improves, as evidenced by an increase in urine volume and improvement in serum biochemistry. Dialysis or haemofiltration, if they have been required, can be discontinued. A careful watch on clinical state, salt and water balance, and serum chemistry is required at this stage, particularly if a major diuretic phase develops owing to recovery of glomerular filtration at a time when renal tubular reabsorptive capacity for sodium, potassium and water remains impaired. Intravenous fluid replacement is sometimes required together with supplements of sodium chloride and potassium. Typically, the diuretic phase lasts for only a few days.

Acute cortical necrosis

Renal hypoperfusion results in diversion of blood flow from the cortex to the medulla, with a drop in GFR. Medullary ischaemic damage is largely reversible owing to the capacity of the tubular cells for regeneration. In contrast, glomerular ischaemic injury does not heal with regeneration but with scarring – glomerulosclerosis. Prolonged cortical ischaemia may lead to irreversible loss of renal function termed 'cortical necrosis'. This may be patchy or complete. Any cause of acute tubular necrosis, if sufficiently severe or prolonged, may lead to cortical necrosis. This outcome is particularly common if acute kidney injury has been accompanied by derangements of the vascular endothelial system or coagulation system, such as occurs in haemolytic uraemic syndrome and complications of pregnancy.

Contrast nephropathy

In patients with impaired renal function, iodinated radiological contrast media may be nephrotoxic, possibly by causing renal vasoconstriction and by a direct toxic effect upon renal tubules. The effect is dose-dependent and therefore more commonly seen in procedures that require large amounts of contrast media, such as angiography with or without angioplasty. In many patients the effect is mild, transient, fully reversible and of no clinical significance. The risk and severity of contrast nephropathy is amplified by the presence of hypovolaemia and renal impairment, especially if due to diabetic nephropathy. Diabetes per se is not a risk

factor. However, metformin can precipitate acidosis and should be stopped and not restarted until renal function returns to the baseline level.

Prevention involves minimization as far as possible of the dose of contrast employed and use of an iso-osmolar or low-osmolality contrast medium. Superiority of iso-osmolar agents has been questioned by randomized trials in patients with severe renal impairment (eGFR <30 mL/min) undergoing coronary angiogram, where no differences were seen between the two types of contrast agent. Pre-hydration with intravenous saline is of proven benefit. A popular regimen involves infusion of 1 L of 0.9% saline during the 12 hours before and 12 hours after contrast exposure. Care must be taken to avoid volume overload in susceptible patients.

N-acetylcysteine (NAC; a potent antioxidant) given 48 hours prior to radiological intervention may be of benefit in preventing worsening of re-existing renal impairment following intravenous contrast, but the effect on morbidity and mortality is unknown. Routine use of dopamine, theophylline (an adenosine antagonist) and prophylactic haemodialysis (removing contrast agent from circulation) is of no benefit. In patients with advanced CKD who are undergoing coronary angiography, periprocedural haemofiltration given in an ITU setting appears to be effective in preventing the deterioration of renal function due to contrast agent-induced nephropathy and is associated with improved in-hospital and long-term outcomes.

When deterioration in renal function occurs after intra-arterial injection of contrast (for example, after coronary angiography) it may be difficult to differentiate the effects of contrast-induced damage from those of atheromatous embolization (see p. 609). The latter carries a worse prognosis.

N.B. Gadolinium used in contrast MRI causes problems in chronic renal failure (see p. 580).

Hepatorenal syndrome (HRS)

The renal failure observed in HRS results from profound renal vasoconstriction with histologically normal kidneys (p. 354). Although many of the features of HRS resemble prerenal ARF, the defining feature is a lack of improvement in renal function with volume expansion. Renal recovery is usually observed after restoration of hepatic function after successful liver transplantation.

CHRONIC KIDNEY DISEASE (CKD)

The term 'CKD' is now replacing terms such as chronic renal failure or insufficiency. CKD implies long-standing, and usually progressive, impairment in renal function. In many instances, no effective means are available to reverse the primary disease process. Exceptions include correction of urinary tract obstruction, immunosuppressive therapy for systemic vasculitis and Goodpasture's syndrome, treatment of accelerated hypertension, and correction of critical narrowing of renal arteries causing renal impairment. The rate of deterioration in renal function can, however, be slowed (see p. 632). A list of causes of CKD is given in Table 11.19.

Table 11.19 Causes of chronic kidney disease

Congenital and inherited disease
- Polycystic kidney disease (adult and infantile forms)
- Medullary cystic disease
- Tuberous sclerosis
- Oxalosis
- Cystinosis
- Congenital obstructive uropathy

Glomerular disease
- *Primary glomerulonephritides* including focal glomerulosclerosis
- *Secondary glomerular disease* (systemic lupus, polyangiitis, Wegener's granulomatosis, amyloidosis, diabetic glomerulosclerosis, accelerated hypertension, haemolytic uraemic syndrome, thrombotic thrombocytopenic purpura, systemic sclerosis, sickle cell disease)

Vascular disease
- Hypertensive nephrosclerosis (common in black Africans)
- Reno-vascular disease
- Small and medium-sized vessel vasculitis

Tubulointerstitial disease
- Tubulointerstitial nephritis – idiopathic, due to drugs (especially nephrotoxic analgesics), immunologically mediated
- Reflux nephropathy
- Tuberculosis
- Schistosomiasis
- Nephrocalcinosis
- Multiple myeloma (myeloma kidney)
- Balkan nephropathy
- Renal papillary necrosis (diabetes, sickle cell disease and trait, analgesic nephropathy)
- Chinese herb nephropathy

Urinary tract obstruction
- Calculus disease
- Prostatic disease
- Pelvic tumours
- Retroperitoneal fibrosis
- Schistosomiasis

Wide geographical variations in the incidence of disorders causing CKD exist. The most common cause of glomerulonephritis in sub-Saharan Africa is malaria. Schistosomiasis is a common cause of renal failure due to urinary tract obstruction in parts of the Middle East, including southern Iraq. The incidence of end-stage renal failure varies between racial groups. End-stage renal failure is three to four times as common in black Africans in the UK and USA as it is in whites, and hypertensive nephropathy is a much more frequent cause of end-stage renal failure in this group. The prevalence of diabetes mellitus and hence of diabetic nephropathy is higher in some Asian groups than in whites. The age is of relevance; CKD due to atherosclerotic renal vascular disease is much more common in the elderly than in the young. Over 70% of all cases with CKD are due to diabetes mellitus, hypertension and atherosclerosis.

Patients with CKD should be referred to a nephrologist with access to facilities for renal replacement therapy at an early stage since late referral has been shown to be associated with increased mortality and morbidity when such

FURTHER READING

Coca SG, Yalavarthy R, Concato J, Parikh CR. Biomarkers for the diagnosis and risk stratification of acute kidney injury: a systematic review. *Kidney International* 2008; **73**: 1008–1016.

Lameire N, Van Biesen W, Vanholder R (2005) Acute renal failure. *Lancet* **365**: 417–430.

Rondon-Berrios H, Palevsky PM. Treatment of acute kidney injury: an update on the management of renal replacement therapy. *Current Opinion in Nephrology and Hypertension*. 2007; **16**: 64–70.

Table 11.20 Stages of chronic kidney disease

Stage	Description	GFR (mL/min/1.73 m^2)
1	Kidney damage with normal or ↑ GFR	≥90
2	Kidney damage with mild ↓ GFR	60–89
3	Moderate ↓ GFR	30–59
4	Severe ↓ GFR	15–29
5	Kidney failure	<15 (or dialysis)

Chronic kidney disease is defined as either kidney damage or GFR <60 mL/min/1.73 m^2 for ≥3 months
Kidney damage is defined as pathological abnormalities or markers of damage, including abnormalities in blood or urine tests or imaging studies
From Kidney Disease Outcomes Quality Initiative of the National Kidney Foundation, USA

patients commence renal replacement therapy. Old age is no bar to referral in the reasonably fit elderly patient.

For CKD referral and treatment pathways the Kidney Disease Outcomes Quality Initiative (K/DOQI) of the National Kidney Foundation established a classification of CKD in the USA in 2002, which is internationally accepted (Table 11.20). This classification defines CKD as a GFR <60 mL/min/1.73 m^2 or a GFR >60 mL/min/1.73 m^2 together with the presence of kidney damage, present for >3 months. Use of this classification requires doctors to estimate GFR, which has been made possible by the adoption of the MDRD formula in most chemical pathology laboratories. Large prevalences of CKD have now been found in population-based surveys using the MDRD method to estimate GFR.

Clinical approach to the patient with CKD or any other form of renal disease

History

Particular attention should be paid to:

- *duration of symptoms*
- *drug ingestion,* including non-steroidal anti-inflammatory agents, analgesic and other medications, and unorthodox treatments such as herbal remedies
- *previous medical and surgical history,* e.g. previous chemotherapy, multisystem diseases such as SLE, malaria
- *previous occasions* on which urinalysis or measurement of urea and creatinine might have been performed, e.g. pre-employment or insurance medical examinations, new patient checks
- *family history* of renal disease.

Symptoms

The early stages of renal failure are often completely asymptomatic, despite the accumulation of numerous metabolites. Serum urea and creatinine concentrations are measured in renal failure since methods for their determination are available and a rough correlation exists between urea and creatinine concentrations and symptoms. These substances are, however, in themselves not particularly toxic. The nature of the metabolites that are involved in the genesis of symptoms is unclear. Such metabolites must be products of protein catabolism (since dietary protein restriction may reverse symptoms associated with renal failure) and many of them must be of relatively small molecular size (since haemodialysis employing membranes which allow through only relatively small molecules improves symptoms). Little else is known with certainty.

Symptoms are common when the serum urea concentration exceeds 40 mmol/L, but many patients develop uraemic symptoms at lower levels of serum urea. Symptoms include:

- malaise, loss of energy
- loss of appetite
- insomnia
- nocturia and polyuria due to impaired concentrating ability
- itching
- nausea, vomiting and diarrhoea
- paraesthesiae due to polyneuropathy
- 'restless legs' syndrome (overwhelming need to frequently alter position of lower limbs) (p. 1197)
- bone pain due to metabolic bone disease
- paraesthesiae and tetany due to hypocalcaemia
- symptoms due to salt and water retention – peripheral or pulmonary oedema
- symptoms due to anaemia (see p. 392)
- amenorrhoea in women; erectile dysfunction in men.

In more advanced uraemia (serum urea >50–60 mmol/L), these symptoms become more severe, and CNS symptoms are common:

- mental slowing, clouding of consciousness, and seizures
- myoclonic twitching.

Severe depression of glomerular filtration can result in oliguria. This can occur with either acute kidney injury or in the terminal stages of chronic kidney disease. However, even if the GFR is profoundly depressed, failure of tubular reabsorption may lead to very high urine volumes; the urine output is therefore not a useful guide to renal function.

Examination

There are few physical signs of uraemia per se. Findings include short stature (in patients who have had chronic kidney disease in childhood), pallor (due to anaemia), increased photosensitive pigmentation (which may make the patient look misleadingly healthy), brown discoloration of the nails, scratch marks due to uraemic pruritus, signs of fluid overload (p. 655), pericardial friction rub, flow murmurs (mitral regurgitation due to mitral annular calcification, aortic and pulmonary regurgitant murmurs due to volume overload), and glove and stocking peripheral sensory loss (rare).

The kidneys themselves are usually impalpable unless grossly enlarged as a result of polycystic disease, obstruction or tumour. Rectal and vaginal examination may disclose evidence of an underlying cause of renal failure, particularly urinary obstruction, and should always be performed.

In addition to these findings, there may be physical signs of any underlying disease which may have caused the renal failure, for instance:

- cutaneous vasculitic lesions in systemic vasculitides
- retinopathy in diabetes
- evidence of peripheral vascular disease
- evidence of spina bifida or other causes of neurogenic bladder.

An assessment of the central venous pressure, skin turgor, blood pressure both lying and standing and peripheral circulation should also be made. The major symptoms and signs of chronic kidney disease are shown in Figure 11.46.

Investigations

The following investigations are common for all renal patients. This includes patients with glomerular or non-glomerular disease, renal involvement in systemic diseases, AKI and CKD, as renal symptoms and signs are non-specific.

Urinalysis

- *Haematuria* may indicate glomerulonephritis, but other sources must be excluded. Haematuria should not be assumed to be due to the presence of an indwelling catheter.
- *Proteinuria*, if heavy, is strongly suggestive of glomerular disease. Urinary infection may also cause proteinuria. *Glycosuria* with normal blood glucose is common in CKD.
- *Urine culture*, including early-morning urine samples for TB.

Urine microscopy (see p. 579)

- *White cells* in the urine usually indicate active bacterial urinary infection, but this is an uncommon cause of renal failure; sterile pyuria suggests papillary necrosis or renal tuberculosis.
- *Eosinophiluria* is strongly suggestive of allergic tubulointerstitial nephritis or cholesterol embolization.
- *Casts. Granular casts* are formed from abnormal cells within the tubular lumen, and indicate active renal disease. *Red-cell casts* are highly suggestive of glomerulonephritis.
- *Red cells in the urine* may be from anywhere between the glomerulus and the urethral meatus (Fig. 11.5).

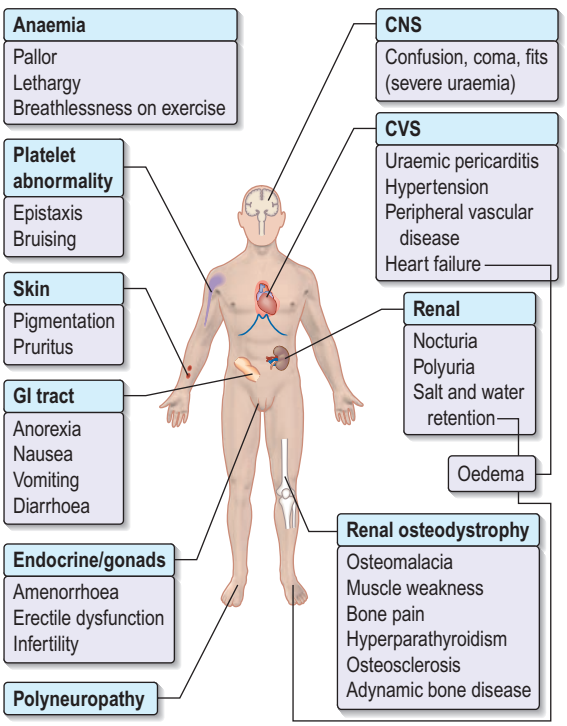


Fig. 11.46 **Symptoms and signs of chronic kidney disease.** Oedema may be due to a combination of primary renal salt and water retention and heart failure.

Urine biochemistry

- *24-hour creatinine clearance* is sometimes useful in assessing the severity of renal disease.
- *Measurements of urinary electrolytes* are unhelpful in chronic kidney disease. The use of urinary sodium concentration in the distinction between prerenal and intrinsic renal disease is discussed on page 619.
- *Urine osmolality* is a measure of concentrating ability. A low urine osmolality is normal in the presence of a high fluid intake but indicates renal disease when the kidney should be concentrating urine, such as in hypovolaemia or hypotension.
- *Urine electrophoresis and immunofixation* is necessary for the detection of light chains, which can be present without a detectable serum paraprotein.

Serum biochemistry

Urea and creatinine with calculation of eGFR. Electrophoresis and immunofixation for myeloma. Elevations of creatine kinase and a disproportionate elevation in serum creatinine and potassium compared to urea suggest rhabdomyolysis.

Haematology

- *Eosinophilia* suggests vasculitis, allergic tubulointerstitial nephritis, or cholesterol embolism.
- *Markedly raised viscosity* or ESR suggests myeloma or vasculitis.
- *Fragmented red cells and/or thombocytopenia* suggest intravascular haemolysis due to accelerated hypertension, haemolytic uraemic syndrome or thrombotic thrombocytopenic purpura.
- *Tests for sickle cell disease* should be performed when relevant.

Immunology

- *Complement components* may be low in active renal disease due to SLE, mesangiocapillary glomerulonephritis, post-streptococcal glomerulonephritis, and cryoglobulinaemia.
- *Autoantibody screening* is useful in detection of SLE (p. 541), scleroderma (p. 545), Wegener's granulomatosis and microscopic polyangiitis (p. 871), and Goodpasture's syndrome (p. 874).
- *Cryoglobulins* in unexplained glomerular disease, particularly mesangiocapillary glomerulonephritis.
- *Antibodies to streptococcal antigens* (antistreptolysin O titre (ASOT), anti-DNAase B) if post-streptococcal glomerulonephritis is possible.
- *Antibodies to hepatitis B and C* may point to polyarteritis or membranous nephropathy (hepatitis B) or to cryoglobulinaemic renal disease (hepatitis C).
- *Antibodies to HIV* raise the possibility of HIV-associated renal disease.

Radiological investigation

- *Ultrasound*. Every patient should undergo ultrasonography (for renal size and to exclude hydronephrosis), and plain abdominal radiography and CT (without contrast) to exclude low-density renal stones or nephrocalcinosis, which may be missed on ultrasound.
- *CT* is also useful for the diagnosis of retroperitoneal fibrosis and some other causes of urinary obstruction, and may also demonstrate cortical scarring.
- *MRI*. Magnetic resonance angiography in renovascular disease. For gadolinium used as contrast in renal failure, see page 580.
- *Intravenous urography* is seldom necessary in advanced renal disease.

Renal biopsy (see p. 582)

This should be performed in every patient with unexplained renal failure and normal-sized kidneys, unless there are strong contraindications. If rapidly progressive glomerulonephritis is possible, this investigation must be performed within 24 hours of presentation if at all possible.

COMPLICATIONS OF CHRONIC KIDNEY DISEASE

Anaemia

Several factors have been implicated:

- *erythropoietin deficiency* (the most significant)
- *bone marrow toxins* retained in renal failure
- *bone marrow fibrosis* secondary to hyperparathyroidism
- *haematinic deficiency* – iron, vitamin B_{12}, folate
- *increased red cell destruction*
- *abnormal red cell membranes* causing increased osmotic fragility
- *increased blood loss* – occult gastrointestinal bleeding, blood sampling, blood loss during haemodialysis or because of platelet dysfunction
- *ACE inhibitors* (may cause anaemia in CKD, probably by interfering with the control of endogenous erythropoietin release).

Red cell survival is reduced in renal failure. Increased red cell destruction may occur during haemodialysis owing to mechanical, oxidant and thermal damage.

Bone disease: renal osteodystrophy

The term 'renal osteodystrophy', more appropriately described as bone mineral disorder, embraces the various forms of bone disease that may develop alone or in combination in chronic kidney disease – hyperparathyroid bone disease, osteomalacia, osteoporosis, osteosclerosis and adynamic bone disease (Fig. 11.47). Most patients with chronic kidney disease are found, histologically, to have mixed bone disease. Covert renal osteodystrophy is present in many patients with moderate renal impairment and in almost all of those with end-stage renal failure.

Pathogenesis of bone disease

- Decreased renal production of the 1α-hydroxylase enzyme results in reduced conversion of 25-$(OH)_2D_3$ to the more metabolically active 1,25-$(OH)_2D_3$.

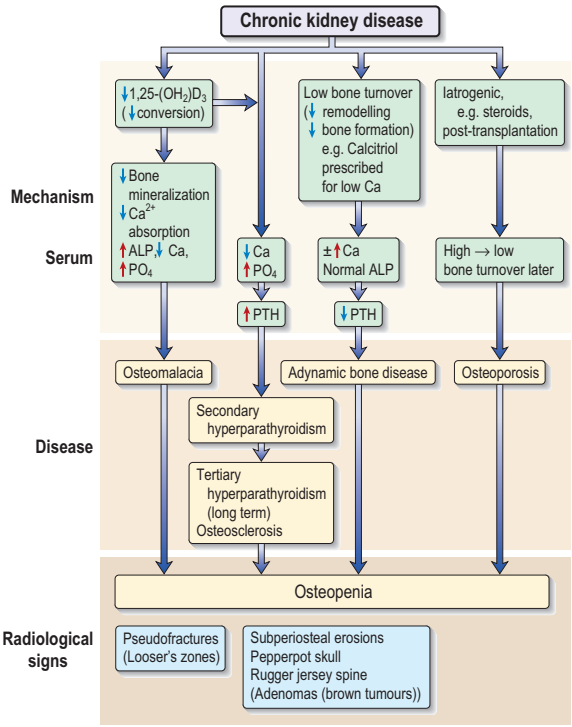


Fig. 11.47 **Renal osteodystrophy**: Pathogenesis and radiological features of renal bone disease. ALP, alkaline phosphatase.

- Reduced activation of vitamin D receptors (VDR) in the parathyroid glands leads to increased release of parathyroid hormone (*primary hyperparathyroidism*).
- The calcium sensing receptors (CaR), expressed in the parathyroid glands, react rapidly to acute changes in extracellular calcium concentrations and a low calcium also leads to increased release of PTH.
- 1,25-Dihydroxycholecalciferol deficiency also results in gut calcium malabsorption.
- Phosphate retention owing to reduced excretion by the kidneys, also indirectly by lowering ionized calcium (and probably directly via a putative but unrecognized phosphate receptor), results in an increase in PTH synthesis and release.
- PTH promotes reabsorption of calcium from bone and increased proximal renal tubular reabsorption of calcium, and this opposes the tendency to develop hypocalcaemia induced by 1,25-$(OH)_2D_3$ deficiency and phosphate retention. This 'secondary' hyperparathyroidism leads to increased osteoclastic activity, cyst formation and bone marrow fibrosis (osteitis fibrosa cystica).

Radiologically, digital subperiosteal erosions and 'pepperpot skull' are seen. Long-standing *secondary hyperparathyroidism* ultimately leads to hyperplasia of the glands with autonomous or '*tertiary*' *hyperparathyroidism* in which hypercalcaemia is present. Serum alkaline phosphatase concentration is raised in both secondary and tertiary hyperparathyroidism. Long-standing parathyroid hormone excess is also thought to cause increased bone density (osteosclerosis) seen particularly in the spine where alter-

nating bands of sclerotic and porotic bone give rise to a characteristic 'rugger jersey' appearance on X-ray.

$1,25\text{-}(OH)_2D_3$ deficiency and hypocalcaemia result in impaired mineralization of osteoid (osteomalacia). Such impaired mineralization also occurs when osteoblasts are inhibited by, for example, aluminium given as gut phosphorus binders, or accumulated in bone as a result of exposure to aluminium in source water used to make up dialysate for haemodialysis. In this situation, serum alkaline phosphatase concentration tends to be low or normal.

The condition of 'adynamic bone disease' in which both bone formation and resorption are depressed (in the absence of aluminium bone disease or overtreatment with vitamin D) is also seen. The pathogenesis of this condition is unclear and it is not known whether it leads to an increased risk of fractures or other complications. There may be hypercalcaemia; the serum alkaline phosphatase is normal, the PTH is low. X-rays and dual X-ray absorptiometry (DXA) scan show osteopenia. No treatment is of proven benefit.

Osteoporosis is commonly found in CKD, often after transplantation and the use of corticosteroids. Monitoring is with yearly DXA scan.

Management is discussed on page 563.

Skin disease

Pruritus (itching) is common in severe renal failure and is usually attributed in the main to retention of nitrogenous waste products of protein catabolism as it improves following the institution of dialysis. Other causes of pruritus include hypercalcaemia, hyperphosphataemia, elevated calcium × phosphate product, hyperparathyroidism (even if calcium and phosphate levels are normal) and iron deficiency.

In dialysis patients, inadequate dialysis is usually the cause of pruritus. Nevertheless, a significant number of dialysis patients who are well dialysed and in whom other causes of pruritus can be excluded suffer persistent itching. The cause is unknown and no effective treatment exists.

Many patients with renal failure suffer from dry skin for which simple aqueous creams are helpful. Eczematous lesions, particularly in relation to the region of an arteriovenous fistula, are relatively common. Chronic kidney disease may also cause porphyria cutanea tarda (PCT), a blistering photosensitive skin rash. This results from a decrease in hepatic uroporphyrinogen decarboxylase combined with a decreased clearance of porphyrins in the urine or by dialysis. Pseudoporphyria, a condition similar to PCT but without enzyme deficiency, is also seen in CKD with increased frequency.

Nephrogenic systemic fibrosis (NSF)

NSF is a systemic fibrosing disorder with predominant skin involvement and can simulate a variety of other disorders such as scleroderma, scleromyxoedema and eosinophilic fasciitis. It is seen only in patients with moderate to severe kidney failure (eGFR <30 mL/min), particularly patients on dialysis. Gadolinium-containing contrast agents, which are excreted exclusively by the kidney, have been implicated in the causation of over 95% cases of NSF. *Skin involvement* occurs universally with plaques, papules and/or nodules, the affected skin becoming thickened and firm and assuming a peau d'orange appearance. *Systemic involvement* occurs in a few patients with muscle stiffness, joint contracture, and fibrosis of the lungs, pleura, diaphragm, myocardium, pericardium and dura mater.

The *diagnosis* is based upon a biopsy of an involved site, showing proliferation of dermal fibrocytes with excessive collagen deposition. Special testing may show gadolinium.

NSF usually follows a chronic and unremitting course, with 30% having no improvement, 20% having modest improvement and 30% dying. No single therapy or combination of therapies has shown consistent benefit in NSF with exception of improvement in renal function. Although unproven, improvement in the NSF may follow renal transplantation. *Prevention* is by avoiding the use of gadolinium-based contrast agents in patents with severe renal impairment (eGFR <30 mL/min) or those on dialysis therapy.

Gastrointestinal complications

These include decreased gastric emptying and increased risk of reflux oesophagitis, peptic ulceration, acute pancreatitis and constipation, particularly in patients on continuous ambulatory peritoneal dialysis (CAPD).

Elevations of serum amylase of up to three times normal may be found in CKD without any evidence of pancreatic disease, owing to retention of high-molecular-weight forms of amylase normally excreted in the urine.

Metabolic abnormalities

Gout. Urate retention is a common feature of CKD. Treatment of clinical gout is complicated by the nephrotoxic potential of NSAIDs. Colchicine is useful for the acute attack, and allopurinol should be introduced under colchicine cover to prevent further attacks. The dose of allopurinol should be reduced in renal impairment, e.g. 100 mg on alternate days.

Insulin. Insulin is catabolized by and to some extent excreted via the kidneys. For this reason, insulin requirements in diabetic patients decrease as renal failure progresses. By contrast, end-organ resistance to insulin is a feature of advanced renal impairment resulting in modestly impaired glucose tolerance when a standard glucose tolerance test is carried out. Insulin resistance may contribute to hypertension and lipid abnormalities.

Lipid metabolism abnormalities. These are common in renal failure, and include:

- impaired clearance of triglyceride-rich particles
- hypercholesterolaemia (particularly in advanced renal failure).

The situation is further complicated in end-stage renal disease, when regular heparinization (in haemodialysis), excessive glucose absorption (in CAPD) and immunosuppressive drugs (in transplantation) may all contribute to lipid abnormalities. Correction of lipid abnormalities by, for example, HMG-CoA reductase inhibitor therapy (statins) is used in renal failure patients, although without formal proof of benefit derived from prospective controlled trials.

Endocrine abnormalities

These include:

- Hyperprolactinaemia, which may present with galactorrhoea in men as well as women.

- Increased luteinizing hormone (LH) levels in both sexes, and abnormal pulsatility of LH release.
- Decreased serum testosterone levels (only seldom below the normal level). Erectile dysfunction and decreased spermatogenesis are common.
- Absence of normal cyclical changes in female sex hormones, resulting in oligomenorrhoea or amenorrhoea.
- Complex abnormalities of growth hormone secretion and action, resulting in impaired growth in uraemic children (pharmacological treatment with recombinant growth hormone and insulin-like growth factor is used).
- Abnormal thyroid hormone levels, partly because of altered protein binding. Measurement of thyroid-stimulating hormone (TSH) is the best way to assess thyroid function. True hypothyroidism occurs with increased frequency in renal failure.

Posterior pituitary gland function is normal in renal failure.

Muscle dysfunction

Uraemia appears to interfere with muscle energy metabolism, but the mechanism is uncertain. Decreased physical fitness (cardiovascular deconditioning) also contributes.

Nervous system

Central nervous system

Severe uraemia causes an unusual combination of depressed cerebral function and decreased seizure threshold. However, convulsions in a uraemic patient are much more commonly due to other causes such as accelerated hypertension, thrombotic thrombocytopenic purpura or drug accumulation. Asterixis, tremor and myoclonus are also features of severe uraemia.

Rapid correction of severe uraemia by haemodialysis leads to '*dialysis disequilibrium*' owing to osmotic cerebral swelling. This can be avoided by correcting uraemia gradually by short, repeated haemodialysis treatments or by the use of peritoneal dialysis.

'*Dialysis dementia*' is a syndrome of progressive intellectual deterioration, speech disturbance, myoclonus and fits, which is due to aluminium intoxication; it may be accompanied by aluminium bone disease and by microcytic anaemia. Low-grade aluminium exposure may also cause more subtle, subclinical deterioration in intellectual function. Prevention involves removal of aluminium from source water used to manufacture dialysis fluid, and restriction or avoidance of aluminium-containing gut phosphorus binders. Treatment is with the chelating agent desferrioxamine.

Psychiatric problems are common. Patients can have anxiety, depression, phobias and psychoses.

Autonomic nervous system

Increased circulating catecholamine levels associated with down-regulation of α-receptors, impaired baroreceptor sensitivity and impaired efferent vagal function are common in renal impairment.

Overactivity of the sympathetic nervous system in chronic renal failure is believed to play a part in the genesis of hypertension in this condition. All of these abnormalities improve to some extent after institution of regular dialysis and resolve after successful renal transplantation.

Peripheral nervous system

Median nerve compression in the carpal tunnel is common, usually due to β_2-microglobulin-related amyloidosis. This can be avoided by haemofiltration and haemodiafiltration. 'Restless legs' syndrome (p. 1197) is common in uraemia. The syndrome is difficult to treat. Iron deficiency should be treated if present. Attention should be paid to adequacy of dialysis. Symptoms may improve with the correction of anaemia by erythropoietin. Clonazepam is sometimes useful. Renal transplantation cures the problem. A polyneuropathy occurs in patients who are inadequately dialysed.

Calciphylaxis

Also known as calcific uraemic arteriolopathy, this is a rare but serious life-threatening complication in CKD patients. It is increasingly recognized as a contributing factor to death in dialysis patients. Aetiological factors include reduced serum levels of a calcification inhibitory protein (fetuin-A) and abnormalities in smooth muscle cell biology in uraemia. It presents as painful non-healing eschars with panniculitis and dermal necrosis. The characteristic feature on histology is vascular calcification and superimposed small vessel thrombosis (Fig. 11.48).

Hyperparathyroidism and elevated concentrations of serum phosphate, morbid obesity and warfarin use remain consistent clinical features of most cases reported. Control of hyperparathyroidism is with either surgical intervention or with a calcimimetic agent.

Promising new treatment options include hyperbaric oxygen therapy and sodium thiosulfate infusion. Benefits from bisphosphonates and tissue plasminogen activator have also been reported.

Cardiovascular disease

Life expectancy remains severely reduced compared with the normal population owing to a greatly increased (16-fold) incidence of cardiovascular disease, particularly myocardial infarction, cardiac failure, sudden cardiac death and stroke.

Risk factors

Hypertension is a frequent complication of renal failure. Diabetes mellitus is the commonest cause of CKD. Dyslipidaemia is universal in uraemic patients. Furthermore, smoking is as common as in the general population and male gender is over-represented in patients with CKD. Ventricular hypertrophy is common, as is systolic and diastolic dysfunction. Diastolic dysfunction is largely attributable to left ventricular hypertrophy and contributes to hypotension during fluid removal on haemodialysis. Systolic dysfunction may be due to:

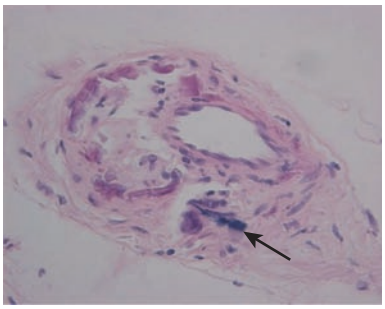


Fig. 11.48 **Calciphylaxis**: Characteristic arterial calcification of the skin microvasculature in the absence of vasculitic change (arrow).

- myocardial fibrosis
- abnormal myocyte function owing to uraemia
- calcium overload and hyperparathyroidism
- carnitine and selenium deficiency.

Left ventricular hypertrophy is a risk factor for early death in renal failure, as in the general population. Systolic dysfunction is also a marker for early death in renal failure.

Coronary artery calcification. Traditional risk factors (e.g. smoking, diabetes) can only partly explain the risk in patients with chronic nephropathies.

Coronary artery calcification is more common in patients with *end-stage renal failure* than in normal individuals and it is highly likely that this contributes significantly to cardiovascular mortality. Vascular calcification is frequent in all sizes of vessel in renal failure.

In addition to the classical risk factors for atherosclerosis:

- A *raised [calcium × phosphate] product* causes medial calcification.
- *Hyperparathyroidism* may also contribute independently to the pathogenesis by increasing intracellular calcium.
- *Vascular calcification* in uraemia is now thought to be an active process whereby vascular smooth muscle cells acquire osteoblast-like characteristics, possibly in response to elevated phosphate or [calcium × phosphate] product.
- *Inflammation* is a potent mediator of vascular calcification by inhibition of fetuin (a glycoprotein synthesized by liver, which is a potent inhibitor of vascular calcification).

The impact of vascular calcification is the *reduction of vascular compliance*, which manifests by increased pulse pressure and pulse wave velocity, and increased afterload contributes further to left ventricular hypertrophy. In addition to myocardial abnormalities, vascular calcification with its associated biomechanical vessel wall alterations is a strong predictor of all-cause and cardiovascular morbidity and mortality in patients with CKD. Diffuse calcification of the myocardium is also common; the causes are similar.

Other cardiovascular risk factors. These include hyperhomocysteinaemia, *Chlamydia pneumoniae* infection, oxidative stress and elevated endogenous inhibitor of nitric oxide synthase and asymmetric dimethyl arginine (ADMA) levels. High ADMA levels in uraemia are in part caused by oxidative stress and can possibly explain the 52% increase in the risk of death and 34% increase in the risk of cardiovascular events in uraemic patients. The use of antioxidants, vitamin E or *N*-acetylcysteine has been associated with a significant reduction in all-cause and cardiovascular mortality. However, recent trials to reduce levels of homocysteine with statins in dialysis patients with diabetes, and folic acid supplementation have been unsuccessful.

Pericarditis

This is common and occurs in two clinical settings:

- Uraemic pericarditis is a feature of severe, pre-terminal uraemia or of underdialysis. Haemorrhagic pericardial effusion and atrial arrhythmias are often associated. There is a danger of pericardial tamponade, and anticoagulants should be used with caution. Pericarditis usually resolves with intensive dialysis.
- Dialysis pericarditis occurs as a result of an intercurrent illness or surgery in a patient receiving apparently adequate dialysis.

Malignancy

The incidence of malignancy is raised in patients with CKD and with dialysis. Malignant change can occur in multicystic kidney disease. Lymphomas, primary liver cancer and thyroid cancers also occur.

Management

Successful renal transplantation improves some, but not all, of the complications of CKD, therefore attempts should be made to prevent these complications by careful monitoring with ECG, echocardiography, angiography (if necessary) and nuclear imaging. CT (spiral CT) and/or MRI are useful in the assessment of arterial calcification. Treatment is with the control of risk factors (p. 630) as well as the treatment of hypercalcaemia and hyperparathyroidism (p. 631).

Ask the authors
www.studentconsult.com

PROGRESSION OF CHRONIC RENAL IMPAIRMENT

Chronic renal impairment tends to progress inexorably to end-stage renal failure, although the rate of progression may depend upon the underlying nephropathy. Patients with chronic glomerular diseases tend to deteriorate more quickly than those with chronic tubulointerstitial nephropathies. Hypertension and heavy proteinuria are bad prognostic indicators. A non-specific renal scarring process common to renal disorders of different aetiologies may be responsible for progression.

Possible causes of glomerular scarring and proteinuria include:

- a rise in intraglomerular capillary pressure
- adaptive glomerular hypertrophy due to reduced arteriolar resistance and increased glomerular blood flow when there is reduced nephron mass.

This glomerular hyperfiltration, in response to nephron loss, was postulated as a common pathway for the progression of renal failure.

Since the afferent arteriolar tone decreases more than efferent arteriolar tone, intraglomerular pressure and the amount of filtrate formed by a single nephron rises. Angiotensin II produced locally modulates intraglomerular capillary pressure and GFR, predominantly causing vasoconstriction of postglomerular arterioles, thereby increasing the glomerular hydraulic pressure and filtration fraction (Fig. 11.49). In addition, by its effect on mesangial cells and podocytes, it increases the pore sizes and impairs the size-selective function of basement membrane for macromolecules.

Angiotensin II also modulates cell growth directly and indirectly by upregulating TGF-β, a potent fibrogenic cytokine, increases collagen synthesis and also causes epithelial cell transdifferentiation to myofibroblasts which contribute to excessive matrix formation. Furthermore, angiotensin II by upregulating plasminogen activator inhibitor-1 (PAI-1) inhibits matrix proteolysis by plasmin, resulting in accumulation of excessive matrix and scarring both in the glomeruli and interstitium.

Renal interstitial scarring is also a factor. The cause of this progressive interstitial damage and fibrosis is mutifactorial. In addition to non-haemodynamic effects of angiotensin II, *proteinuria* per se (by exposing tubular cells to albumin and its bound fatty acids and cytokines) promotes secretion of pro-inflammatory mediators, which promote interstitial inflammatory cell infiltrate and further augment fibrosis and progression of renal failure. The prognosis for renal function in chronic glomerular disorders is judged more accurately by interstitial histological appearances than by glomerular morphology.

Therapeutic manoeuvres (Box 11.6) aimed at inhibiting angiotensin II and reducing proteinuria mainly by ACEI and angiotensin-receptor antagonists have beneficial effects in slowing the rate of progression of renal failure in both diabetic and non-diabetic renal diseases in humans.

Experimentally, hepatocyte growth factor and bone morphogenic protein 7, both potent inhibitors of TGF-β, have resulted in the reversal of fibrosis and restoration of normal renal achitecture with improvement in renal function. If clinical trials are performed then nephrologists for the first time will be talking of *reversal* of CKD rather than *halting* the progression of renal impairment.

MANAGEMENT OF CHRONIC KIDNEY DISEASE

The underlying cause of renal disease should be treated aggressively wherever possible.

Renoprotection

The multidrug approach to chronic nephropathies has been formalized in an international protocol (Box 11.6).

Correction of complications

Hyperkalaemia

Hyperkalaemia often responds to dietary restriction of potassium intake. Drugs which cause potassium retention (see p. 667) should be stopped. Occasionally it may be necessary to prescribe ion-exchange resins to remove potassium in the gastrointestinal tract. Emergency treatment of severe hyperkalaemia is described on page 668.

Acidosis

Correction of acidosis helps to correct hyperkalaemia in CKD, and may also decrease muscle catabolism. Sodium bicarbonate supplements are often effective (4.8 g (57 mmol) of Na^+ and HCO_3^- daily), and the possibility of oedema and hypertension owing to the so-called 'extracellular fluid expansion' was not borne out by a recent trial. Calcium carbonate, also used as a calcium supplement and phosphate binder, has a beneficial effect on acidosis.

Calcium and phosphate control and suppression of PTH

Hypocalcaemia and hyperphosphataemia should be treated aggressively, preferably with regular (e.g. 3-monthly) measurements of serum PTH to assess how effectively hyperparathyroidism is being suppressed. Suppression of PTH levels to below two or three times the upper limit of 'normal'

Box 11.6 Renoprotection

Goals of treatment

BP <120/80
Proteinuria <0.3 g/24 hours

Treatment

Patients with chronic kidney disease and proteinuria >1 g/24 hours:

- ACE inhibitor increasing to maximum dose
- Add angiotensin receptor antagonist if goals are not achieved*
- Add diuretic to prevent hyperkalaemia and help to control BP
- Add calcium-channel blocker (verapamil or diltiazem) if goals not achieved

Additional measures

- Statins to lower cholesterol to <4.5 mmol/L
- Stop smoking (threefold higher rate of deterioration in CKD)
- Treat diabetes (HbA1c <7%)
- Normal protein diet (0.8–1 g/kg bodyweight)

*In type 2 diabetes start with angiotensin receptor antagonist

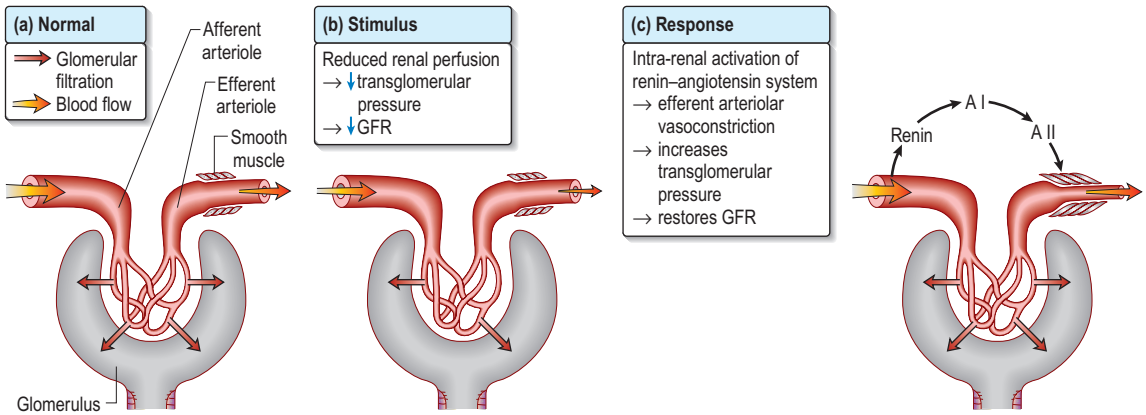


Fig. 11.49 **Glomerular dynamics**: Effect of the renin–angiotensin system. AI, angiotensin I; AII, angiotensin II.

carries a high risk of development of adynamic bone disease.

Dietary restriction of phosphate is seldom effective alone, because so many foods contain it. Oral calcium carbonate or acetate reduces absorption of dietary phosphate but is contraindicated where there is hypercalcaemia or hypercalciuria. Aluminium-containing gut phosphate binders are very effective but absorption of aluminium poses the risk of aluminium bone disease and development of cognitive impairment. They are now rarely used.

Treatment

- *Gut phosphate binders*. The polymer sevelamer reduces the calcium load and attenuates vascular calcification and also lowers cholesterol levels by 10%. However, it has not been shown to reduce mortality. Lanthanum carbonate is a new non-calcium, non-aluminium phosphate binder that is effective and has a good safety profile.
- *Nicotinamide*, an alternative to phosphate binders, blocks the intestinal sodium/inorganic phosphate (Na/Pi) cotransporter. Preliminary results show that it reduces phosphate levels and PTH levels alongside improvement in the lipid profile in dialysis patients.
- *Calcitriol or a vitamin D analogue,* such as alfacalcidol, used in early renal impairment has no deleterious effect upon renal function provided hypercalcaemia is avoided. New vitamin D metabolites (22-oxacalcitriol, paricalcitol, doxercalciferol) are less calcaemic. However, with the exception of paricalcitol (19-nor-1,25 dihydroxyvitamin D_2, which may have survival advantage), their usefulness over conventional but less expensive calcitriol or alfacalcidol remains to be established. (Treatment with vitamin D analogues should be started only if serum PTH level is three times or more above the upper limit of normal, in order to prevent the development of adynamic bone disease (see p. 628). Vitamin D therapy has the disadvantage that it increases not only calcium but also phosphate absorption and may therefore exacerbate hyperphosphataemia and ectopic calcification including calciphylaxis (calcification of small vessels)).
- *Calcimimetic agents* (e.g. cinacalcet, a calcium-sensing receptor agonist, see p. 628) have also been tried in established secondary hyperparathyroidism with successful suppression of PTH levels and lowering of calcium × phosphate product. The long-term safety and efficacy of these agents remains to be established.

Drug therapy

This should be minimized in patients with chronic renal impairment. Tetracyclines (with the possible exception of doxycycline) should be avoided in view of their anti-anabolic effect and tendency to worsen uraemia. Drugs excreted by the kidneys, such as gentamicin, should be prescribed with caution and drug levels monitored if feasible. Non-steroidal anti-inflammatory drugs (NSAIDs) should be avoided. Potassium-sparing agents, such as spironolactone and amiloride, pose particular dangers, as do artificial salt substitutes, all of which contain potassium.

Anaemia

The anaemia of erythropoietin (EPO) deficiency can be treated with synthetic (recombinant) human EPO (epoetin-alfa or -beta, or the longer-acting darbopoetin-alfa or polyethylene glycol-bound epoetin-beta). The intravenous route is used, initially with 50 U/kg of epoetin-alfa over 1–5 minutes, three times weekly. Subcutaneous administration of *epoetin-alfa* can also be used (see below). Blood pressure, haemoglobin concentration and reticulocyte count are measured every 2 weeks and the dose adjusted to maintain a target haemoglobin of 11–12 g/dL. Darbepoietin-alfa can be used once a week. Continuous erythropoiesis receptor activator (CERA), a recently licensed epoietin in Europe, is a pegylated epoietin beta and can be given once a month. In trials, CERA has a similar activity to other types of epoietins. However, CERA use is extremely limited and data about loss of flexibility of dosing, overshooting the target haemoglobin and its efficacy in combination with short-acting epoietins are required before its widespread use.

The target haemoglobin to be achieved during the treatment of anaemia is being revised to between 11 and 12 g/dL as studies in pre-dialysis CKD patients have not revealed any outcome benefits in patients who were treated to achieve higher haemoglobin targets (>12 g/dL).

Failure to respond to 300 U/kg weekly, or a fall in haemoglobin after a satisfactory response, may be due to iron deficiency, bleeding, malignancy, infection, inflammation or formation of anti-EPO neutralizing antibodies. The demand for iron by the bone marrow is enormous when erythropoietin is commenced. Patients on EPO therapy are regularly monitored for iron status and considered iron deficient if plasma ferritin is <100 μg/L, hypochromic RBCs >10%, transferrin saturation <20%. Functional iron deficiency due to poor mobilization of iron, despite adequate iron stores (ferritin >500 μg/L) and a transferrin saturation of >20%, is a common finding in patients with chronic inflammation. It is caused by hepcidin (see p. 394), an acute phase reactant produced by the liver in response to cytokines, particularly IL-6. Intravenous (rather than oral) iron supplements optimize response to EPO treatment by repletion of iron stores. A recent randomized trial has demonstrated a beneficial effect of i.v. iron even in patients with ferritin >800 μg/L and transferrin saturation of 20%.

Correction of anaemia with EPO improves quality of life, exercise tolerance and sexual and cognitive function in dialysis patients, and leads to regression of left ventricular hypertrophy. Avoidance of blood transfusion also reduces the chance of sensitization to HLA antigens, which may otherwise be a barrier to successful renal transplantation.

The *disadvantages of erythropoietin therapy* are that it is expensive and causes a rise in blood pressure in up to 30% of patients, particularly in the first 6 months. Peripheral resistance rises in all patients, owing to loss of hypoxic vasodilatation and to increased blood viscosity. A rare complication is encephalopathy with fits, transient cortical blindness and hypertension. Several reports of anti-EPO antibody-mediated pure red cell aplasia in patients receiving subcutaneous EPO therapy (particularly EPO alpha) have been described. The exact cause is unknown but interventions such as using the intravenous route and changes in manufacture of prefilled syringes have reduced the number of cases by 80%. Other causes of anaemia should be looked for and treated appropriately (see p. 392).

Several new erythropoiesis-stimulating agents are in clinical trials. An EPO mimetic is an engineered peptide which stimulates EPO receptor but still has to be given by injection. Oral agents which inhibit prolyl hydoxlase and prolong the

life of HIF (hypoxia inducible factors) 1 alpha, a transcription factor for endogenous production of EPO, have shown promise in phase 2 trials.

Male erectile dysfunction

Testosterone deficiency should be corrected. The oral phosphodiesterase inhibitors, e.g. sildenafil, tadalafil and vardenafil, are effective in end-stage renal failure and are the first line of therapy. The use of nitrates is a contraindication to this treatment. Other treatments are discussed on page 1000.

EARLY REFERRAL OF PATIENTS WITH CHRONIC KIDNEY DISEASE

Patients need time to adjust to the demands of chronic renal failure and its treatment, and to absorb information. Veins required in the future for fashioning of an arteriovenous fistula should not be rendered useless by cannulation (Fig. 11.50).

If the patient opts for regular haemodialysis, fashioning of an arteriovenous fistula should be carried out well in advance of the need for dialysis, when serum creatinine is of the order 400–500 μmol/L in non-diabetics and at an even earlier stage in diabetics with poorer vasculature. Such fistulae require several weeks to mature and become usable for vascular access.

RENAL REPLACEMENT THERAPY

Approximately 100 white individuals per million population commence renal replacement therapy in the UK each year. The corresponding figure in black Africans and Asians in the UK is three to four times higher, largely owing to diabetic and hypertensive nephropathy. The aim of all renal replacement techniques is to mimic the excretory functions of the normal kidney, including excretion of nitrogenous wastes, maintenance of normal electrolyte concentrations, and maintenance of a normal extracellular volume.

Haemodialysis

Basic principles

In haemodialysis, blood from the patient is pumped through an array of semipermeable membranes (the dialyser, often called an 'artificial kidney') which bring the blood into close contact with dialysate, flowing countercurrent to the blood. The plasma biochemistry changes towards that of the dialysate owing to diffusion of molecules down their concentration gradients (Fig. 11.51).

The dialysis machine comprises a series of blood pumps, with pressure monitors and bubble detectors and a proportionating unit, also with pressure monitors and blood leak detectors. Blood flow during dialysis is usually 200–300 mL per minute and the dialysate flow usually 500 mL per minute. The efficiency of dialysis in achieving biochemical change depends on blood and dialysate flow and the surface area of the dialysis membrane.

Dialysate is prepared by a proportionating unit, which mixes specially purified water with concentrate, resulting in fluid with the composition described in Table 11.21. Bicarbonate has replaced acetate which contributed to hypotension. Highly permeable synthetic membranes allow more rapid haemodialysis than cellulose-based membranes (high-flux haemodialysis).

Access for haemodialysis

Adequate dialysis requires a blood flow of at least 200 mL per minute. The most reliable long-term way of achieving this is surgical construction of an arteriovenous fistula (Fig. 11.52a, b), using the radial or brachial artery and the cephalic vein. This results in distension of the vein and thickening ('arterialization') of its wall, so that after 6–8 weeks large-bore needles may be inserted to take blood to and from the dialysis machine. In patients with poor-quality veins or arterial disease (e.g. diabetes mellitus) arteriovenous polytetrafluoroethylene (PTFE) grafts are used for access. However, these grafts have a very high incidence of thrombosis and 2-year graft patency is only 50–60%. Dipyridamole or fish oils

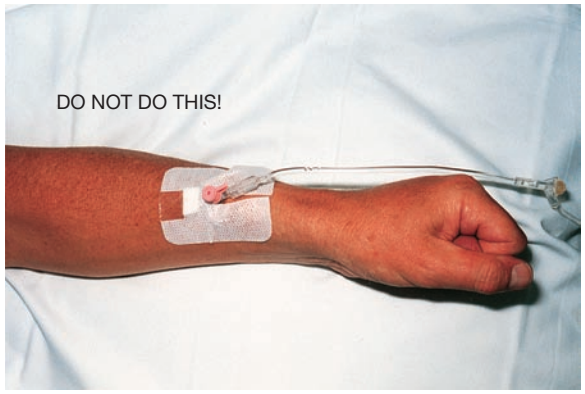


Fig. 11.50 **Intravenous cannula in exactly the WRONG place**, in a right-handed patient with chronic renal impairment who will in future need a left (non-dominant arm) radiocephalic fistula.

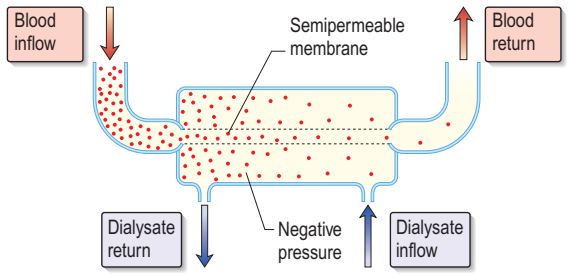


Fig. 11.51 **Changes across a semipermeable dialysis membrane.**

Table 11.21 **Range of concentrations (mmol/L) in routinely available final dialysates used for haemodialysis**

Sodium	130–145
Potassium	0.0–4.0
Calcium	1.0–1.6
Magnesium	0.25–0.85
Chloride	99–108
Bicarbonate	35–40
Glucose	0–10

(a)

(b)

Fig. 11.52 **Arteriovenous (radiocephalic) fistula. (a)** In the left forearm. **(b)** One way of reminding the anaesthetic, nursing and surgical team about the presence of an existing arteriovenous fistula in a haemodialysis patient about to undergo surgery. In particular, the fistula should not be compressed during surgery to avoid thrombosis within it.

improve graft patency but warfarin, aspirin and clopidogrel do not and are associated with a high incidence of complications.

If dialysis is needed immediately, a large-bore double-lumen cannula may be inserted into a central vein – usually the subclavian, jugular or femoral. *Semipermanent dual-lumen venous catheters* can also be used, usually inserted via a skin tunnel to lessen the risk of infection. Nevertheless, local and systemic sepsis is high, with increased morbidity and mortality. Stenosis of the subclavian vein is also common, and the jugular route is preferred.

Dialysis prescription

Dialysis must be tailored to an individual patient to obtain optimal results.

Dry weight

This is the weight at which a patient is neither fluid overloaded nor depleted. Patients are weighed at the start of each dialysis session and the transmembrane pressure adjusted to achieve fluid removal equal to the amount by which they exceed their dry weight.

The dialysate buffer

The dialysate buffer is usually acetate or bicarbonate. The sodium and calcium concentrations of the dialysate buffer are carefully monitored. A high dialysate sodium causes thirst and hypertension. A high dialysate calcium causes hypercalcaemia, whilst a low-calcium dialysate combined with poor compliance of medication with oral calcium carbonate and vitamin D may result in hyperparathyroidism.

Frequency and duration

Frequency and duration of dialysis are adjusted to achieve adequate removal of uraemic metabolites and to avoid excessive fluid overload between dialysis sessions. An adult of average size usually receives 4–5 hours' treatment three times a week. Twice-weekly dialysis is adequate only if the patient has considerable residual renal function.

Short-duration dialysis using very biocompatible high-flux membranes is commonly employed. Advantages include shorter duration of treatment and hence increased patient convenience. Disadvantages include higher cost of the membranes employed and, in all probability, higher prevalence of hypertension in such patients, requiring hypotensive medication. It should not be forgotten that normal kidneys work for 24 hours a day, 7 days a week, and that dialysis is a poor substitute for the natural state.

Adequate/optimal dialysis should be adjusted to individual patients' needs. All patients are anticoagulated (usually with heparin) during treatment as contact with foreign surfaces activates the clotting cascade. In the UK, a small number of patients manage self-supervised home haemodialysis.

Complications

Hypotension during dialysis is the major complication. Contributing factors include: an excessive removal of extracellular fluid, inadequate 'refilling' of the blood compartment from the interstitial compartment during fluid removal, abnormalities of venous tone, autonomic neuropathy, acetate intolerance (acetate acts as a vasodilator) and left ventricular hypertrophy.

Very rarely patients may develop anaphylactic reactions to ethylene oxide, which is used to sterilize most dialysers. Patients receiving ACE inhibitors are at risk of anaphylaxis if polyacrylonitrile dialysers are used.

Other potential, rare, complications include the hard-water syndrome (caused by failure to soften water resulting in a high calcium concentration prior to mixing with dialysate concentrate), haemolytic reactions and air embolism.

Adequacy of dialysis

Dialysis treatment is empirical since the size, number and nature of 'uraemic toxins' is unclear. The only true measure of adequacy is patient mortality and morbidity. Adequate nutrition of the patient as well as adequate dialysis is necessary to reduce morbidity and mortality.

Symptoms of underdialysis are non-specific and include insomnia, itching, fatigue despite adequate correction of anaemia, restless legs and a peripheral sensory neuropathy.

Adequacy of dialysis may be assessed by computerized calculation of urea kinetics, requiring measurement of the residual renal urea clearance, the rate of rise of urea concentration between dialysis sessions, and the reduction in urea concentration during dialysis. The dialysis dose is normally defined in terms of urea reduction ratio (URR) and/or equilibrated urea clearance, *eKt/V* (where *K* is the dialyser clearance, *t* is the duration of dialysis in minutes, and *V* is the urea distribution volume estimated as total body water). *Kt/V* of 1.0–1.2 and/or URR of 65% per dialysis session is the minimum threshold required for well-nourished dialysis patients dialysed three times per week. It is unclear whether a higher e*Kt/V* is associated with a better outcome, although no additional benefits of high (1.53) compared to standard

(1.16) e*Kt/V* were seen over 5 years' follow-up. It is likely that duration of haemodialysis session is a factor in itself in addition to the efficiency with which small molecules such as urea are cleared. The best haemodialysis outcome data are seen in renal units where long hours (8 hours per session) of dialysis are routinely practised.

Haemodialysis is the most efficient way of achieving rapid biochemical improvement, for instance in the treatment of acute kidney injury or severe hyperkalaemia. This advantage is offset by disadvantages such as haemodynamic instability, especially in acutely ill patients with multiorgan disease, and over-rapid correction of uraemia can lead to 'dialysis disequilibrium'. This is characterized by nausea and vomiting, restlessness, headache, hypertension, myoclonic jerking, and in severe instances seizures and coma owing to rapid changes in plasma osmolality leading to cerebral oedema.

These problems have led to the increasing adoption of gentler continuous methods for the treatment of acute kidney injury (see below).

Haemofiltration

This involves removal of plasma water and its dissolved constituents (e.g. K^+, Na^+, urea, phosphate) by convective flow across a high-flux semipermeable membrane, and replacing it with a solution of the desired biochemical composition (Fig. 11.53). Lactate is used as buffer in the replacement solution because rapid infusion of acetate causes vasodilatation, and bicarbonate may cause precipitation of calcium carbonate.

Haemofiltration can be used for both acute and chronic renal failure and is used in mainland Europe for CKD patients with haemodynamic instability. High volumes need to be exchanged in order to achieve adequate small molecule removal, typically a 22 L exchange three times a week for maintenance treatment and 1 L per hour in acute kidney injury. Financial costs of disposable items (such as filters and replacement fluid) are high and only a selected group of patients with end-stage renal failure are managed in this way. Modern dialysis machines have built-in facilities to generate online ultrapure water which has minimized the cost of the procedure and given an option to the clinician to use this technique either as haemofiltration or in combination with dialysis as haemodiafiltration to increase middle molecule clearance (e.g. β_2-microglobulin) and prevent long-term dialysis complications such as dialysis-related amyloidosis, particularly in young, highly sensitized, non-transplantable patients.

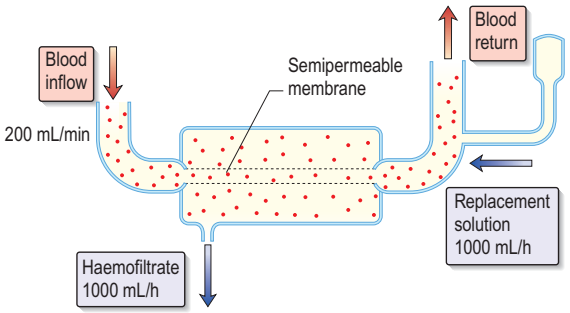


Fig. 11.53 **Principles of haemofiltration.**

Peritoneal dialysis

Peritoneal dialysis utilizes the peritoneal membrane as a semipermeable membrane, avoiding the need for extracorporeal circulation of blood. This is a very simple, low-technology treatment compared to haemodialysis. The principles are simple (Fig. 11.54).

- A tube is placed into the peritoneal cavity through the anterior abdominal wall (Fig. 11.55).
- Dialysate is run into the peritoneal cavity, usually under gravity.
- Urea, creatinine, phosphate and other uraemic toxins pass into the dialysate down their concentration gradients.
- Water (with solutes) is attracted into the peritoneal cavity by osmosis, depending on the osmolarity of the dialysate. This is determined by the glucose or polymer (icodextrin) content of the dialysate (Table 11.22).
- The fluid is changed regularly to repeat the process.

Chronic peritoneal dialysis requires insertion of a soft catheter, with its tip in the pelvis, exiting the peritoneal cavity in

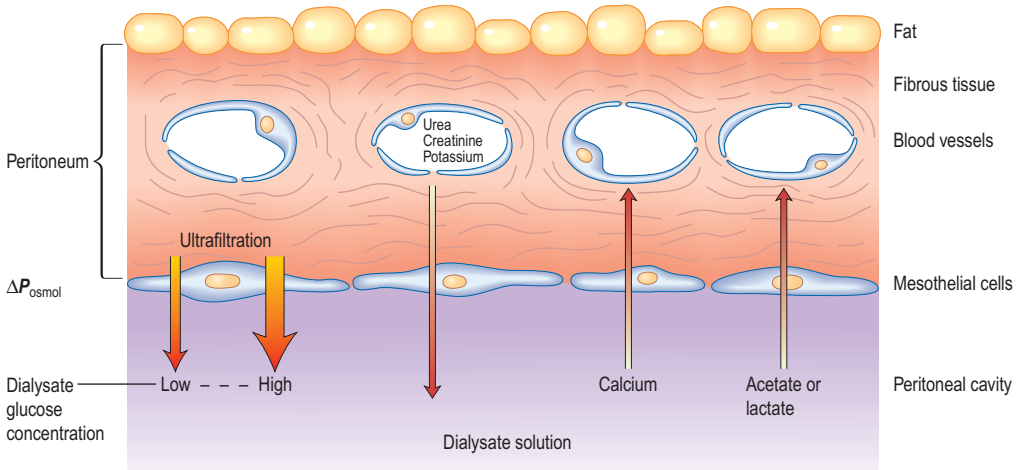


Fig. 11.54 **Principles of peritoneal dialysis.** Water is attracted into the peritoneal cavity depending on the osmolarity of the dialysate.

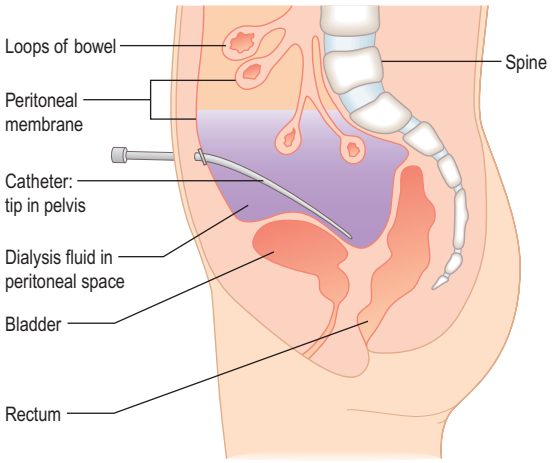


Fig. 11.55 **The siting of a Tenckhoff peritoneal dialysis catheter.**

Table 11.22 **Range of concentrations (mmol/L) in routinely available CAPD dialysate***

Sodium	130–134
Potassium	0
Calcium	1.0–1.75
Magnesium	0.25–0.75
Chloride	95–104
Lactate	35–40
Glucose	77–236
Total osmolality	356–511 mOsm/kg

*Glucose content is often expressed as g/dL of anhydrous glucose (e.g. 1.36% = 77 mmol/L). An even more hypertonic dialysate (6.36%) is available for acute (intermittent) peritoneal dialysis

the midline and lying in a skin tunnel with an exit site in the lateral abdominal wall (Fig. 11.55).

This form of dialysis can be adapted in several ways.

- *Continuous ambulatory peritoneal dialysis (CAPD).* Dialysate is present within the peritoneal cavity continuously, except when dialysate is being exchanged. Dialysate exchanges are performed three to five times a day, using a sterile no-touch technique to connect 1.5–3 L bags of dialysate to the peritoneal catheter; each exchange takes 20–40 minutes. This is the technique most often used for maintenance peritoneal dialysis in patients with end-stage renal failure.
- *Nightly intermittent peritoneal dialysis (NIPD),* also called automated peritoneal dialysis (APD). An automated device is used to perform exchanges each night while the patient is asleep. Sometimes dialysate is left in the peritoneal cavity during the day in addition, to increase the time during which biochemical exchange is occurring. Few trials have demonstrated superiority of APD over CAPD with regards to complications such as peritonitis, fluid status and in anuric patients.
- *Tidal dialysis.* A residual volume is left within the peritoneal cavity with continuous cycling of smaller volumes in and out.

Table 11.23 **Some causes of CAPD peritonitis***

	Approximate percentage of cases
Staphylococcus epidermidis	40–50
Escherichia coli, Pseudomonas and other Gram-negative organisms	25
Staphylococcus aureus	15
Mycobacterium tuberculosis	2
Candida and other fungal species	2

*In approximately 20%, no bacteria are found

Osmotic removal of excess plasma water and solutes is achieved using hypertonic dialysate (either due to high glucose concentration or icodextrin), which exerts an osmotic 'drag'. Depending on the patient's fluid intake and residual urine output, it may be necessary to use one or more hypertonic dialysate bags daily to achieve fluid balance in CAPD. Fluid overload is a relatively common problem in CAPD, and is due to failure of transport across the peritoneal membrane.

Complications

Peritonitis

Bacterial peritonitis is the most common serious complication of CAPD and other forms of peritoneal dialysis. Clinical presentations include abdominal pain of varying severity (guarding and rebound tenderness are unusual), and a cloudy peritoneal effluent, without which the diagnosis cannot be made. Microscopy reveals a neutrophil count above 100 cells per mL. Nausea, vomiting, fever and paralytic ileus may be seen if peritonitis is severe. The incidence of CAPD-associated peritonitis has been much reduced (to about one episode every two patient years) by use of a Y-disconnect system in preference to previous methods.

CAPD peritonitis must be investigated with culture of peritoneal effluent. Empirical antibiotic treatment is started, with a spectrum which covers both Gram-negative and Gram-positive organisms. Antibiotics may be given by the oral, intravenous or intraperitoneal route; most centres rely on intraperitoneal antibiotics. Common causative organisms are listed in Table 11.23.

Staph. aureus peritonitis should lead to a search for nasal carriage of this organism, and *Staph. epidermidis* peritonitis may indicate contamination from the patient's (or helper's) skin. Relapsing *Staph. epidermidis* peritonitis with an organism with the same antibiotic sensitivity pattern on each occasion may indicate that the Tenckhoff catheter has become colonized: often this is difficult to eradicate without replacement of the catheter under antibiotic cover.

Gram-negative peritonitis may complicate septicaemia from urinary or bowel infection. A mixed growth of Gram-negative and anaerobic organisms strongly suggests bowel perforation, and is an indication for laparotomy.

Fungal peritonitis often follows antibacterial treatment but may occur de novo. Clinical presentation is very variable. It is rare to be able to cure fungal peritonitis without catheter removal as well as antifungal treatment. Intraperitoneal amphotericin has been associated with the formation of peritoneal adhesions.

Infection around the catheter site

Infection where the catheter exits through the skin is relatively common. It should be treated aggressively (with systemic and/or local antibiotics) to prevent spread of the infection into the subcutaneous tunnel and the peritoneum. The most common causative organisms are staphylococci, including MRSA. Exit site infections can be reduced by routine use of mupirocin or gentamicin ointments locally and intranasally in those with colonized nares. Exit site infection due to *Pseudomonas* is treated by antibiotics and resiting the exit following catheter exchange.

Other complications

CAPD is often associated with constipation, which in turn may impair flow of dialysate in and out of the pelvis. Occasionally dialysate may leak through a diaphragmatic defect into the thoracic cavity, causing a massive pleural 'effusion'. The glucose content of the effusion is usually diagnostic, or the diagnosis may be made by instillation of methylthioninium chloride (methylene blue) with dialysate and the demonstration of a blue colour on pleural tap. Dialysate may also leak into the scrotum down a patent processus vaginalis.

Failure of peritoneal membrane function is a predictable complication of long-term CAPD, resulting in worsening biochemical exchange and decreased ultrafiltration with hypertonic dialysate. It is thought that this problem may be accelerated by excessive reliance on hypertonic dialysate to remove fluid.

Sclerosing peritonitis is a potentially fatal complication of CAPD. The cause is often unclear, but recurrent peritonitis and exposure of the peritoneum to unphysiological high glucose concentrations is responsible in most cases. Progressive thickening of the peritoneal membrane occurs in association with adhesions and strictures, turning the small bowel into a mass of matted loops and causing repeated episodes of small bowel obstruction. CAPD should be abandoned. Improvement may follow renal transplantation or treatment with prednisolone or azathioprine.

Contraindications

There are few absolute contraindications apart from unwillingness or inability on the patient's part to learn the technique.

- Previous peritonitis causing peritoneal adhesions may make peritoneal dialysis impossible, but the extent of adhesions is difficult to predict and it may be worth an attempted surgical placement of a dialysis catheter.
- The presence of a stoma (colostomy, ileostomy, ileal urinary conduit) makes successful placement of a dialysis catheter extremely unlikely.
- Active intra-abdominal sepsis, for instance due to diverticular abscesses, is an absolute contraindication to peritoneal dialysis although diverticular disease per se is not.
- Abdominal hernias may often expand during CAPD as a result of increased intra-abdominal pressure, and should ideally be repaired before or at the time of CAPD catheter insertion.
- Visual impairment may make it difficult for a patient to perform dialysate exchanges, but completely blind patients can be trained in the technique if adequately motivated.
- Severe arthritis makes it difficult to perform the exchanges, but a large number of mechanical aids are available. Sterilization of connections by heat or ultraviolet light reduces the risk of peritonitis.
- Some studies suggest that end-stage renal failure patients with coronary artery disease or congestive heart failure have significantly higher morbidity and mortality on CAPD compared to HD. Moreover, elderly patients do worse on CAPD than younger patients. It is advisable to consider HD as the first-choice dialysis modality in these patients where it is possible.

Adequacy of peritoneal dialysis

No consensus exists on the optimum degree of removal of urea and other waste products to be obtained in unit time but there is general agreement that minimum dose delivered should not be less than 1.7. Based on observational studies, weekly *Kt/V* of 2.0 coupled with a creatinine clearance target of 60 L per week had been recommended. However, in a randomized prospective study (ADEMEX), standard daily exchanges of 8 L in total (creatinine clearance 40 L per week) compared to a creatinine clearance of 60 L per week yielded similar patient and technique survival. Peritoneal dialysis inadequacy becomes common when residual renal function declines to zero. With increasing time on treatment, adequacy may become impaired owing to alterations in the efficiency of the peritoneal membrane in transporting waste products, fluid and electrolytes.

Complications of all long-term dialysis

Cardiovascular disease (see p. 630) and sepsis are the leading causes of death in long-term dialysis patients.

Causes of fatal sepsis include peritonitis complicating peritoneal dialysis and *Staph. aureus* infection (including endocarditis) complicating the use of indwelling access devices for haemodialysis.

Dialysis amyloidosis

This is the accumulation of amyloid protein (p. 1073) as a result of failure of clearance of β_2-microglobulin, a molecule of 11.8 kDa. This protein is the light chain of the class I HLA antigens and is normally freely filtered at the glomerulus but is not removed by cellulose-based haemodialysis membranes. The protein polymerizes, possibly after modification by non-enzymic glycosylation, to form amyloid deposits, which may cause median nerve compression in the carpal tunnel or a dialysis arthropathy – a clinical syndrome of pain and disabling stiffness in the shoulders, hips, hands, wrists and knees. β_2-microglobulin-related amyloid may be demonstrated in the synovium. Amyloid deposits (see p. 1072) can also cause pathological bone cysts and fractures, pseudotumours and gastrointestinal bleeding caused by amyloid deposition around submucosal blood vessels. The extent of amyloid deposition is best assessed by nuclear imaging, either using [^{99m}Tc]DMSA, or, more specifically, by the use of radiolabelled serum amyloid P component.

Rapid improvement after renal transplantation is probably due to steroid therapy as low-dose prednisolone alone can also cause an improvement. A change to a biocompatible synthetic membrane has also been reported to be of benefit: again, the mechanism for this improvement is not clear. β_2-microglobulin clearance is several times higher in patients treated with haemofiltration or haemodiafiltration and these

techniques are increasingly used in patients at high risk of developing this complication.

Transplantation

Successful renal transplantation offers the potential for almost complete rehabilitation in end-stage renal failure. This mode of renal replacement therapy has significant survival advantage compared to dialysis patients on transplant waiting lists. It allows freedom from dietary and fluid restriction; anaemia and infertility are corrected; and the need for parathyroidectomy is reduced. It is the treatment of choice for most patients with end-stage renal failure. The supply of donor organs (in the UK, 30/million/year) is greatly exceeded by demand (48/million/year), and donor organs are therefore scarce and a valuable resource that must be used optimally.

The technique involves the anastomosis of an explanted human kidney, usually either from a cadaveric donor or from a living close relative, on to the iliac vessels of the recipient (Fig. 11.56). The donor ureter is placed into the recipient's bladder. Unless the donor is genetically identical (i.e. an identical twin), immunosuppressive treatment is needed, for as long as the transplant remains in place, to prevent rejection. Refinements in patient selection and assessment of donor–recipient compatibility, improvements in surgical techniques and the development of more efficient immunosuppressive regimens have increased patient and graft survival. Eighty per cent of grafts now survive for 5–10 years in the best centres, and 50% for 10–30 years. However the half-life of renal allografts is still 13–16 years. The three most common causes of late graft loss are death with functioning graft, recurrence of renal disease and chronic allograft nephropathy.

Factors affecting success

ABO (blood group) compatibility between donor and recipient is required. However, in some centres, particularly in Japan where a cadaveric organ donation programme is not well established, ABO-incompatible renal transplants are increasingly performed. These transplants follow with immunoadsorption to remove preformed antibodies, splenectomy, anti-CD20 antibodies to remove B lymphocytes, and intravenous pooled immunoglobulins for immunomodulation or anti-idiotypic antibodies. In experienced hands, results are acceptable.

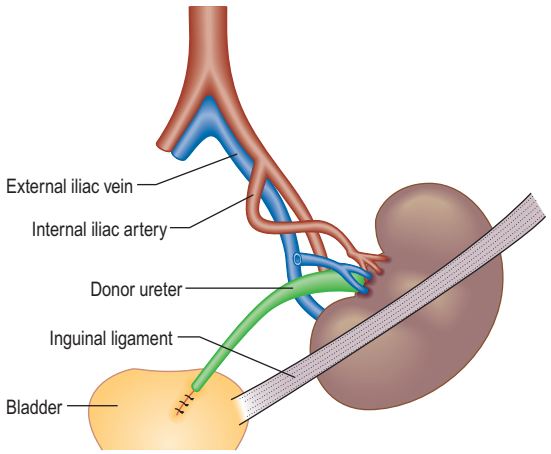


Fig. 11.56 **Anatomy of a renal transplant operation.**

Matching donor and recipient for HLA type

Matching for HLA-DR antigens appears to have the most impact on graft survival. Studies have shown that matching at the HLA-B locus has only a minor effect on graft outcomes. Complete compatibility at A, B and DR offers the best chance of success, followed by a single HLA mismatch (i.e. antigen possessed by the donor and not possessed by the recipient). The effect of further degrees of mismatching upon graft survival in first transplants is of modest degree. Nationwide matching schemes for kidneys retrieved from cadaver donors are in existence. However, with availability of more efficient immunosuppressive agents, the value of HLA matching in the overall transplant outcome is, if any, only modest. Transplantation with completely mismatched kidneys, particularly when the donor is the patient's partner, is routinely practised and results are as good as, if not better than, properly matched cadaveric kidneys.

A newly identified donor factor, an allotype of the C3 complement molecule, may be associated with better long-term outcomes for cadaveric kidney grafts. The C3 allotypes have distinct electrophoretic migration patterns: C3F (fast), present in about 20% of white patients, and C3S (slow). This report raises the possibility that the presence of the C3F allele in a kidney allograft is associated with a better outcome.

Adequate immunosuppressive treatment

See below.

The donor kidney

Cadaveric donation

Most countries allow the removal of kidneys and other organs from patients who have suffered irretrievable brain damage ('brainstem death') while their hearts are still beating (see p. 920).

Living related donation

A close relative may volunteer as a potential donor. A sibling donor may be HLA identical or share one or no haplotypes with the potential recipient. In the UK, donor age must be 18 years or more.

Potential living related donors are subjected to an intensive preoperative evaluation, including clinical examination and measurement of renal function, tests for carriage of hepatitis B, C, HIV and cytomegalovirus, and detailed imaging of renal anatomy, to be sure that transplantation will be technically feasible.

Unrelated living donors may be accepted provided no inducement (financial or otherwise) is involved. Paid live non-related donor transplantation is illegal in the UK.

Immunosuppression for transplantation

Long-term drug treatment for the prevention of rejection is employed in all cases apart from living related donation from an identical twin. Some degree of immunological tolerance does develop, and the risk of rejection is highest in the first 3 months after transplantation. In the early months rejection

episodes occur in less than 30% of cadaver kidney recipients. Most are reversible. A combination of immunosuppressive drugs is usually used.

Corticosteroids. Corticosteroids have a non-specific anti-inflammatory (inhibition of phospholipase A_2 and arachidonic acid cascade) and immunosuppressive action (inhibition of gamma-interferon-dependent adhesion molecules and thereby dendritic cell and T lymphocyte interaction). High-dose methylprednisolone is used as the primary treatment for acute rejection.

Azathioprine. Azathioprine and its metabolite 6-mercaptopurine prevent cell-mediated rejection by blocking purine synthesis and replication of lymphocytes. They can convert a CD28-dependent costimulatory Rac1 signal (proliferation and differentiation) upon dendritic cells and T-lymphocyte interaction into an apoptotic response by their ability to form 6-thio-GTP. 6-Thio-GTP competes with GTP for Rac1, preventing its anti-apoptotic action. Adverse effects (see p. 290) include suppression of red cell and platelet production, an increased incidence of infections (particularly viral), and hepatotoxicity. It also interacts with allopurinol, which increases levels of its active metabolite (6-mercaptopurine), resulting in toxicity.

Mycophenolate mofetil is metabolized to mycophenolic acid and blocks inosine monophosphate dehydrogenase (IMPD), an enzyme in the de novo synthetic pathway of purine synthesis. It may be more specific for lymphocytes because, unlike in other cells, an alternative purine synthetic salvage pathway is absent from lymphocytes. It is more potent than azathioprine, as trials have shown a reduction in rejection episodes. Gastrointestinal symptoms such as nausea, vomiting and diarrhoea are more common than with azathioprine but the drug can be used with allopurinol.

Ciclosporin (CSA). Ciclosporin, a calcineurin inhibitor, prevents the activation of T lymphocytes in response to new antigens and is highly effective in preventing rejection, while leaving the functioning of the rest of the immune system largely intact. Intracellularly, CSA forms complexes with cyclophilin (isomerase) and inhibits calmodulin–calcineurin-induced phosphorylation of NF-AT transcription factor for IL-2 and possibly other T-cell activation genes. Its introduction has revolutionized organ transplantation. Disadvantages include high cost and nephrotoxicity. Even with careful adjustment of the dose in response to trough blood levels, renal function may be adversely affected.

Tacrolimus, also a calcineurin inhibitor, blocks T-cell activation by a mechanism very similar to that of ciclosporin with the exception that tacrolimus forms a complex with immunophilin FK binding protein-12 (FKBP-12) instead of cyclophilin. The downstream actions are the same as with CSA. It is more potent than CSA and is therefore used both as a rescue agent for treating rejection and as a maintenance agent. Its side-effects are broadly similar to those of CSA except it is more diabetogenic.

Sirolimus and the structural analogue everolimus are immunosuppressants that are synergistic with CSA but lack nephrotoxicity. Sirolimus, like tacrolimus, binds to FKBP (particularly the -25 rather than the -12 isoform) but instead of calcineurin inhibition, acts on targets of rapamycin signalling (mammalian target of rapamycin – mTOR) and prevents cell entry into G1 phase of the cell cycle. It may lead to poor wound healing and delay the recovery of proximal tubular cell injury sustained during transplantation. Its main side-effects are thrombocytopenia and hyperlipidaemia. Its role in renal transplantation is being explored.

Antibodies. Both polyclonal anti-thymocytic globulin (ATG) and anti-lymphocytic globulin (ALG) and monoclonal (OKT3) and anti-CD52 (CAMPATH) antibodies are potent immunosuppressive agents. Antibodies may be derived from mouse, rabbit, horse, or 'humanized', and directed against any of a number of lymphocyte surface marker proteins, enabling neutralization or killing of lymphocytes with certain functions (e.g. T cells, activated T cells, cells expressing adhesion molecules). They are mainly used for the treatment of steroid-resistant rejections and as an induction therapy for high immunological risk patients (previously sensitized with circulating anti-HLA antibodies). Basiliximab (chimeric) and dacluzumab (humanized) anti-CD25 monoclonal antibodies bind to IL-2 receptors on activated T lymphocytes only (therefore having few side-effects), inhibiting IL-2-driven proliferative responses. They have largely replaced ATG and ALG in transplant induction protocols of patients with low or intermediate immunological risk for rejection.

A new anti-B7 monoclonal antibody, LEA29Y (belatacept), given by a monthly injection and lacking the nephrotoxicity of other calcineurin inhibitors, has been found to be as effective as ciclosporin. It prevents T-cell activation induced by CD28.

Complications

Acute tubular necrosis (ATN)

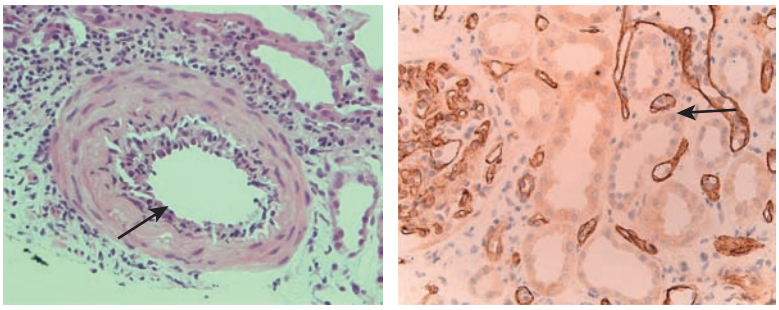
ATN is the commonest cause of cadaveric graft dysfunction (up to 40–50%), particularly after asystolic donation and prolonged cold ischaemia time (>24 hours). Kidneys from elderly (>55 years) donors and those with a history of hypertension are liable to develop ATN. Delayed graft function due to ATN is associated with worse long-term outcome and also predisposes the graft to rejection. It is usual practice to use induction with antibodies and use 50% of the starting dose of calcineurin inhibitors in recipient of grafts at high risk for ATN.

Technical failures

There may be occlusion or stenosis of the arterial anastomosis, occlusion of the venous anastomosis, and urinary leaks owing to damage to the lower ureter, or defects in the anastomosis between ureter and recipient bladder. If urine output drops, Doppler ultrasonography, DTPA scanning and/or renal angiography is used. Surgical re-implantation may be required.

Acute rejection (AR)

AR is seen in between 10 and 30% of transplant recipients and usually presents with declining renal function within the first 3 months. Renal biopsy confirms the diagnosis and also assesses the severity and type of rejection (cellular or vascular or antibody mediated, also called humoral rejection). Therapy in cellular rejection is high-dose pulse steroid; in acute vascular rejection (diagnosed by the presence of additional endothelial inflammation), ATG, ALG or OKT3 is used.



(a) (b)

Fig. 11.57 **Histology of kidney in rejection. (a) Acute vascular rejection**: The mononuclear inflammatory cell infiltrate is limited to the expanded intima and does not involve the entire vascular wall as in systemic vasculitis. **(b) Antibody mediated rejection**: An acute tubular necrosis with capillary glomerulitis and peritubular and glomerular capillary C4d positivity.

Histological appearances similar to vascular rejection are seen in a small number of patients where the culprit is an agonistic antibody directed against angiotensin receptor 1 (ATR 1) which usually presents with sudden decline in renal function but unlike vascular rejection also manifests with malignant phase hypertension and seizures. It responds to angiotensin receptor blockers in addition to conventional antirejection strategies (Fig. 11.57a). Humoral rejection (diagnosed by the presence of circulating donor-specific anti-HLA antibodies and evidence of complement activation on renal biopsy by C4 staining) (Fig. 11.57b) is usually treated empirically by a combination of i.v. polyclonal immunoglobulin infusion to neutralize and promote the clearance of anti-HLA antibodies, plasmapheresis (to remove antibodies) and anti-CD20 antibody administration (to deplete B lymphocytes) with variable success.

More than one rejection within the first 3 months, vascular and/or humoral rejection, delayed rejection (requirement of dialysis within the first week after transplantation), and failure of serum creatinine to return to baseline (less than 130 μmol/L) are associated with worse long-term outcome. The outcome of renal transplantation during an episode of acute rejection is difficult to predict, even with an allograft biopsy. Urinary levels of messenger RNA (mRNA) for FOXP3, a specification and functional factor for regulatory T lymphocytes, improve the prediction of outcome of acute rejection of renal transplants by a noninvasive means. It predicts with a reasonable degree of sensitivity and specificity successful outcome or otherwise in a renal transplant with acute rejections.

Infections

In the first month post transplantation, infections tend to be from bacterial sources seen typically in the surgical population. Cytomegalovirus (CMV) infections develop weeks or months after transplantation in 70% of CMV-seronegative recipients receiving grafts from a seropositive donor and in patients receiving biological agents (antibodies) as induction or therapy for rejection, unless prophylaxis with valganciclovir or valaciclovir is given. Prophylaxis is also routinely given against *Pneumocystis jiroveci* (co-trimoxazole) and oral candidiasis (nystatin or amphotericin lozenges). Polyomavirus infections (*BK nephropathy*) result in graft dysfunction and eventual loss due to mainly tubulointerstitial nephritis diagnosed by renal biopsy and presence of cellular inclusion bodies positive for SV40 (Fig. 11.58). There is no known specific treatment with the exception of tapering immunosuppression and with cautious use of leflunomide and cidofovir.

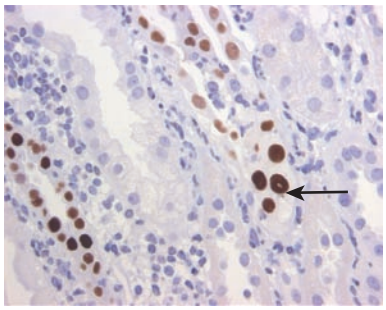


Fig. 11.58 **BK nephropathy** showing the characteristic intranuclear viral inclusion with positive immunohistological stain for SV40.

Post-transplantation lymphoproliferative disorders

Epstein–Barr virus-associated malignancies are common in patients who received biological agents and in children. Tapering of ciclosporin or tacrolimus and monitoring for the reappearance of cytotoxic lymphocytes has improved the outcome.

Chronic allograft nephropathy (CAN)

CAN remains the most common cause of late graft failure. The process is mediated by immunological and non-immunological factors and results in a progressive irreversible decline in graft function with mild to modest proteinuria (<3 g/day). Unfortunately there is no established therapy of proven efficiency.

Malignancy

Immunosuppressive therapy increases the risk of skin tumours, including basal and squamous cell carcinoma. In white recipients, exposure to ultraviolet light should be minimized and sun-block creams employed. Other common cancers are renal, cervical and vaginal. In female recipients, regular yearly cervical smears should be carried out.

Cardiovascular disease

Cardiovascular disease is the cause of death post transplantation in 50% of cases. This is due to increased incidence of hypertension, obesity, diabetes and insulin resistance lipid disorders. Use of a statin (fluvastatin) was associated with reduction in cardiac end-points (myocardial infarction or revascularization) post transplantation but overall mortality remained unchanged in a randomized prospective study.

Post-transplant osteoporosis

This is common following transplantation owing to treatment with steroids. Maximum bone loss occurs within the first 3 months and regular DXA scans are necessary. Bisphosphonates (alendronate, pamidronate), and alfacalcidol with or without calcium carbonate have proven to be effective in control studies.

Recurrent disease

Recurrence of renal disease is surprisingly common. Primary FSGS often recurs and causes early graft loss. Mesangiocapillary GN, diabetic nephropathy and IgA nephropathy also commonly recur but seldom cause renal insufficiency.

FURTHER READING

Bakris GL. Slowing nephropathy progression: focus on proteinuria reduction. *Clinical Journal of American Society of Nephrology* 2008; **3** Suppl 1: S3–10.

Ekberg H, Tedesco-Silva H, Demirbas A et al. Reduced exposure to calcineurin inhibitors in renal transplantation. *New England Journal of Medicine* 2007; **347**: 2562–2575.

Meguid EL, Nahas A, Bello AK. Chronic kidney disease: the global challenge. *Lancet* 2005; **365**: 331–340.

Meyer TW, Hostetter TH. Uraemia. *New England Journal of Medicine* 2007; **357**: 1316–1325.

Zoccali C. Traditional and emerging cardiovascular and renal risk factors: an epidemiologic perspective. *Kidney International* 2006; **70**: 26–33.

CYSTIC RENAL DISEASE

Solitary or multiple renal cysts are common, especially with advancing age: 50% of those aged 50 years or more have one or more such cysts. They have no special significance except in the differential diagnosis of renal tumours (see p. 644). Such cysts are often asymptomatic and are found on ultrasound examination performed for some other reason. Occasionally they may cause pain and/or haematuria owing to their large size, or bleeding may occur into the cyst. Cystic degeneration (the formation of multiple cysts which enlarge with time) occurs regularly in the kidneys of patients with end-stage renal failure treated by dialysis and/or transplantation. Malignant tumour formation seems to be more common in such kidneys than in the general population.

Autosomal-dominant polycystic kidney disease

Autosomal-dominant polycystic kidney disease (ADPKD) is an inherited disorder usually presenting in adult life. It is characterized by the development of multiple renal cysts, variably associated with extrarenal (mainly hepatic and cardiovascular) abnormalities. ADPKD is by far the most common inherited nephropathy, with a prevalence rate ranging from 1:400 to 1:1000 in white populations. It accounts for 3–10% of all patients commencing regular dialysis in the West.

In about 85% of cases, the gene responsible (*PKD1*) has been located on chromosome 16. A second gene, *PKD2*, which has been mapped to chromosome 4, accounts for the vast majority of other cases. These genetic abnormalities are distinct from the autosomal recessive form of polycystic disease (due to mutations in the *PKHD1* gene on chromosome 6p21.1–p12), which is often lethal in early life.

The protein corresponding to the *PKD1* gene, polycystin 1, is an integral membrane glycoprotein involved in cell-to-cell and/or cell-to-matrix interaction and functions as a mechanosensor. The protein corresponding to the *PKD2* gene appears to function as a calcium ion channel, regulating calcium influx and/or release from intracellular stores. Polycystin-1 acts as the regulator of PKD2 channel activity by its co-localization on cilia of collecting tubular cells. Disruption of the polycystin pathway results in reduced cytoplasmic calcium, which in principal cells (p. 653) of the collecting duct causes an increase in cAMP via stimulation of calcium-inhibitable adenyl cyclase and inhibition of cAMP phosphodiesterases. Defective ciliary signalling results in disoriented division of the cells along the nephron, resulting in cyst formation. In ADPKD, <1% of all nephrons acquire a second somatic mutation that in combination with a germline mutation results in cystogenesis. Progressive loss of renal function is usually attributed to mechanical compression, apoptosis of the healthy tissue and reactive fibrosis. Patients with ADPKD experience declining renal function at a variable rate which is due to discrepancy in the growth and size of the cysts; patients with rapid growth in cyst size as determined by MRI lose renal function more rapidly. Strategies to slow the growth rate of cysts have been very effective in preserving renal function in animal models. These therapies include the vasopressin V_2 receptor inhibitor (to reduce cAMP in the principal cells), roscovitine (a cyclin-dependent kinase inhibitor) and antiproliferative therapy with sirolimus (mTOR inhibitor).

Clinical features

Clinical presentation may be at any age from the second decade. Presenting symptoms include:

- acute loin pain and/or haematuria owing to haemorrhage into a cyst, cyst infection or urinary tract stone formation
- loin or abdominal discomfort owing to the increasing size of the kidneys
- subarachnoid haemorrhage associated with berry aneurysm rupture
- complications of hypertension
- complications of associated liver cysts
- symptoms of uraemia and/or anaemia associated with chronic renal failure.

Erythraemia is a rare complication and presentation of ADPKD.

The natural history of the disease is one of progressive renal impairment, sometimes punctuated by acute episodes of loin pain and haematuria, and commonly associated with the development of hypertension. The rate of progression to

renal failure (see above) is variable. The determinants of progression are both genetic and non-genetic. In the PKD2 form, renal cysts develop more slowly and end-stage renal failure (ESRF) occurs 10–15 years later than in the PKD1 form. Gender affects renal prognosis. Males with ADPKD reach end-stage renal failure 5–6 years earlier than females. There is a large variability in the age at ESRF within families, even between affected monozygotic twins.

Complications and associations

Pain

A minority of patients suffer chronic renal pain resistant to common analgesics, presumably owing to the pressure effect of large cysts. Surgical decompression of such cysts appears to be of benefit in about two-thirds of patients. Laparoscopic cyst decortication is a minimally invasive alternative technique.

Cyst infection

The response to standard antibacterial therapy is often poor owing to poor penetration of conventional antibiotics across the cyst wall. Lipophilic antibiotics active against Gram-negative bacteria, such as co-trimoxazole and fluoroquinolones, penetrate into the cysts better and their use has greatly improved the treatment of this complication.

Renal calculi

These are diagnosed in about 10–20% of patients with ADPKD. Frequently they are composed of uric acid and hence radiolucent (see Fig. 11.40). Obstructing or painful stones are treated no differently than are stones in patients with normal urinary tracts. Percutaneous stone removal and extracorporeal lithotripsy is safe.

Hypertension

Hypertension is an early and very common feature of ADPKD. Elevation of blood pressure, still within the normal range, is detectable in young affected individuals and is associated with an increase in left ventricular mass. Left ventricular hypertrophy occurs to a greater degree for a given rise in blood pressure in ADPKD compared with other renal disorders and with essential hypertension. Intrarenal activation of the renin–angiotensin system is involved in pathogenesis, and ACE inhibitors are logical first-line agents in treatment. Early control of blood pressure is essential as cardiovascular complications are a major cause of death in ADPKD.

Progressive renal failure

This is the most serious complication of ADPKD. At glomerular filtration rates below 50 mL/min, the rate of decline in GFR averages 5 mL/min each year, which is more rapid than in other primary renal disorders. The probability of being alive without requiring dialysis or transplantation by the age of 70 years is of the order of 30%. Survival rates on regular haemodialysis and after renal transplantation in ADPKD are better than those in patients with other primary renal diseases.

Hepatic cysts

Approximately 30% of patients have hepatic cysts and in a minority of the patients massive enlargement of the polycystic liver is seen. Pain, infection of cysts and, more rarely, compression of the bile duct, portal vein or hepatic venous outflow occur. Rarely, percutaneous drainage of painful cysts, laparoscopic fenestration or even partial hepatectomy is necessary. Infected cysts may require drainage.

Intracranial aneurysm formation

About 10% of ADPKD patients have an asymptomatic intracranial aneurysm (see p. 1134) and the prevalence is twice as high in the subgroup of patients with a family history of such aneurysms or of subarachnoid haemorrhage. Such haemorrhage is preceded in from 20% to 40% of cases by premonitory headaches from a few hours up to 2 weeks before the onset of subarachnoid bleeding. Headache of sudden onset or unusual character or severity in a patient with ADPKD should prompt investigation. Contrast-enhanced spiral CT or MR angiography are the best investigations. Screening for intracranial aneurysm in ADPKD is currently recommended for patients aged 18–40 years who have a positive family history.

Mitral valve prolapse

This is found in 20% of individuals with ADPKD.

Diagnosis

Physical examination commonly reveals large, irregular kidneys and possibly hepatomegaly. Definitive diagnosis is established by ultrasound examination (Fig. 11.59). However, such renal imaging techniques may be equivocal, especially in subjects under the age of 20 years.

Screening

The children and siblings of patients with established ADPKD should, in general, be offered screening. Affected individuals should have regular blood pressure checks and should be offered genetic counselling. Screening by ultrasonography should not be carried out before the age of 20 years, as excluding the condition may be difficult and hypertension is unusual before this age. Even at age 20, renal ultrasonography may give a false-negative result. Gene linkage analysis can be utilized in many families.

Therapy

No therapy is yet available but potential agents include *vasopressin receptor antagonists*, which have been studied in animal models of polycystic kidney disease, closely related to human ADPKD, where they have halted cyst progression or caused disease regression. ADPKD kidneys may be particularly vulnerable to *adenylyl cyclase agonists* or *cAMP phosphodiesterase inhibitors*. Caffeine at clinically relevant

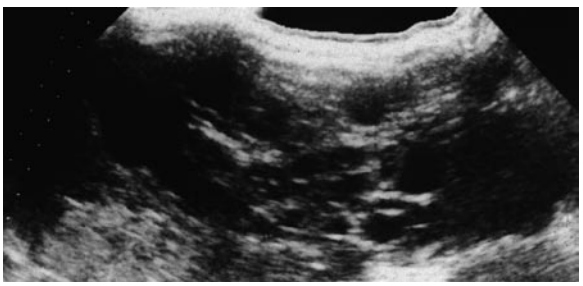


Fig. 11.59 **Ultrasound scan of a polycystic kidney,** showing an enlarged kidney with many cysts of varying size.

concentrations has been found to enhance the effect of desmopressin to stimulate chloride secretion in cultured epithelial cells from ADPKD cysts. These conditions were long considered untreatable hereditable defects. Over the last couple of years, several therapeutic approaches have emerged, e.g. sirolimus, tolvaptan which are already being used in clinical trials and their results are eagerly awaited.

FURTHER READING

Chapman AB. Autosomal dominant polycystic kidney disease: time for a change? *Journal of the American Society of Nephrology* 2007; **18**: 1399–1407.

Yoder BK. Role of primary cilia in the pathogenesis of polycystic kidney disease. *Journal of the American Society of Nephrology* 2007; **18**: 1381–1388.

Medullary cystic disease ('juvenile nephronophthisis')

Juvenile nephronophthisis that develops early in childhood is commonly inherited in an autosomal recessive manner. Mutations in the genes *NPHP1–4* are present. The proteins mutated, nephrocystin and inversin, are co-localized in the cilia of the renal tubules. A similar condition developing later in childhood (medullary cystic disease) is inherited as an autosomal dominant trait, but sporadic cases occur in both conditions. Despite its name, the dominant histological finding is interstitial inflammation and tubular atrophy, with later development of medullary cysts. Progressive glomerular failure is a secondary consequence.

The dominant features are polyuria, polydipsia and growth retardation. Diagnosis is based on the family history and renal biopsy, the cysts rarely being visualized by imaging techniques.

Medullary sponge kidney

Medullary sponge kidney is an uncommon but not rare condition that usually presents with renal colic or haematuria. Although it is most often sporadic, a few affected families have been reported. The condition is characterized by dilatation of the collecting ducts in the papillae, sometimes with cystic change. In severe cases the medullary area has a sponge-like appearance. The condition may affect one or both kidneys or only part of one kidney. Cyst formation is commonly associated with the development of small calculi within the cyst. In about 20% of patients there is associated hypercalciuria or renal tubular acidosis (see p. 676). Hemihypertrophy of the skeleton has been described in this condition.

The diagnosis is made by excretion urography, which shows small calculi in the papillary zones with an increase in radiodensity around these following injection of contrast medium as the dilated or cystic collecting ducts are filled with contrast (see Fig. 11.38).

The natural history is one of intermittent colic with passage of small stones or haematuria. Renal function is usually well maintained and renal failure is unusual, except where obstructive nephropathy develops owing to the presence of stones in the pelvis or ureters.

TUMOURS OF THE KIDNEY AND GENITOURINARY TRACT

MALIGNANT RENAL TUMOURS

These comprise 1–2% of all malignant tumours, and the male to female ratio is 2:1.

Renal cell carcinoma

Renal cell carcinomas (RCC) arise from proximal tubular epithelium. They are the most common renal tumour in adults. They rarely present before the age of 40 years, the average age of presentation being 55 years.

In von Hippel–Lindau disease, an autosomal dominant disorder, bilateral renal cell carcinomas are common and haemangioblastomas, phaeochromocytomas and renal cysts are also found. Polymorphic probes from chromosome 3p, the region implicated in renal cell carcinoma, have demonstrated genetic linkage between this tumour and von Hippel–Lindau disease. It seems likely, therefore, that mutation of the same tumour suppressor gene is responsible for both renal cell carcinoma and von Hippel–Lindau disease. Deletion of the short arm of chromosome 3 is the most consistent cytogenetic finding in sporadic tumours.

Renal cell carcinomas are highly vascular tumours; microscopically most tumours are composed of large cells containing clear cytoplasm.

Clinical features

Patients are often asymptomatic but can present with haematuria, loin pain and a mass in the flank. Malaise, anorexia and weight loss (30%) occur, and 5% of patients have polycythaemia (see p. 419). Thirty per cent of patients have hypertension (due to secretion of renin by the tumour) and anaemia, due to depression of erythropoietin in approximately the same number. Pyrexia is present in about one-fifth of patients and approximately one-quarter present with metastases. Rarely, a left-sided varicocele may be associated with left-sided tumours that have invaded the renal vein and caused obstruction to drainage of the left testicular vein.

Diagnosis

Ultrasonography is used to demonstrate the solid lesion and to examine the patency of the renal vein and inferior vena cava. CT scanning is used to identify the renal lesion and involvement of the renal vein or inferior vena cava. MRI is better than CT for tumour staging. Renal arteriography will reveal the tumour's circulation but is now seldom employed. Urine cytology for malignant cells is of no value. The ESR is usually raised. Liver biochemistry may be abnormal, returning to normal after surgery.

Treatment

- A nephrectomy is performed unless bilateral tumours are present or the contralateral kidney functions poorly, in which case conservative surgery such as partial nephrectomy may be indicated. If metastases are present, nephrectomy may still be warranted since regression of metastases has been reported after removal of the main tumour mass.
- Medroxyprogesterone acetate is of some value in controlling metastatic disease.
- Interleukin-2 and interferon produce a remission in about 20% of cases. A striking regression of metastases has been reported after non-myeloablative chemotherapy followed by allogeneic (sibling) peripheral blood stem cell transplantation.
- Newer therapies. Temsirolimus, a specific inhibitor of rapamycin kinase, has been shown to improve overall survival among patients with metastatic renal cell carcinoma and a poor prognosis and was more effective

than interferon. Everolimus, a similar drug, is also available.

- Bevacizumab (a neutralizing antibody to VEGF) has been used in trials (due to increased mutations in the tumour suppression gene *VHL* and the oversecretion of vascular endothelial growth factor (VEGF)). This has shown significant slowing in the rate of progression of metastatic renal cell carcinoma but did not prolong overall patient survival. In a recent study, *adjuvant autologous renal tumour cell vaccine* given intradermally every 4 weeks for 6 months at the time of radical nephrectomy for non-metastatic RCC was associated with an increase in 5-year disease progression-free survival – 77% compared to 68% in the placebo group. Sunitinib and sorafenib, tyrosine kinase inhibitors, have also improved clinical outcomes in randomized trials. Combinations of these agents are being evaluated.

Prognosis

The prognosis depends upon the degree of differentiation of the tumour and whether or not metastases are present. The 5-year survival rate is 60–70% with tumours confined to the renal parenchyma, 15–35% with lymph node involvement, and only approximately 5% in those who have distant metastases.

Nephroblastoma (Wilms' tumour)

This tumour is seen mainly within the first 3 years of life and may be bilateral. It presents as an abdominal mass, rarely with haematuria. Diagnosis is established by ultrasound, CT and MRI. A combination of nephrectomy, radiotherapy and chemotherapy has much improved survival rates, even in children with metastatic disease. Overall, the 5-year survival rate is 90%.

UROTHELIAL TUMOURS

The calyces, renal pelvis, ureter, bladder and urethra are lined by transitional cell epithelium. Transitional cell tumours account for about 3% of deaths from all forms of malignancy. Such tumours are uncommon below the age of 40 years, and the male to female ratio is 4:1. Bladder tumours are about 50 times as common as those of the ureter or renal pelvis.

Predisposing factors include:

- cigarette smoking
- exposure to industrial carcinogens such as β-naphthylamine and benzidine (workers in the chemical, cable and rubber industries are at particular risk) or ingestion of aristolochic acid found in some herbal weight-loss preparations
- exposure to drugs (e.g. phenacetin, cyclophosphamide)
- chronic inflammation (e.g. schistosomiasis, usually associated with squamous carcinoma).

Presentation

Painless haematuria is the most common presenting symptom of bladder malignancy, although pain may occur owing to clot retention. Symptoms suggestive of UTI develop in the absence of significant bacteriuria. In patients with bladder cancer, pain also results from local nerve involvement. Presenting symptoms may result from local metastases.

Transitional cell carcinomas in the kidney and ureter also present with haematuria. They may also give rise to flank pain, particularly if urinary tract obstruction is present.

Investigations

Cytological examination of urine for malignant cells and renal imaging (ultrasonography, CT and MRI) should be performed in all patients. Cystoscopy is necessary unless pathology is found in the upper urinary tract. It may be omitted in men under 20 and women under 30 years if significant bacteriuria accompanies the haematuria and ceases following control of the infection, provided urine cytology and renal imaging are normal. With these exceptions, haematuria should always be investigated. In cases where the tumour is not clearly outlined on ultrasonography or CT, retrograde ureterography may be helpful.

Treatment

Pelvic and ureteric tumours

These are treated by nephroureterectomy. Radiotherapy and chemotherapy appear to be of little or no value. Subsequently, cystoscopy should be regularly carried out, since about half the patients will develop bladder tumours.

Bladder tumours

Treatment depends upon the stage of the tumour (in particular whether it has penetrated the bladder muscle) and its degree of differentiation.

Superficial bladder tumours are treated by transurethral resection or local diathermy with follow-up check cystoscopies and cytological examination of the urine. Bladder instillation of doxorubicin, mitomycin and thiotepa is used for recurrent tumours, as is instillation of BCG (bacille Calmette–Guérin).

Invasive bladder tumours are treated with radical cystectomy in patients under 70 years and radical radiotherapy in those over 70 years with salvage cystectomy for recurrences. Cystectomy requires a new bladder to be made out of small bowel, joining this to the urethra if possible or making an ileal conduit.

Prognosis

The prognosis ranges from a 5-year survival rate of 80–90% for lesions not involving bladder muscle to 5% for those presenting with metastases.

Testicular tumours are discussed on page 492.

DISEASES OF THE PROSTATE GLAND

Benign enlargement of the prostate gland

Benign prostatic enlargement occurs most often in men over the age of 60 years. Such enlargement is much less common in Asian individuals. It is unknown in eunuchs. The aetiology of the condition is unknown. Microscopically, hyperplasia affects the glandular and connective tissue elements of the prostate. Enlargement of the gland stretches and distorts the urethra, obstructing bladder outflow.

Clinical features

Frequency of urination, usually first noted as nocturia, is a common early symptom. Difficulty or delay in initiating urination, with variability and reduced forcefulness of the urinary stream and post-void dribbling, are often present. Acute retention of urine (see below) or retention with overflow incontinence may occur. Occasionally, severe haematuria results from rupture of prostatic veins or as a consequence of bacteriuria or stone disease. Occasional patients present with severe renal failure.

Abdominal examination for bladder enlargement together with examination of the rectum is essential. A benign prostate feels smooth. An accurate impression of prostatic size cannot be obtained on rectal examination.

Management and prognosis

Patients with mild-to-moderate symptoms should be managed by 'watchful waiting', because symptoms following therapy are sometimes greater than those with no therapy at all.

Patients with moderate prostatic symptoms can be treated medically. A number of drugs have been employed, including alpha-blockers such as tamsulosin. Finasteride is a competitive inhibitor of 5α-reductase, which is the enzyme involved in the conversion of testosterone to dihydrotestosterone. This is the androgen primarily responsible for prostatic growth and enlargement. Finasteride decreases prostatic volume with an increase in urine flow. Deterioration in renal function or the development of upper tract dilatation requires surgery.

In acute retention or retention with overflow, the first priorities are to relieve pain and to establish urethral catheter drainage. If urethral catheterization is impossible, suprapubic catheter drainage should be carried out. The choice of further management is then between immediate prostatectomy, a period of catheter drainage followed by prostatectomy, or the acceptance of a permanent indwelling suprapubic or urethral catheter.

Prostatic carcinoma

Prostatic carcinoma accounts for 7% of all cancers in men and is the sixth most common cancer in the world. Malignant change within the prostate becomes increasingly common with advancing age. By the age of 80 years, 80% of men have malignant foci within the gland, but most of these appear to lie dormant. Histologically, the tumour is an adenocarcinoma. Hormonal factors are thought to play a role in the aetiology.

Clinical features

Presentation is usually with symptoms of lower urinary tract obstruction; less common are symptoms of metastatic spread, e.g. back pain, weight loss or anaemia. The diagnosis is also made by the incidental finding of a hard irregular gland on rectal examination, or as an unexpected histological result after prostatectomy for what was believed to be benign prostatic hypertrophy. In developed countries, patients now present as a result of screening for prostate cancer by measurement of prostate-specific antigen (PSA). However, on the evidence available, national programmes of screening are not justified. Treatment of well people carries a high morbidity of urinary incontinence and sexual dysfunction with no evidence as yet of increased overall survival. In future, screening of 'at-risk' groups may be useful. PSA >4 ng/mL is abnormal but between 4 and 10 ng/mL this can be due to benign hypertrophy and cancer. If PSA is over 10 ng/mL, a prostatic biopsy will show cancer in over 50% of cases.

Investigations

These include transrectal ultrasound of the prostate and prostatic biopsy. A histological diagnosis is essential before treatment. The Gleason scoring system is based on the histological appearances. If metastases are present, serum prostate-specific antigen levels are usually markedly elevated (>16 ng/mL) but can be normal; it is a myth that elevated levels occur as a result of rectal examination.

Ultrasonography and transrectal ultrasonography are also of value in defining the size of the gland and staging any tumour present. Endorectal coil MRI helps to detect extraprostatic extension. The upper renal tracts can be examined by ultrasonography for evidence of dilatation. Bone metastases appear as osteosclerotic lesions on X-ray and are also detected by isotopic bone scans.

Treatment (see also p. 492)

Prophylactic finasteride therapy can prevent or delay the appearance of prostatic carcinoma in men above the age of 55 years but usually at the expense of sexual side-effects and increased risk of high-grade prostatic cancer. Microscopic, impalpable tumours can sometimes be managed expectantly. Treatment for disease confined to the gland is radical prostatectomy (provided the patient is fit for the procedure) or radiotherapy. A randomized trial comparing radical prostatectomy with active surveillance in early prostate cancer revealed that radical prostatectomy significantly reduced disease-specific mortality, but there was no significant difference between surgery and active surveillance in terms of overall survival. Locally extensive disease is managed with radiotherapy with or without androgen ablation therapy. Metastatic disease can be treated with orchidectomy, but many men refuse. Luteinizing hormone-releasing hormone (LHRH) analogues such as buserelin or goserelin are equally effective and preferred by many. Abiraterone inhibits CYP17A1, an enzyme necessary for androgen production and has been used for aggressive prostate cancer. Non-hormonal chemotherapy is usually unhelpful. The duration of survival depends on the age of the patient and the degree of differentiation and extent of the tumour.

RENAL DISEASE IN THE ELDERLY

Progressive sclerosis of glomeruli occurs with ageing and this, together with the development of atheromatous renal vascular disease, accounts for the progressive reduction in GFR seen with advancing years. A GFR of 50–60 mL/min (about half the normal value for a young adult) may be regarded as 'normal' in patients in their eighties. The reduction in muscle mass often seen with ageing may mask this deterioration in renal function in that the serum creatinine concentration may be less than 120 μmol/L in an elderly patient whose eGFR is 50 mL/min or lower. The use of serum creatinine as a measure of renal function in the elderly must take this into account. This is especially so in the elderly when prescribing drugs whose excretion is in whole or in part by the kidney.

www.studentconsult.com
Ask the authors

Urinary tract infections

UTIs are common in the elderly, in whom impaired bladder emptying due to prostatic disease in males and neuropathic bladder – especially common in females – is frequently found. Symptoms may be atypical, the major complaints being incontinence, nocturia, smelly urine or vague change in well-being with little in the way of dysuria. Demonstration of significant bacteriuria in the presence of such symptoms requires treatment.

Urinary incontinence

This is defined as involuntary passage of urine sufficient to be a health or social problem. It is common in the elderly with 25% of women and 15% of men over 65 having a problem.

- *Urge incontinence* is usually due to detrusor overactivity with leakage of urine because the bladder is perceived to be full. This is common in the elderly; it occurs as an isolated event or secondary to local factors, e.g. bladder infection or stones, or to central factors, e.g. stroke, dementia or Parkinson's disease.
- *Stress incontinence* occurs when the intra-abdominal pressure is increased, e.g. after a cough or sneeze and there is a weak pelvic floor or urethral sphincter. It is common in women after childbirth.
- *Overflow incontinence* occurs with leakage of urine from a full distended bladder. It occurs commonly in men with prostatic obstruction, following spinal cord injury or in women with cystoceles or after gynaecological surgery.
- *Functional incontinence*. Passage of urine occurs owing to inability to get to a toilet because of disability, e.g. stroke, trauma, the unavailability of toilet facilities or dementia.

Management

- Physical examination for local problems, e.g. prostatic enlargement in men, gynaecological disorders in women, and for central problems, e.g. neurological disorders or dementia.
- Urine analysis, e.g. glycosuria and culture for UTI.
- Treatment of contributing causes, e.g. constipation, drug therapy, other co-existing disease.
- Urge incontinence – bladder training, antimuscarinics, e.g. oxybutynin, tolterodine, solifenacin and darifenacin.
- Stress incontinence – pelvic floor exercises. Moreover, recently it was demonstrated that transurethral injections of autologous myoblasts can aid in regeneration of the rhabdosphincter, and fibroblasts in reconstruction of the urethral submucosa. Effectiveness and tolerability of ultrasonography-guided injections of autologous cells was much superior compared with those of endoscopic injections of collagen for stress incontinence.
- Overflow – removal of obstruction.
- Functional – improve facilities, regular urine voiding, absorbent padding.

Further evaluation with urodynamics is necessary in patients who do not respond and have a potentially curable problem. An expert and committed incontinence advisory and treatment service combining nursing and medical skills is invaluable for elderly patients with this distressing problem. Home visits to ensure the availability of commodes and toilets are essential. For established incontinence, catheterization may be necessary but should be avoided if at all possible. Incontinence and its treatment are matters of major importance and are by no means confined solely to the elderly.

FURTHER READING

Cohen HT, McGovern FJ. Renal-cell carcinoma. *New England Journal of Medicine* 2005; **353**: 2477–2490.

Damber JE, Aus G. Prostate cancer. *Lancet* 2008; **371**: 1710–1721.

Motzer RJ, Bosch E. Targeted drugs for metastatic renal cell carcinoma. *Lancet* 2007; **370**: 2071–2073.

Powles T, Eu PJ. Does PET imaging have a role in renal cancers after all? *Lancet Oncology* 2007; **8**: 279–281.

Walsh PC, Deweese TL, Eisenberg MA. Localised prostate cancer. *Lancet* 2007; **357**: 2696–2705.

CHAPTER BIBLIOGRAPHY

Cameron JS, Davison AM, Grunfeld JP et al (eds) (2004) *Oxford Textbook of Clinical Nephrology*. Oxford: Oxford University Press.

Current Opinion in Nephrology and Hypertension is a monthly journal with review articles, each issue devoted to one or two topics.

Journal of the American Society of Nephrology is the highest impact journal in nephrology with bimonthly self-assessment programme (SAP) supplements.

Kidney International is the major journal associated with the International Society of Nephrology – monthly with original and review articles.

Nephrology, Dialysis, Transplantation is the major European journal devoted to the subject, with review articles, editorial comments and original papers.

SIGNIFICANT WEBSITES

http://www.tinkershop.net/nephro.htm
Nephrology calculator

http://www.nephronline.org
For healthcare professionals involved in the management of patients with kidney disease

http://www.renalnet.org
Kidney information clearing house database

http://www.kidney.org.uk/
UK charity run by and for patients

This page intentionally left blank

12 Water, electrolytes and acid–base balance

Distribution and composition of body water 649

Disorders of sodium concentration 661
Hyponatraemia 661
Hypernatraemia 663

Disorders of potassium concentration 664
Hypokalaemia 665
Hyperkalaemia 666

Disorders of magnesium concentration 668
Hypomagnesaemia 669
Hypermagnesaemia 669

Disorders of phosphate concentration 670
Hypophosphataemia 670
Hyperphosphataemia 671

Acid–base disorders 672
Respiratory acidosis and alkalosis 675
Metabolic acidosis 675
Metabolic alkalosis 678

Questions and case studies 679

In health, the volume and biochemical composition of both extracellular and intracellular fluid compartments in the body remains remarkably constant. Many different disease states result in changes of control, either of extracellular fluid volume or of the electrolyte composition of extracellular fluid. An understanding of these abnormalities is therefore essential for the management of a wide range of clinical disorders.

DISTRIBUTION AND COMPOSITION OF BODY WATER

In normal persons, the total body water constitutes 50–60% of lean bodyweight in men and 45–50% in women. In a healthy 70 kg male, total body water is approximately 42 L. This is contained in three major compartments:

- intracellular fluid (28 L, about 35% of lean bodyweight)
- extracellular – the interstitial fluid that bathes the cells (9.4 L, about 12%)
- plasma (also extracellular) (4.6 L, about 4–5%).

In addition, small amounts of water are contained in bone, dense connective tissue, and epithelial secretions, such as the digestive secretions and cerebrospinal fluid.

The intracellular and interstitial fluids are separated by the cell membrane; the interstitial fluid and plasma are separated by the capillary wall (Fig. 12.1). In the absence of solute, water molecules move randomly and in equal numbers in either direction across a permeable membrane. However, if solutes are added to one side of the membrane, the intermolecular cohesive forces reduce the activity of the water molecules. As a result, water tends to stay in the solute-containing compartment because there is less free diffusion across the membrane. This ability to hold water in the compartment can be measured as the osmotic pressure.

Osmotic pressure

Osmotic pressure is the primary determinant of the distribution of water among the three major compartments. The concentrations of the major solutes in these fluids differ, with each compartment having one solute that is primarily limited to that compartment and therefore determines its osmotic pressure. The intracellular fluid contains mainly potassium (K^+) (most of the cell Mg^{2+} is bound and osmotically inactive) and in the extracellular compartment, Na^+ salts predominate in the *interstitial fluid*, and proteins in the plasma. Regulation of the plasma volume is somewhat more complicated because of the tendency of the plasma proteins to hold water in the vascular space by an oncotic effect which is in part counterbalanced by the hydrostatic pressure in the capillaries that is generated by cardiac contraction (Fig. 12.1). The composition of intracellular and extracellular fluids is shown in Table 12.1.

A characteristic of an osmotically active solute is that it cannot freely leave its compartment. The capillary wall, for example, is relatively impermeable to plasma proteins, and the cell membrane is 'impermeable' to Na^+ and K^+ because the Na^+/K^+-ATPase pump largely restricts Na^+ to the extracellular fluid and K^+ to the intracellular fluid. By contrast, Na^+ freely crosses the capillary wall and achieves similar concentrations in the interstitium and plasma; as a result, it does not contribute to fluid distribution between these compartments. Similarly, urea crosses both the capillary wall and the cell membrane and is osmotically inactive. Thus, the retention of urea in renal failure does not alter the distribution of the total body water.

A conclusion from these observations is that body Na^+ stores are the primary determinant of the extracellular fluid volume. Thus the extracellular volume – and therefore tissue perfusion – are maintained by appropriate alterations in Na^+ excretion. For example, if Na^+ intake is increased, the extra Na^+ will initially be added to the extracellular fluid. The associated increase in extracellular osmolality will cause water to

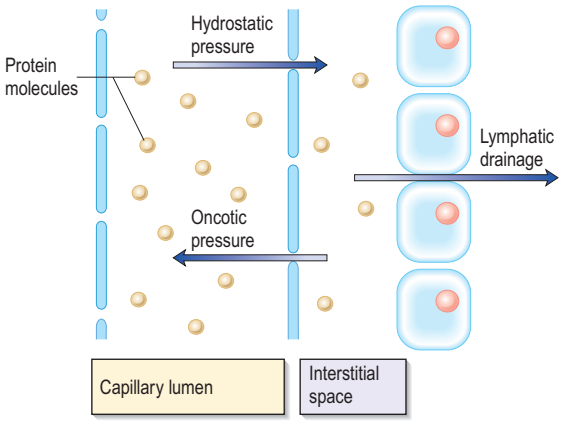


Fig. 12.1 **Distribution of water between the vascular and extravascular (interstitial) spaces.** This is determined by the equilibrium between hydrostatic pressure, which tends to force fluid out of the capillaries, and oncotic pressure, which acts to retain fluid within the vessel. The net flow of fluid outwards is balanced by 'suction' of fluid into the lymphatics, which returns it to the bloodstream. Similar principles govern the volume of the peritoneal and pleural spaces.

Table 12.1 **Electrolyte composition of intracellular and extracellular fluids**

	Plasma (mmol/L)	Interstitial fluid (mmol/L)	Intracellular fluid (mmol/L)
Na^+	142	144	10
K^+	4	4	160
Ca^{2+}	2.5	2.5	1.5
Mg^{2+}	1.0	0.5	13
Cl^-	102	114	2
HCO_3^-	26	30	8
PO_4^{2-}	1.0	1.0	57
SO_4^{2-}	0.5	0.5	10
Organic acid	3	4	3
Protein	16	0	55

move out of the cells, leading to extracellular volume expansion. Balance is restored by excretion of the excess Na^+ in the urine.

Distribution of different types of replacement fluids

Figure 12.2 shows the relative effects on the compartments of the addition of identical volumes of water, saline and colloid solutions. Thus, 1 L of water given intravenously as 5% dextrose is distributed equally into all compartments, whereas the same amount of 0.9% saline remains in the extracellular compartment. The latter is thus the correct treatment for extracellular water depletion – sodium keeping the water in this compartment. The addition of 1 L of colloid with its high oncotic pressure stays in the vascular compartment and is the treatment for hypovolaemia.

Regulation of extracellular volume (Fig. 12.3)

The extracellular volume is determined by the sodium concentration. The regulation of extracellular volume is dependent upon a tight control of sodium balance, which is exerted by normal kidneys. Renal Na^+ excretion varies directly with the effective circulating volume. In a 70 kg man, plasma fluid constitutes one-third of extracellular volume (4.6 L), of which 85% (3.9 L) lies in the venous side and only 15% (0.7 L) resides in the arterial circulation. The unifying hypothesis of extracellular volume regulation in health and disease proposed by Schrier states that the fullness of the arterial vascular compartment – or the so-called *effective arterial blood volume (EABV)* – is the primary determinant of renal sodium and water excretion. Thus effective arterial blood volume constitutes effective circulatory volume for the purposes of body fluid homeostasis. The fullness of the arterial compartment depends upon a normal ratio between cardiac output and peripheral arterial resistance. Thus, diminished EABV is initiated by a fall in cardiac output or a fall in peripheral arterial resistance (an increase in the holding capacity of the arterial vascular tree). When the EABV is expanded, the urinary Na^+ excretion is increased and can exceed 100 mmol/L. By contrast, the urine can be rendered virtually free of Na^+ in the presence of EABV depletion and normal renal function.

These changes in Na^+ excretion can result from alterations both in the filtered load, determined primarily by the glomerular filtration rate (GFR), and in tubular reabsorption, which is affected by multiple factors. In general, it is changes in tubular reabsorption that constitute the main adaptive response to fluctuations in the effective circulating volume. How this occurs can be appreciated from Table 12.2 and

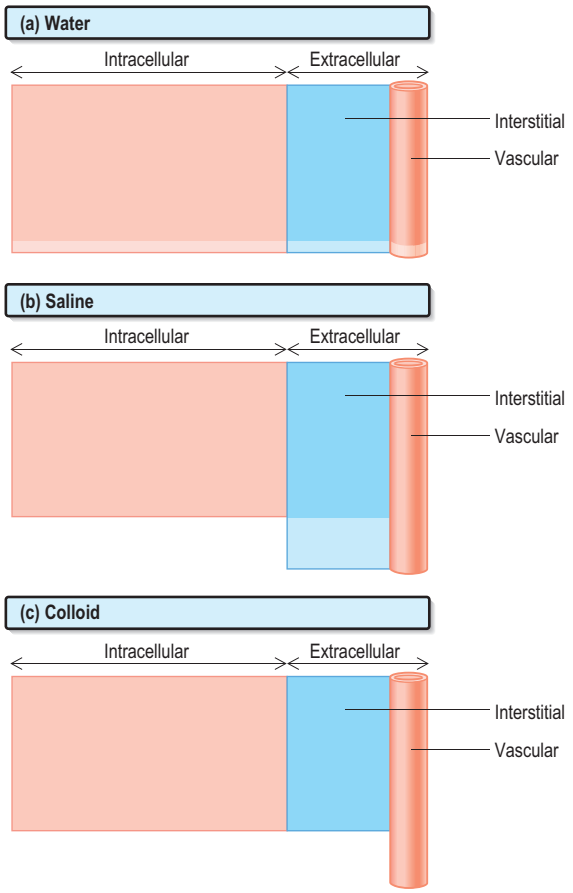


Fig. 12.2 **Relative effects of the addition of 1 L of (a) water, (b) saline 0.9% and (c) a colloid solution.**

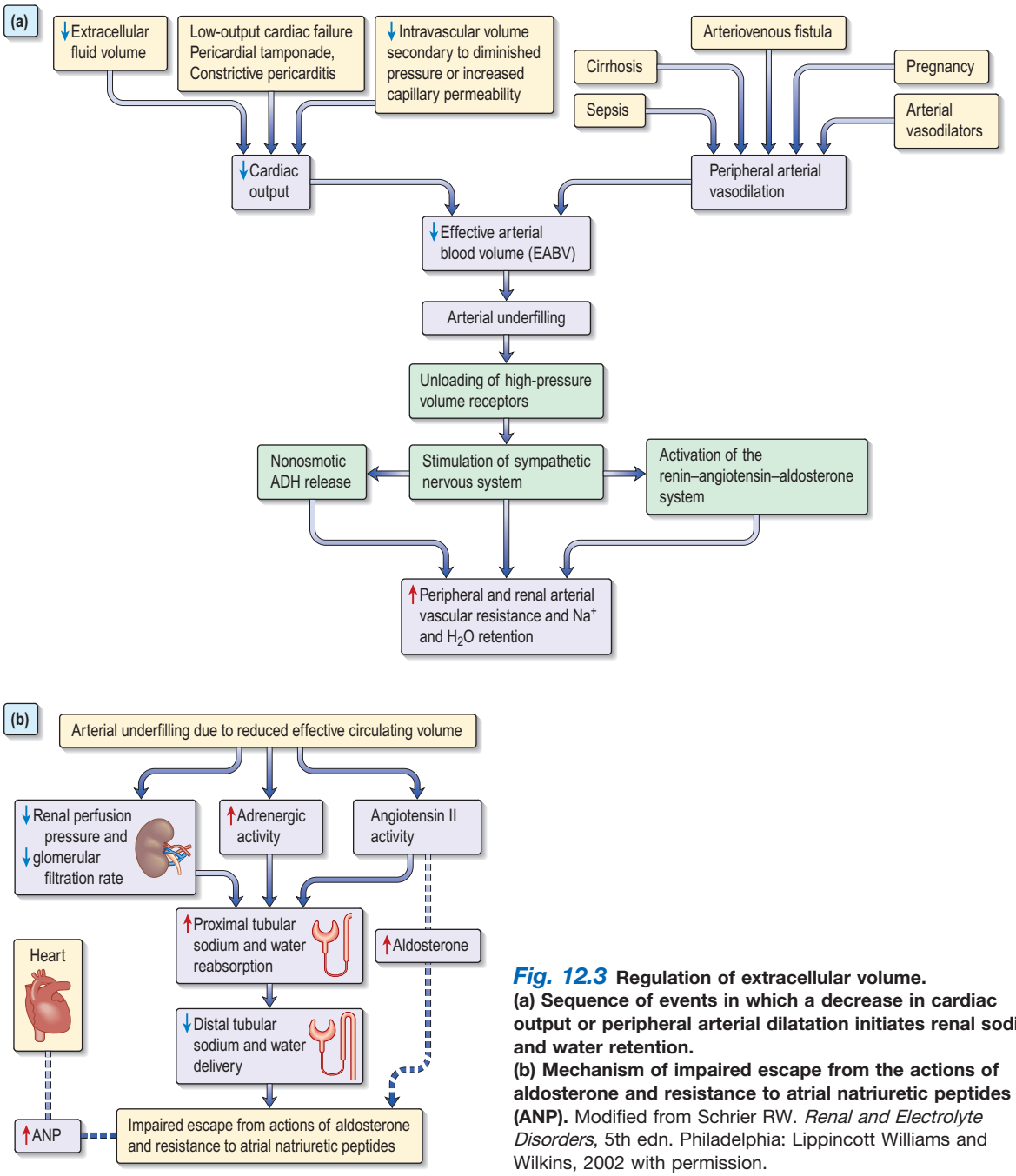


Fig. 12.3 **Regulation of extracellular volume.**
(a) Sequence of events in which a decrease in cardiac output or peripheral arterial dilatation initiates renal sodium and water retention.
(b) Mechanism of impaired escape from the actions of aldosterone and resistance to atrial natriuretic peptides (ANP). Modified from Schrier RW. *Renal and Electrolyte Disorders*, 5th edn. Philadelphia: Lippincott Williams and Wilkins, 2002 with permission.

Figure 12.4 and Figure 11.2 (see p. 573), which depicts the sites and determinants of segmental Na^+ reabsorption. Although the loop of Henle and distal tubules make a major overall contribution to net Na^+ handling, transport in these segments primarily varies with the amount of Na^+ delivered; that is, reabsorption is flow-dependent. In comparison, the neurohumoral regulation of Na^+ reabsorption according to body needs occurs primarily in the proximal tubules and collecting ducts.

Neurohumoral regulation of extracellular volume

This is mediated by volume receptors that sense changes in the EABV rather than alterations in the sodium concentration. These receptors are distributed in both the renal and cardiovascular tissues.

- *Intrarenal receptors*. Receptors in the walls of the afferent glomerular arterioles respond, via the juxtaglomerular apparatus, to changes in renal perfusion, and control the activity of the renin–angiotensin–aldosterone system (see p. 1023). In addition, sodium concentration in the distal tubule and sympathetic nerve activity alter renin release from the juxtaglomerular cells. Prostaglandins I_2 and E_2 are also generated within the kidney in response to angiotensin II, acting to maintain glomerular filtration rate and sodium and water excretion, modulating the sodium-retaining effect of this hormone.

Table 12.2 Mechanisms of sodium transport in the various nephron segments

	Reabsorbed (%)	Luminal Na^+ entry	Transport
Proximal tubule	60–70	Na^+–H^+ exchange and cotransport of Na^+ with glucose, phosphate and other organic solutes	Angiotensin II Norepinephrine (noradrenaline)
Loop of Henle	20–25	Na^+–K^+–$2Cl^-$ cotransport	Flow dependent Pressure natriuresis mediated by nitric oxide
Distal tubule	5	Na^+–Cl^- cotransport	Flow dependent
Collecting ducts	4	Na^+ channels	Aldosterone Atrial natriuretic peptide

- *Extrarenal receptors*. These are located in the vascular tree in the left atrium and major thoracic veins, and in the carotid sinus body and aortic arch. These volume receptors respond to a slight reduction in effective circulating volume and result in increased sympathetic nerve activity and a rise in catecholamines. In addition, volume receptors in the cardiac atria control the release of a powerful natriuretic hormone – atrial natriuretic peptide (ANP) – from granules located in the atrial walls (see p. 1023).

High-pressure arterial receptors (carotid, aortic arch, juxtaglomerular apparatus) predominate over low-pressure volume receptors in volume control in mammals. The low-pressure volume receptors are distributed in thoracic tissues (cardiac atria, right ventricle, thoracic veins, pulmonary vessels) and their role in the volume regulatory system is marginal.

Aldosterone and possibly ANP are responsible for day-to-day variations in Na^+ excretion, by their respective ability to augment and diminish Na^+ reabsorption in the collecting ducts.

A *salt load*, for example, leads to an increase in the effective circulatory and extracellular volume, raising both renal perfusion pressure, and atrial and arterial filling pressure. The increase in the renal perfusion pressure reduces the secretion of renin, and subsequently that of angiotensin II and aldosterone (Fig. 18.27), whereas the rise in atrial and arterial filling pressure increases the release of ANP. These factors combine to reduce Na^+ reabsorption in the collecting duct, thereby promoting excretion of excess Na^+.

By contrast, in patients on a *low* Na^+ intake or in those who become volume-depleted as a result of vomiting and diarrhoea, the ensuing decrease in effective volume enhances the activity of the renin–angiotensin–aldosterone system and reduces the secretion of ANP. The net effect is enhanced Na^+ reabsorption in the collecting ducts, leading to a fall in Na^+ excretion. This increases the extracellular volume towards normal.

With more marked hypovolaemia, a decrease in GFR leads to an increase in proximal and thin ascending limb Na^+ reabsorption which contributes to Na^+ retention. This is brought about by enhanced sympathetic activity acting directly on the kidneys and indirectly by stimulating the secretion of renin/angiotensin II (see Fig. 12.3b) and non-osmotic release of antidiuretic hormone (ADH), also called vasopressin. The pressure natriuresis phenomenon may be the final defence against changes in the effective circulating volume. Marked persistent hypovolaemia leads to systemic hypotension and increased salt and water absorption in the

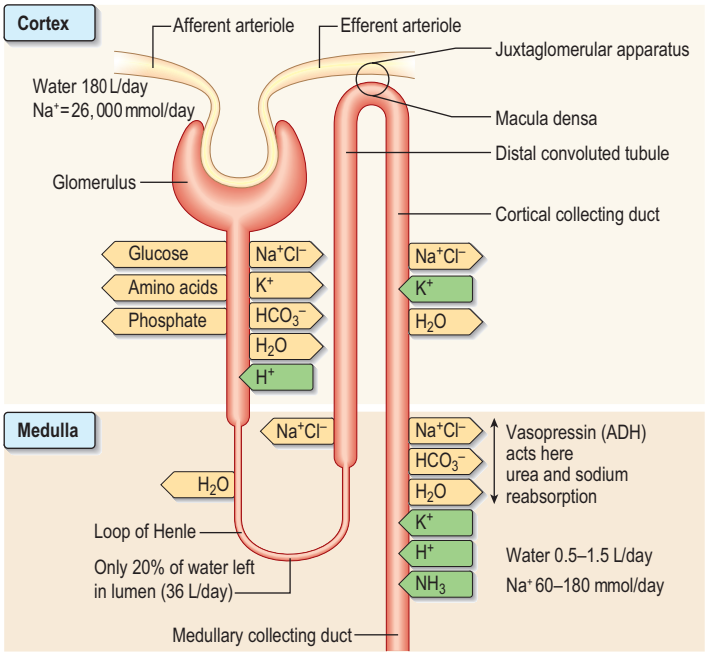


Fig. 12.4 **The nephron – electrolyte and water exchange.** 180 L of water and 26 000 mmol of sodium per day enter the nephrons via the afferent arterioles to the kidneys. Removal or addition of electrolytes results in the excretion of approximately 1 L of water and 60–180 mmol of sodium per day.

proximal tubules and ascending limb of Henle. This process is partly mediated by changes in renal interstitial hydrostatic pressure and local prostaglandin and nitric oxide production.

Volume regulation in oedematous conditions

Sodium and water are retained despite increased extracellular volume in oedematous conditions such as cardiac failure, hepatic cirrhosis and hypoalbuminaemia. Here the principal mediator of salt and water retention is the concept of arterial underfilling due either to reduced cardiac output or diminished peripheral arterial resistance. Arterial underfilling in these settings leads to reduction of pressure or stretch (i.e. 'unloading' of arterial volume receptors), which results in activation of the sympathetic nervous system, activation of the renin–angiotensin–aldosterone system and non-osmotic release of ADH. These neurohumoral mediators promote salt and water retention in the face of increased extracellular volume. The common nature of the degree of arterial fullness and neurohumoral pathway in the regulation of extracellular volume in health and disease states forms the basis of Schrier's unifying hypothesis of volume homeostasis (Fig. 12.3a).

Mechanism of impaired escape from actions of aldosterone and resistance to ANP

Not only is the activity of the renin–angiotensin–aldosterone system increased in oedematous conditions such as cardiac failure, hepatic cirrhosis and hypoalbuminaemia, but also the action of aldosterone is more persistent than in normal subjects and patients with Conn's syndrome, who have increased aldosterone secretion (see p. 1023).

In normal subjects, high doses of mineralocorticoids initially increase renal sodium retention so that the extracellular volume is increased by 1.5–2 L. However, renal sodium retention then ceases, sodium balance is re-established, and there is no detectable oedema. This escape from mineralocorticoid-mediated sodium retention explains why oedema is not a characteristic feature of primary hyperaldosteronism (Conn's syndrome). The escape is dependent on an increase in delivery of sodium to the site of action of aldosterone in the collecting ducts. The increased distal sodium delivery is achieved by high extracellular volume-mediated arterial overfilling. This suppresses sympathetic activity and angiotensin II generation, and increases cardiac release of ANP with resultant increase in renal perfusion pressure and GFR. The net result of these events is reduced sodium absorption in the proximal tubules and increased distal sodium delivery which overwhelms the sodium-retaining actions of aldosterone.

In patients with the above oedematous conditions, escape from the sodium-retaining actions of aldosterone **does not occur** and therefore they continue to retain sodium in response to aldosterone. Accordingly they have substantial natriuresis when given spironolactone, which blocks mineralocorticoid receptors. Alpha-adrenergic stimulation and elevated angiotensin II increase sodium transport in the proximal tubule, and reduced renal perfusion and GFR further increase sodium absorption from the proximal tubules by presenting less sodium and water in the tubular fluid. Sodium delivery to the distal portion of the nephron, and thus the collecting duct, is reduced. Similarly, increased cardiac ANP release in these conditions requires optimum sodium concentration at the site of its action in the collecting duct for its desired natriuretic effects. Decreased sodium delivery to the collecting duct is therefore the most likely explanation for the persistent aldosterone-mediated sodium retention, absence of escape phenomenon and resistance to natriuretic peptides in these patients (Fig. 12.3b).

Regulation of water excretion

Body water homeostasis is effected by thirst and the urine concentrating and diluting functions of the kidney. These in turn are controlled by intracellular osmoreceptors, principally in the hypothalamus, to some extent by volume receptors in capacitance vessels close to the heart, and via the renin–angiotensin system. Of these, the major and best-understood control is via osmoreceptors. Changes in the plasma Na^+ concentration and osmolality are sensed by osmoreceptors that influence both thirst and the release of ADH (vasopressin) from the supraoptic and paraventricular nuclei.

ADH plays a central role in urinary concentration by increasing the water permeability of the normally impermeable cortical and medullary collecting ducts. There are three major G-protein coupled receptors for vasopressin (ADH):

- V_{1A} found in vascular smooth muscle cells
- V_{1B} in anterior pituitary and throughout the brain
- V_2 receptors in the principal cells of the kidney distal convoluting tubule and collecting ducts.

Activation of the V_{1A} receptors induces vasoconstriction while V_{1B} receptors appear to mediate the effect of ADH on the pituitary, facilitating the release of ACTH. The V_2 receptors mediate the antidiuretic response as well as other functions.

The ability of ADH to increase the urine osmolality is related indirectly to transport in the ascending limb of the loop of Henle, which reabsorbs NaCl without water. This process, which is the primary step in the countercurrent mechanism, has two effects: it makes the tubular fluid dilute and the medullary interstitium concentrated. In the absence of ADH, little water is reabsorbed in the collecting ducts, and a dilute urine is excreted. By contrast, the presence of ADH promotes water reabsorption in the collecting ducts down the favourable osmotic gradient between the tubular fluid and the more concentrated interstitium. As a result, there is an increase in urine osmolality and a decrease in urine volume.

The cortical collecting duct has two cell types (see also p. 573) with very different functions:

- *Principal cells* (about 65%) have sodium and potassium channels in the apical membrane and, as in all sodium-reabsorbing cells, Na^+/K^+-ATPase pumps in the basolateral membrane.
- *Intercalated cells*, in comparison, do not transport NaCl (since they have a lower level of Na^+/K^+-ATPase activity) but play a role in hydrogen and bicarbonate handling and in potassium reabsorption in states of potassium depletion.

The ADH-induced increase in collecting duct water permeability occurs primarily in the principal cells. ADH acts on V_2 (vasopressin) receptors located on the basolateral surface of principal cells, resulting in the activation of adenyl cyclase. This leads to protein kinase activation and to preformed cytoplasmic vesicles that contain unique water channels (called aquaporins) moving to and then being inserted into the luminal membrane. Four renal aquaporins have been well

characterized and are localized in different areas of the cells of the collecting duct. The water channels span the luminal membrane and permit water movement into the cells down a favourable osmotic gradient (Fig. 12.5). This water is then rapidly returned to the systemic circulation across the basolateral membrane. When the ADH effect has worn off, the water channels aggregate within clathrin-coated pits, from which they are removed from the luminal membrane by endocytosis and returned to the cytoplasm. A defect in any step in this pathway, such as in attachment of ADH to its receptor or the function of the water channel, can cause resistance to the action of ADH and an increase in urine output. This disorder is called *nephrogenic diabetes insipidus*.

Plasma osmolality

In addition to influencing the rate of water excretion, ADH plays a central role in osmoregulation because its release is directly affected by the plasma osmolality. At a plasma osmolality of less than 275 mOsm/kg, which usually represents a plasma Na^+ concentration of less than 135–137 mmol/L, there is essentially no circulating ADH. As the plasma osmolality rises above this threshold, however, the secretion of ADH increases progressively.

Two simple examples will illustrate the basic mechanisms of osmoregulation, which is so efficient that the plasma Na^+ concentration is normally maintained within 1–2% of its baseline value.

1. *Ingestion of a water load* leads to an initial reduction in the plasma osmolality, thereby diminishing the release of ADH. The ensuing reduction in water reabsorption in the collecting ducts allows the excess water to be excreted in a dilute urine.
2. *Water loss* resulting from sweating is followed by, in sequence, a rise in both plasma osmolality and ADH secretion, enhanced water reabsorption, and the appropriate excretion of a small volume of concentrated urine. This renal effect of ADH minimizes further water loss but does not replace the existing water deficit. Thus, optimal osmoregulation requires an increase in water intake, which is mediated by a concurrent stimulation of thirst. The importance of thirst can also be illustrated by studies in patients with central diabetes insipidus, who are deficient in ADH. These patients often complain of marked polyuria, which is caused by the decline in water reabsorption in the collecting ducts. However, they do not typically become hypernatraemic, because urinary water loss is offset by the thirst mechanism.

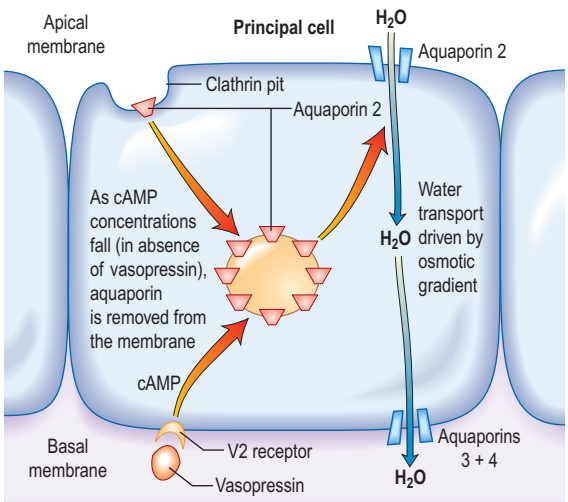


Fig. 12.5 **Aquaporin-mediated water transport in the renal collecting duct.** Stimulation of the vasopressin 2 receptor causes cAMP-mediated insertion of the aquaporin into the apical membrane, allowing water transport down the osmotic gradient. Adapted from Connolly DL, Shanahan CM, Weissberg PL. Water channels in health and disease. *Lancet* 1996; **347**: 211 with permission from Elsevier.

Osmoregulation versus volume regulation

A common misconception is that regulation of the plasma Na^+ concentration is closely correlated with the regulation of Na^+ excretion. However, it is related to volume regulation, which has different sensors and effectors (volume receptors) from those involved in water balance and osmoregulation (osmoreceptors).

The roles of these two pathways should be considered separately when evaluating patients.

- A *water load* is rapidly excreted (in 4–6 hours) by inhibition of ADH release so that there is little or no water reabsorption in the collecting ducts. This process is normally so efficient that volume regulation is not affected and there is no change in ANP release or in the activity of the renin–angiotensin–aldosterone system. Thus, a dilute urine is excreted, and there is little alteration in the excretion of Na^+.
- *Isotonic saline* administration, by contrast, causes an increase in volume but no change in plasma osmolality. In this setting, ANP secretion is increased, aldosterone secretion is reduced and ADH secretion does not change. The net effect is the appropriate excretion of the excess Na^+ in a relatively iso-osmotic urine.

In some cases, both volume and osmolality are altered and both pathways are activated. For example, if a person with normal renal function eats salted potato chips and peanuts without drinking any water, the excess Na^+ will increase the plasma osmolality, leading to osmotic water movement out of the cells and increased extracellular volume. The rise in osmolality will stimulate both ADH release and thirst (the main reason why many restaurants and bars supply free salted foods), whereas the hypervolaemia will enhance the secretion of ANP and suppress that of aldosterone. The net effect is increased excretion of Na^+ without water.

This principle of separate volume and osmoregulatory pathways is also evident in the **syndrome of inappropriate ADH secretion (SIADH)**. Patients with SIADH (see p. 1018) have impaired water excretion and hyponatraemia (dilutional) caused by the persistent presence of ADH. However, the release of ANP and aldosterone is not impaired and, thus, Na^+ handling remains intact. These findings have implications for the correction of the hyponatraemia in this setting which initially requires restriction of water intake.

ADH is also secreted by non-osmotic stimuli such as stress (e.g. surgery, trauma), markedly reduced effective circulatory volume (e.g. cardiac failure, hepatic cirrhosis), psychiatric disturbance and nausea, irrespective of plasma osmolality. This is mediated by the effects of sympathetic overactivity on supraoptic and paraventricular nuclei. In addition to water retention, ADH release in these conditions promotes vasoconstriction owing to the activation of V_{1B}

(vasopressin) receptors distributed in the vascular smooth muscle cells.

Regulation of cell volume

Maintenance of a constant volume in the face of extracellular and intracellular osmotic alterations is a critical problem faced by all cells. Most cells respond to swelling or shrinkage by activating specific metabolic or membrane-transport processes that return cell volume to its normal resting state. Within minutes after exposure to hypotonic solutions and resulting cell swelling, a common feature of many cells is the increase in plasma membrane potassium and chloride conductance. Although extrusion of intracellular potassium certainly contributes to a regulatory volume decrease, the role of chloride efflux itself is modest, given the relatively low intracellular chloride concentration. Other intracellular osmolytes, such as taurine and other amino acids, are transported out of the cell to achieve a regulatory volume decrease. By contrast, these regulatory mechanisms are operative in reverse to protect cell volume under hypertonic conditions, as is the case in the renal medulla. The tubular cells at the tip of renal papillae, which are constantly exposed to a hypertonic extracellular milieu, maintain their cell volume on a long-term basis by actively taking up smaller molecules, such as betaine, taurine and myoinositol, and by synthesizing more sorbitol and glycerophosphocholine.

Increased extracellular volume

Increased extracellular volume occurs in numerous disease states. The physical signs depend on the distribution of excess volume and on whether the increase is local or systemic. According to Starling principles, distribution depends on:

- venous tone, which determines the capacitance of the blood compartment and thus hydrostatic pressure
- capillary permeability
- oncotic pressure – mainly dependent on serum albumin
- lymphatic drainage.

Depending on these factors, fluid accumulation may result in expansion of interstitial volume, blood volume, or both.

Clinical features

Peripheral oedema is caused by expansion of the extracellular volume by at least 2 L (15%). The ankles are normally the first part of the body to be affected, although they may be spared in patients with lipodermatosclerosis (where the skin is tethered and cannot expand to accommodate the oedema). Oedema may be noted in the face, particularly in the morning. In a patient in bed, oedema may accumulate in the sacral area. Expansion of the interstitial volume also causes pulmonary oedema, pleural effusion, pericardial effusion and ascites. Expansion of the blood volume (overload) causes a raised jugular venous pressure, cardiomegaly, added heart sounds, basal crackles as well as a raised arterial blood pressure in certain circumstances.

Causes

Extracellular volume expansion is due to sodium chloride retention. Increased oral salt intake does not normally cause volume expansion because of rapid homeostatic mechanisms which increase salt excretion. However, a rapid intravenous infusion of a large volume of saline will cause volume expansion. Most causes of extracellular volume expansion are associated with renal sodium chloride retention.

Heart failure

Reduction in cardiac output and the consequent fall in effective circulatory volume and arterial filling lead to activation of the renin–angiotensin–aldosterone system, non-osmotic release of ADH, and increased activity of the renal sympathetic nerves via volume receptors and baroreceptors (Fig. 12.3a). Sympathetic overdrive also indirectly augments ADH and renin–angiotensin–aldosterone response in these conditions. The cumulative effect of these mediators results in increased peripheral and renal arteriolar resistance and water and sodium retention. These factors result in extracellular volume expansion and increased venous pressure, causing oedema formation.

Hepatic cirrhosis

The mechanism is complex, but involves peripheral vasodilatation (possibly owing to increased nitric oxide generation) resulting in reduced effective arterial blood volume (EABV) and arterial filling. This leads to an activation of a chain of events common to cardiac failure and other conditions with marked peripheral vasodilatation (Fig. 12.3). The cumulative effect results in increased peripheral and renal resistance, water and sodium retention, and oedema formation.

Nephrotic syndrome

Interstitial oedema is a common clinical finding with hypoalbuminaemia, particularly in the nephrotic syndrome. Expansion of the interstitial compartment is secondary to the accumulation of sodium in the extracellular compartment. This is due to an imbalance between oral (or parenteral) sodium intake and urinary sodium output, as well as alterations of fluid transfer across capillary walls. The intrarenal site of sodium retention is the cortical collecting duct (CCD) where Na^+/K^+-ATPase expression and activity are increased threefold along the basolateral surface (Fig. 12.4). In addition, amiloride-sensitive epithelial sodium channel activity is also increased in the CCD. The renal sodium retention should normally be counterbalanced by increased secretion of sodium in the inner medullary collecting duct, brought about by the release of ANP. This regulatory pathway is altered in patients with nephrotic syndrome by enhanced *kidney specific* catabolism of cyclic GMP (the second messenger for ANP) following phosphodiesterase activation.

Oedema generation was classically attributed to the decrease in the plasma oncotic pressure and the subsequent increase in the transcapillary oncotic gradient. However, the oncotic pressure and transcapillary oncotic gradient remain unchanged and the transcapillary hydrostatic pressure gradient is not altered. Conversely, capillary hydraulic conductivity (a measure of permeability) is increased. This is determined by intercellular macromolecular complexes between the endothelial cells consisting of tight junctions (made of occludins, claudins and ZO proteins) and adherens junctions (made of cadherin, catenins and actin cytoskeleton). Elevated TNF-α levels in nephrotic syndrome activate protein kinase C, which changes phosphorylation of occludin and capillary permeability. In addition, increased circulating ANP can increase capillary hydraulic conductivity by altering the permeability of intercellular junctional complexes. Furthermore, reduction in effective circulatory volume and the

consequent fall in cardiac output and arterial filling can lead to a chain of events as in cardiac failure and cirrhosis (see above and Fig. 12.3). These factors result in extracellular volume expansion and oedema formation.

Sodium retention

A decreased GFR decreases the renal capacity to excrete sodium. This may be acute, as in the acute nephritic syndrome (see p. 592), or may occur as part of the presentation of chronic kidney disease. In end-stage renal failure, extracellular volume is controlled by the balance between salt intake and its removal by dialysis.

Numerous drugs cause renal sodium retention, particularly in patients whose renal function is already impaired:

- *Oestrogens* cause mild sodium retention, due to a weak aldosterone-like effect. This is the cause of weight gain in the premenstrual phase.
- *Mineralocorticoids and liquorice* (the latter potentiates the sodium-retaining action of cortisol) have aldosterone-like actions.
- *NSAIDs* cause sodium retention in the presence of activation of the renin–angiotensin–aldosterone system by heart failure, cirrhosis and in renal artery stenosis.
- *Thiazolidinediones* (TZD) (see p. 1039) are widely used to treat type 2 diabetes. Their mechanism of action is attributed to binding and activation of the PPAR-γ system. PPARs are nuclear transcription factors essential to the control of energy metabolism that are modulated via binding with tissue-specific fatty acid metabolites. Of the three PPAR isoforms, γ has been extensively studied and is expressed at high levels in adipose and liver tissues, macrophages, pancreatic-β cells and principal cells of the collecting duct. These drugs have been asociated with salt and water retention and are contraindicated in patients with heart failure. Recent evidence suggests that TZD induced oedema (like insulin) is also due to upregulation of epithelial Na transporter channel (ENaC) but by different pathways. Diuretics of choice for TZD induced oedema are amiloride, triamterene and spironolactone.

Substantial amounts of sodium and water may accumulate in the body without clinically obvious oedema or evidence of raised venous pressure. In particular, several litres may accumulate in the pleural space or as ascites; these spaces are then referred to as ‘third spaces’. Bone may also act as a ‘sink’ for sodium and water.

Other causes of oedema

- Initiation of insulin treatment for type 1 diabetes and refeeding after malnutrition are both associated with the development of transient oedema. The mechanism is complex but involves upregulation of ENaC in the principal cell of the collecting duct. This transportor is amiloride sensitive which makes amiloride or triamterene the diuretic of choice in insulin induced oedema.
- Oedema may result from increased capillary pressure owing to relaxation of precapillary arterioles. The best example is the peripheral oedema caused by dihydropyridine calcium-channel blockers such as nifedipine which affects up to 10% of the patients. Oedema is usually resolved by stopping the offending drug.
- Oedema is also caused by increased interstitial oncotic pressure as a result of increased capillary permeability to proteins. This can occur as part of a rare complement-deficiency syndrome; with therapeutic use of interleukin 2 in cancer chemotherapy; or in ovarian hyperstimulation syndrome (see p. 1005).

Idiopathic oedema of women

This, by definition, occurs in women without heart failure, hypoalbuminaemia, renal or endocrine disease. Oedema is intermittent and often worse in the premenstrual phase. The condition remits after the menopause. Patients complain of swelling of the face, hands, breasts and thighs, and a feeling of being bloated. Sodium retention during the day and increased sodium excretion during recumbency are characteristic; an abnormal fall in plasma volume on standing caused by increased capillary permeability to proteins may be the cause of this. The oedema may respond to diuretics, but returns when they are stopped. A similar syndrome of diuretic-dependent sodium retention can be caused by abuse of diuretics, for instance as part of an attempt to lose weight; but not all women with idiopathic oedema admit to having taken diuretics, and the syndrome was described before diuretics were introduced for clinical use, so the cause remains unclear.

Local increase in oedema

This does not reflect disturbances of extracellular volume control per se, but can cause clinical confusion. Examples are ankle oedema due to venous damage following thrombosis or surgery, ankle or leg oedema due to immobility, oedema of the arm due to subclavian thrombosis, and facial oedema due to superior vena caval obstruction. Local loss of oncotic pressure may result from increased capillary permeability to proteins, caused by inflammatory mediators such as histamine and interleukins (e.g. a bee sting). Lastly, local loss of lymphatic drainage causes lymphoedema (see p. 1263).

Treatment

The underlying cause should be treated where possible. Heart failure, for example, should be treated, and offending drugs such as NSAIDs withdrawn.

Sodium restriction has only a limited role, but is useful in patients who are resistant to diuretics. Sodium intake can easily be reduced to approximately 100 mmol (2 g) daily; reductions below this are often difficult to achieve without affecting the palatability of food.

Manoeuvres that increase venous return stimulate salt and water excretion by effects on cardiac output and ANP release. This is the rationale for strict bed rest in congestive cardiac failure. Water immersion also causes redistribution of blood towards the central veins, but is seldom of practical use. Venous compression stockings or bandages may help to mobilize oedema in heart failure.

The *mainstay of treatment is the use of diuretic agents*, which increase sodium, chloride and water excretion in the kidney (Table 12.3). These agents act by interfering with membrane ion pumps which are present on numerous cell types; they mostly achieve specificity for the kidney by being secreted into the proximal tubule, resulting in much higher concentrations in the tubular fluid than in other parts of the body.

Table 12.3 Types and clinical uses of diuretics

Class	Major action	Examples	Clinical uses	Potency
Loop diuretics	↓ Na^+–Cl^-–K^+ cotransport in thick ascending limb of loop of Henle	Furosemide Bumetanide Torasemide	Volume overload (CCF, nephrotic syndrome, CKD) ?Acute renal failure SIADH	+++
Thiazide and related diuretics	↓ Na^+–Cl^- cotransport in early distal convoluted tubule	Bendroflumethiazide Chlortalidone Metolazone Indapamide	Hypertension Volume overload (CCF) Hypercalciuria	++
Potassium sparing	↓ Na^+ reabsorption (in exchange for K^+) in collecting duct (principal cells)	Aldosterone antagonists, e.g. spironolactone, eplerenone Others: amiloride, triamterene	Hyperaldosteronism (primary and secondary) Bartter's syndrome Heart failure Cirrhosis with fluid overload Prevention of K^+ deficiency in combination with loop or thiazide	+
Carbonic anhydrase inhibitors	↓ Na^+ HCO_3^- reabsorption in proximal collecting duct ↓ Aqueous humour formation	Acetazolamide	Metabolic alkalosis Glaucoma	+
Vasopressin/ADH Receptor blockers (Aquaretics)	Block V_2 receptor in collecting ducts producing free water diuresis	Lixivaptan Tolvaptan Satavaptan Conivaptan	In trials for heart failure, cirrhosis, SIADH	

CCF, congestive cardiac failure; CKD, chronic kidney disease; SIADH, syndrome of inappropriate antidiuretic hormone secretion

Clinical use of diuretics

Loop diuretics

These potent diuretics are useful in the treatment of any cause of systemic extracellular volume overload. They stimulate excretion of both sodium chloride and water by blocking the sodium–potassium–2-chloride (NKCC2) channel in the thick ascending limb of Henle (Fig. 12.6) and are useful in stimulating water excretion in states of relative water overload. They also act by causing increased venous capacitance, resulting in rapid clinical improvement in patients with left ventricular failure, preceding the diuresis. Unwanted effects include:

- urate retention causing gout
- hypokalaemia
- hypomagnesaemia
- decreased glucose tolerance
- allergic tubulo-interstitial nephritis and other allergic reactions
- myalgia – especially with bumetanide
- ototoxicity (due to an action on sodium pump activity in the inner ear) – particularly with furosemide
- interference with excretion of lithium, resulting in toxicity.

In most situations there is little to choose between the drugs in this class. Bumetanide has a better oral bioavailability than furosemide, particularly in patients with severe peripheral oedema, and has more beneficial effects than furosemide on venous capacitance in left ventricular failure. It can cause severe muscle cramps when used in high doses.

Thiazide diuretics (see p. 738)

These are less potent than loop diuretics. They act by blocking a sodium chloride channel in the *distal convoluted tubule* (Fig. 12.7). They cause relatively more urate retention, glucose intolerance and hypokalaemia than loop diuretics. They interfere with water excretion and may cause hyponatraemia, particularly if combined with amiloride or triamterene. This effect is clinically useful in diabetes insipidus. Thiazides reduce peripheral vascular resistance by mechanisms that are not completely understood but do not appear to depend on their diuretic action, and are widely used in the treatment of essential hypertension. They are also used extensively in mild to moderate cardiac failure. Thiazides reduce calcium excretion. This effect is useful in patients with idiopathic hypercalciuria, but may cause hypercalcaemia. Numerous agents are available, with varying half-lives but little else to choose between them. Metolazone is not dependent for its action on glomerular filtration, and therefore retains its potency in renal impairment.

Potassium-sparing diuretics (see Fig. 12.8)

These are relatively weak and are most often used in combination with thiazides or loop diuretics to prevent potassium depletion. They are of two types.

- *Aldosterone antagonists*, which compete with aldosterone in the *collecting ducts* and reduce sodium absorption, e.g. spironolactone and eplerenone (which has a shorter half-life). Spironolactone is used in patients with heart failure because it significantly reduces the mortality in these patients by antagonizing the fibrotic effect of aldosterone on the heart. Eplerenone is devoid of antiandrogenic or antiprogesterone properties.
- *Amiloride and triamterene* inhibit sodium uptake by blocking epithelial sodium channels in the collecting duct and reduce renal potassium excretion by reducing lumen-negative transepithelial voltage. They are mainly used as potassium sparing agents with thiazide or loop diuretics.

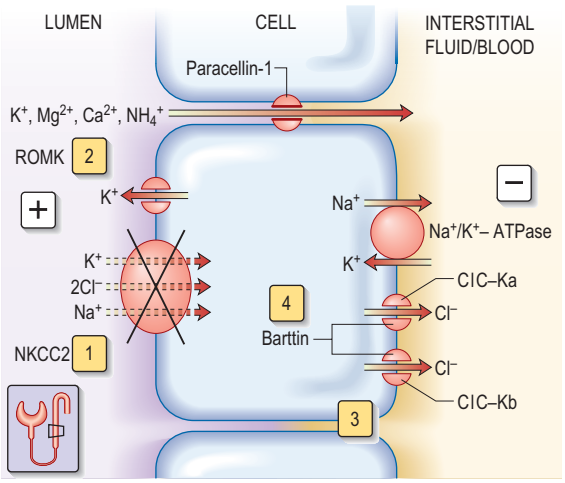


Fig. 12.6 **Transport mechanisms in the thick ascending limb of the loop of Henle.** Sodium chloride is reabsorbed in the thick ascending limb by the *bumetanide-sensitive* sodium–potassium–2-chloride cotransporter (NKCC2). The electroneutral transporter is driven by the low intracellular sodium and chloride concentrations generated by the Na^+/K^+-ATPase and the kidney-specific basolateral chloride channel (ClC-Kb). The availability of luminal potassium is rate-limiting for NKCC2, and recycling of potassium through the ATP-regulated potassium channel (ROMK – renal outer medulla K^+ channel) ensures the efficient functioning of the NKCC2 and generates a lumen-positive transepithelial potential.

Genetic studies have identified putative loss-of-function mutations in the genes encoding NKCC2 [1], ROMK [2], ClC-Kb [3], and barttin [4] in subgroups of patients with **Bartter's syndrome**. In contrast to the normal condition, loss of function of NKCC2 impairs reabsorption of sodium and potassium. Inactivation of the basolateral ClC-Kb and *barttin* reduces transcellular reabsorption of chloride. Loss of function of any of these will reduce the transepithelial potential and thus decrease the driving force for the paracellular reabsorption of cations (K^+, Mg^{2+}, Ca^{2+} and Na^{4+}).

Paracellin-1 is necessary for the paracellular transport of Ca^{2+} and Mg^{2+}. In most patients with **Bartter's syndrome**, urinary calcium excretion is increased. Hypercalcaemia or increased activation of calcium-sensing receptor inactivates ROMK and causes Bartter's syndrome. Ka and Kb, kidney-specific basolateral chloride channel. ROMK, renal outer medullary potassium channel.

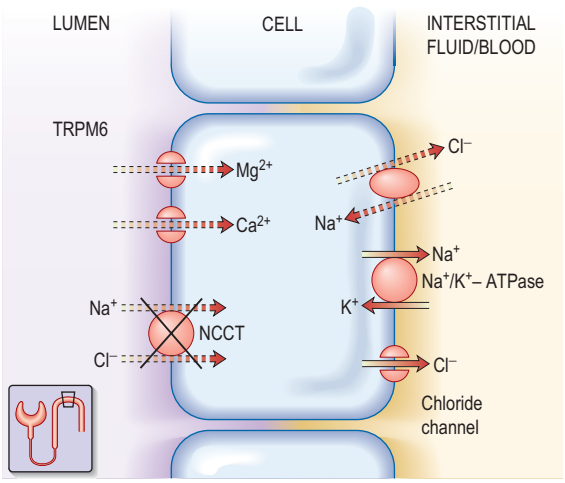


Fig. 12.7 **Transport mechanisms in the distal convoluted tubule.** Under normal conditions, sodium chloride is reabsorbed by the apical *thiazide-sensitive* sodium–chloride cotransporter (NCCT) in the distal convoluted tubule. The electroneutral transporter is driven by the low intracellular sodium and chloride concentrations generated by the Na^+/K^+-ATPase and an, as yet, undefined basolateral chloride channel. In this nephron segment, there is an apical calcium channel and a basolateral sodium-coupled exchanger. Physiological evidence indicates that the mechanisms for the transport of magnesium are similar to those for calcium.

In **Gitelman's syndrome**, putative loss-of-function mutations in the sodium–chloride cotransporter (NCCT (**X**)) lead to decreased reabsorption of sodium chloride and increased reabsorption of calcium. Functional overactivity of NCCT leads to **Gordon's syndrome** by a new mechanism (see text). TRPM6, a member of the transient receptor potential family.

Carbonic anhydrase inhibitors

These are relatively weak diuretics and are seldom used except in the treatment of glaucoma. They cause metabolic acidosis and hypokalaemia.

Aquaretics (vasopressin or ADH receptor blockers)

Vasopressin V_2 receptor antagonists may become very useful agents in the treatment of conditions associated with elevated levels of vasopressin, such as heart failure, cirrhosis and SIADH (see p. 1018). Non-peptide vasopressin V_2 receptor antagonists are efficacious in producing free water diuresis in humans. Studies in patients with heart failure and cirrhosis suggest that such agents will allow normalization of serum osmolality with less water restriction (see p. 663).

Resistance to diuretics

Resistance may occur as a result of:

- poor bioavailability
- reduced GFR, which may be due to decreased circulating volume despite oedema (e.g. nephrotic syndrome, cirrhosis with ascites) or intrinsic renal disease
- activation of sodium-retaining mechanisms, particularly aldosterone.

Management. **Intravenous administration** of diuretics may establish a diuresis. **High doses** of loop diuretics are required to achieve adequate concentrations in the tubule if GFR is depressed. However, the daily dose of furosemide must be limited to a maximum of 2 g for an adult, because of ototoxicity. Intravenous albumin solutions restore plasma oncotic pressure temporarily in the nephrotic syndrome and allow mobilization of oedema but do not increase the natriuretic effect of loop diuretics.

Combinations of various classes of diuretics are extremely helpful in patients with resistant oedema. A loop diuretic plus a thiazide inhibit two major sites of sodium reabsorption; this effect may be further potentiated by addition of a potassium-sparing agent. Metolazone in combination with a loop diuretic is particularly useful in refractory congestive cardiac failure, because its action is less dependent on glomerular filtration. However, this potent combination can cause severe electrolyte imbalance. Both aminophylline and dopamine increase renal blood flow and may be useful in refractory cardiogenic sodium retention. In addition, theophyllines, by inhibiting phosphodiesterase activity in the inner medullary collecting

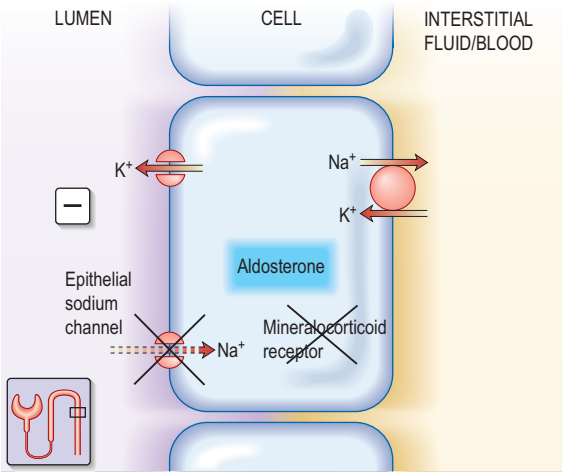


Fig. 12.8 **Aldosterone-regulated transport in the cortical collecting ducts.** Under normal conditions, the epithelial sodium channel is the rate-limiting barrier for the normal entry of sodium from the lumen into the cell. The resulting lumen-negative transepithelial voltage (indicated by the minus sign) drives potassium secretion from the principal cells and proton secretion from the α-intercalated cells (see Fig. 12.12).

In **Liddle's syndrome**, a mutation in the gene encoding the epithelial sodium channel results in persistent unregulated reabsorption of sodium and increased secretion of potassium (not shown).

In **pseudo-hypoaldosteronism type I autosomal recessive**, loss-of-function mutations (X) in this gene inactivate the channel. In the **autosomal dominant variety**, the mutation is in the gene encoding the mineralocorticoid regulation of the activity of the epithelial sodium channel. Either mechanism reduces the activity of the epithelial sodium channel, thus causing salt wasting and decreasing the secretion of potassium and protons.

duct, prolong the action of cyclic GMP (a second messenger of ANP).

Effects on renal function

All diuretics may increase plasma urea concentrations by increasing urea reabsorption in the medulla. Thiazides may also promote protein breakdown. In certain situations diuretics also decrease GFR:

- Excessive diuresis causes volume depletion and prerenal failure.
- Diuretics can cause allergic tubulo-interstitial nephritis.
- Thiazides may directly cause a drop in GFR; the mechanism is complex and not fully understood.

Decreased extracellular volume

Deficiency of sodium and water causes shrinkage both of the interstitial space and of the blood volume and may have profound effects on organ function.

Clinical features

Symptoms are variable. Thirst, muscle cramps, nausea and vomiting, and postural dizziness occur. Severe depletion of circulating volume causes hypotension and impairs cerebral perfusion, causing confusion and eventual coma.

Signs can be divided into those due to loss of interstitial fluid and those due to loss of circulating volume.

Table 12.4 Postural hypotension: some causes of a fall in blood pressure from lying to standing

Decreased circulating volume (hypovolaemia)

Autonomic failure
Diabetes mellitus
Systemic amyloidosis
Shy–Drager syndrome
Parkinson's disease
Ageing

Interference with autonomic function by drugs
Ganglion blockers
Tricyclic antidepressants

Interference with peripheral vasoconstriction by drugs
Nitrates
Calcium-channel blockers
α-Adrenoreceptor blocking drugs

Prolonged bed rest (cardiovascular deconditioning)

- *Loss of interstitial fluid* leads to loss of skin elasticity ('turgor') – the rapidity with which the skin recoils to normal after being pinched. Skin turgor decreases with age, particularly at the peripheries. The turgor over the anterior triangle of the neck or on the forehead is a very useful sign in all ages.
- *Loss of circulating volume* leads to decreased pressure in the venous and (if severe) arterial compartments. Loss of up to 1 L of extracellular fluid in an adult may be compensated for by venoconstriction and may cause no physical signs. Loss of more than this causes the following:

Postural hypotension

Normally the blood pressure rises if a subject stands up, as a result of increased venous return due to venoconstriction (this maintains cerebral perfusion). Loss of extracellular fluid (underfill) prevents this and causes a fall in blood pressure. This is one of the earliest and most reliable signs of volume depletion, as long as the other causes of postural hypotension are excluded (Table 12.4).

Low jugular venous pressure

In hypovolaemic patients, the jugular venous pulsation can be seen only with the patient lying completely flat, or even head down, because the right atrial pressure is lower than 5 cmH_2O.

Peripheral venoconstriction

This causes cold skin with empty peripheral veins, which are difficult to cannulate just when the patient needs intravenous therapy the most! This sign is often absent in sepsis, where peripheral vasodilatation contributes to effective hypovolaemia.

Tachycardia

This is not always a reliable sign. Beta-blockers and other antiarrhythmics may prevent tachycardia, and hypovolaemia may activate vagal mechanisms and actually cause bradycardia.

Causes

Salt and water may be lost from the kidneys, from the gastrointestinal tract, or from the skin. Examples are given in Table 12.5.

In addition, there are a number of situations where signs of volume depletion occur despite a normal or increased body content of sodium and water.

Table 12.5 Causes of extracellular volume depletion

Haemorrhage	**Renal losses**
External	Diuretic use
Concealed, e.g. leaking aortic aneurysm	Impaired tubular sodium conservation
Burns	Reflux nephropathy
Gastrointestinal losses	Papillary necrosis
Vomiting	Analgesic nephropathy
Diarrhoea	Diabetes mellitus
Ileostomy losses	Sickle cell disease
Ileus	

- Septicaemia causes vasodilatation of both arterioles and veins, resulting in greatly increased capacitance of the vascular space. In addition, increased capillary permeability to plasma proteins leads to loss of fluid from the vascular space to the interstitium.
- Diuretic treatment of heart failure or nephrotic syndrome may lead to rapid reduction in plasma volume. Mobilization of oedema may take much longer.
- There may be inappropriate diuretic treatment of oedema (e.g. when the cause is local rather than systemic).

Investigations

Blood tests are in general not helpful in the assessment of extracellular volume. Plasma urea may be raised owing to increased urea reabsorption and, later, to prerenal failure (when the creatinine rises as well), but this is very non-specific. Urinary sodium is low if the kidneys are functioning normally, but is misleading if the cause of the volume depletion involves the kidneys (e.g. diuretics, intrinsic renal disease). Urine osmolality is high in volume depletion (owing to increased water reabsorption), but may also often mislead.

Assessment of volume status is shown in Information box 12.1.

Treatment

The overriding principle is to replace what is missing.

Haemorrhage

The rational treatment of acute haemorrhage is the infusion of a combination of red cells and a plasma substitute or (if unavailable) whole blood. (Chronic anaemia causes salt and water retention rather than volume depletion by a mechanism common to conditions with peripheral vasodilatation.)

Box 12.1 Assessment of volume status

Best achieved by simple clinical observations which you should do yourself. Check:

- Jugular venous pressure
- Central venous pressure both basal and after intravenous fluid challenge (p. 901)
- Serial weights of the patient
- Postural changes in blood pressure
- A chest X-ray.

Loss of plasma

Loss of plasma, as occurs in burns or severe peritonitis, should be treated with human plasma or a plasma substitute (see p. 429).

Loss of water and electrolytes

Loss of water and electrolytes, as occurs with vomiting, diarrhoea, or excessive renal losses, should be treated by replacement of the loss. If possible, this should be with oral water and sodium salts. These are available as slow sodium (600 mg, approximately 10 mmol of each Na^+ and Cl^- per tablet), the usual dose of which is 6–12 tablets per day with 2–3 L of water. It is used in mild or chronic salt and water depletion, such as that associated with renal salt wasting.

Sodium bicarbonate (500 mg, 6 mmol each of Na^+ and HCO_3^- per tablet) is used in doses of 6–12 tablets per day with 2–3 L of water. This is used in milder chronic sodium depletion with acidosis (e.g. chronic renal failure, post-obstructive renal failure, renal tubular acidosis). Sodium bicarbonate is less effective than sodium chloride in causing positive sodium balance. Oral rehydration solutions are described in Box 4.9.

Intravenous fluids are sometimes required (Table 12.6). Rapid infusion (e.g. 1000 mL per hour or even faster) is necessary if there is hypotension and evidence of impaired organ perfusion (e.g. oliguria, confusion); in these situations, plasma expanders (colloids) are often used in the first instance to restore an adequate circulating volume (see p. 904). Repeated clinical assessments are vital in this situation, usually complemented by frequent measurements of central venous pressure (see p. 901, for the management of shock). Severe hypovolaemia induces venoconstriction, which maintains

Table 12.6 Intravenous fluids in general use for fluid and electrolyte disturbances

	Na^+ (mmol/L)	K^+ (mmol/L)	HCO_3^- (mmol/L)	Cl^- (mmol/L)	Indication (see footnote)
Normal plasma values	142	4.5	26	103	
Sodium chloride 0.9%	150	–	–	150	1
Sodium chloride 0.18% + glucose 4%	30	–	–	30	2
Glucose 5% + potassium chloride 0.3%	–	40	–	40	3
Sodium bicarbonate 1.26%	150	–	150	–	4

1. Volume expansion in hypovolaemic patients. Rarely to maintain fluid balance when there are large losses of sodium. The sodium (150 mmol/L) is higher than in plasma and hypernatraemia can result. It is often necessary to add KCl 20–40 mmol/L.
2. Maintenance of fluid balance in normovolaemic, normonatraemic patients.
3. To replace water. Can be given with or without potassium chloride. May be alternated with 0.9% saline as an alternative to (2).
4. For volume expansion in hypovolaemic, acidotic patients alternating with (1). Occasionally for maintenance of fluid balance combined with (2) in salt-wasting, acidotic patients.

venous return; over-rapid correction does not give time for this to reverse, resulting in signs of circulatory overload (e.g. pulmonary oedema) even if a total body extracellular fluid (ECF) deficit remains. In less severe ECF depletion (such as in a patient with postural hypotension complicating acute tubular necrosis), the fluid should be replaced at a rate of 1000 mL every 4–6 hours, again with repeated clinical assessment. If all that is required is avoidance of fluid depletion during surgery, 1–2 L can be given over 24 hours, remembering that surgery is a stimulus to sodium and water retention and that over-replacement may be as dangerous as under-replacement. Regular monitoring by fluid balance charts, bodyweight and plasma biochemistry is mandatory.

Loss of water alone

This causes extracellular volume depletion only in severe cases, because the loss is spread evenly among all the compartments of body water. In the rare situations where there is a true deficiency of water alone, as in diabetes insipidus or in a patient who is unable to drink (after surgery, for instance), the correct treatment is to give water.

If intravenous treatment is required, water is given as 5% dextrose with K^+, because pure water would lead to osmotic lysis of blood cells.

FURTHER READING

Bagshaw SM, Bellomo R, Kellum JA. Oliguria, volume overload, and loop diuretics. *Critical Care Medicine* 2008; **34**(4 Suppl): S172–S178.

Decaux G, Soupart A, Vassart G. Non-peptide arginine-vasopressin antagonists: the vaptans. *Lancet* 2008; **371:** 1624–1632.

Schrier RW. Decreased effective blood volume in edematous disorders: what does this mean? *Journal of the American Society of Nephrology* 2007; **18**(7): 2028–2031.

Schrier RW. The sea within us: disorders of body water homeostasis. *Current Opinion in Investigational Drugs* 2007; **8**(4): 304–311.

Struthers A, Krum H, Williams GH. A comparison of the aldosterone-blocking agents eplerenone and spironolactone. *Clinical Cardiology* 2008; **31**(4): 153–158.

Yool AJ. Functional domains of aquaporin-1: keys to physiology, and targets for drug discovery. *Current Pharmaceutical Design* 2007; **13**(31): 3212–3221.

DISORDERS OF SODIUM CONCENTRATION

These are best thought of as disorders of body water content. As discussed above, sodium content is regulated by volume receptors; water content is adjusted to maintain, in health, a normal osmolality and (in the absence of abnormal osmotically active solutes) a normal sodium concentration. Disturbances of sodium concentration are caused by disturbances of water balance.

HYPONATRAEMIA

Hyponatraemia (Na < 135 mmol/L) is a common biochemical abnormality. The causes depend on the associated changes in extracellular volume:

- hyponatraemia with hypovolaemia (Table 12.7)
- hyponatraemia with euvolaemia (Table 12.8)
- hyponatraemia with hypervolaemia (Table 12.9).

Rarely, hyponatraemia may be a 'pseudo-hyponatraemia'. This occurs in hyperlipidaemia (either high cholesterol or high triglyceride) or hyperproteinaemia where there is a spuriously low measured sodium concentration, the sodium being confined to the aqueous phase but having its concentration expressed in terms of the total volume of plasma. In this situation, plasma osmolality is normal and therefore treatment of 'hyponatraemia' is unnecessary. Artefactual 'hyponatraemia' caused by taking blood from the limb into which fluid of low sodium concentration is being infused should be excluded!

Hyponatraemia with hypovolaemia

This is due to salt loss in excess of water loss; the causes are listed in Table 12.7. In this situation, ADH secretion is initially suppressed (via the hypothalamic osmoreceptors);

Table 12.7 Causes of hyponatraemia with decreased extracellular volume (hypovolaemia)

Extra-renal (urinary sodium < 20 mmol/L)	Kidney (urinary sodium > 20 mmol/L)
Vomiting	Osmotic diuresis (e.g. hyperglycaemia, severe uraemia)
Diarrhoea	Diuretics
Haemorrhage	Adrenocortical insufficiency
Burns	Tubulo-interstitial renal disease
Pancreatitis	Unilateral renal artery stenosis
	Recovery phase of acute tubular necrosis

Table 12.8 Causes of hyponatraemia with normal extracellular volume (euvolaemia)

Abnormal ADH release
Vagal neuropathy (failure of inhibition of ADH release)
Deficiency of adrenocorticotrophic hormone (ACTH) or glucocorticoids (Addison's disease)
Hypothyroidism
Severe potassium depletion

Syndrome of inappropriate antidiuretic hormone (see Table 18.33)

Major psychiatric illness
'Psychogenic polydipsia'
Non-osmotic ADH release?
Anti-depressant therapy

Increased sensitivity to ADH
Chlorpropamide
Tolbutamide

ADH-like substances
Oxytocin
Desmopressin

Unmeasured osmotically active substances stimulating osmotic ADH release
Glucose
Chronic alcohol abuse
Mannitol
Sick-cell syndrome (leakage of intracellular ions)

Table 12.9 Causes of hyponatraemia with increased extracellular volume (hypervolaemia)

Heart failure
Liver failure
Oliguric renal failure
Hypoalbuminaemia

Ask the authors
www.studentconsult.com

but as fluid volume is lost, volume receptors override the osmoreceptors and stimulate both thirst and the release of ADH. This is an attempt by the body to defend circulating volume at the expense of osmolality.

With extrarenal losses and normal kidneys, the urinary excretion of sodium falls in response to the volume depletion, as does water excretion, leading to concentrated urine containing less than 10 mmol/L of sodium. However, in salt-wasting kidney disease, renal compensation cannot occur and the only physiological protection is increased water intake in response to thirst.

Clinical features

With sodium depletion the clinical picture is usually dominated by features of volume depletion (see p. 659). The diagnosis is usually obvious where there is a history of gut losses, diabetes mellitus or diuretic abuse. Examination of the patient is often more helpful than the biochemical investigations, which include plasma and urine electrolytes and osmolality.

Table 12.10 shows the potential daily losses of water and electrolytes from the gut. Losses due to renal or adrenocortical disease may be less easily identified and a urinary sodium concentration of more than 20 mmol/L, in the presence of clinically evident volume depletion, suggests a renal loss.

Treatment

This is directed at the primary cause whenever possible.

In a healthy patient:

- give oral electrolyte–glucose mixtures (see p. 132)
- increase salt intake with slow sodium 60–80 mmol/day.

In a patient with vomiting or severe volume depletion:

- give intravenous fluid with potassium supplements, i.e. 1.5–2 L 5% Dextrose (with 20 mmol K^+) and 1 L 0.9% saline over 24 hours PLUS measurable losses
- correction of acid–base abnormalities is usually not required.

Hyponatraemia with euvolaemia

(see Table 12.8)

This results from an intake of water in excess of the kidney's ability to excrete it (dilutional hyponatraemia) with no change in body sodium content but the plasma osmolality is low. With normal kidney function, dilution hyponatraemia is uncommon even if a patient drinks approximately 1 L per hour. The most common iatrogenic cause is overgenerous infusion of 5% glucose into postoperative patients; in this situation it is exacerbated by an increased ADH secretion in response to stress. Postoperative hyponatraemia is a common clinical problem (almost 1% of patients) with symptomatic hyponatraemia occurring in 20% of these patients. Marathon runners drinking excess water and 'sports drinks' can become hyponatraemic. Premenopausal females are at most risk for developing hyponatraemic encephalopathy postoperatively, with postoperative ADH values in young females being 40 times higher than in young males. To prevent hyponatraemia, avoid using hypotonic fluids postoperatively and administer isotonic saline unless otherwise clinically contraindicated. The serum sodium should be measured daily in any patient receiving continuous parenteral fluid.

Some degree of hyponatraemia is usual in acute oliguric renal failure, while in chronic renal failure it is most often due to ill-given advice to 'push' fluids.

Table 12.10 Average concentrations and potential daily losses of water and electrolytes from the gut

	Na^+ mmol/L	K^+ mmol/L	Cl^- mmol/L	Volume (mL in 24 hours)
Stomach	50	10	110	2500
Small intestine				
Recent ileostomy	120	5	110	1500
Adapted ileostomy	50	4	25	500
Bile	140	5	105	500
Pancreatic juice	140	5	60	2000
Diarrhoea	130	10–30	95	1000–2000+

Clinical features

Dilutional hyponatraemia symptoms are common when this develops *acutely* (<48 hours, often postoperatively). Symptoms rarely occur until the serum sodium is less than 120 mmol/L and are more usually associated with values around 110 mmol/L or lower, particularly when *chronic*. They are principally neurological and are due to the movement of water into brain cells in response to the fall in extracellular osmolality.

Hyponatraemic encephalopathy symptoms and signs include headache, confusion and restlessness leading to drowsiness, myoclonic jerks, generalized convulsions and eventually coma. MRI scan of the brain reveals cerebral oedema but, in the context of electrolyte abnormalities and neurological symptoms, it can help to make a confirmatory diagnosis.

Risk factors for developing hyponatraemic encephalopathy. The brain's adaptation to hyponatraemia initially involves extrusion of blood and CSF, as well as sodium, potassium and organic osmolytes, in order to decrease brain osmolality. Various factors can interfere with successful adaptation. These factors rather than the absolute change in serum sodium predict whether a patient will suffer hyponatraemic encephalopthy.

- *Children* under 16 years are at increased risk due to their relatively larger brain-to-intracranial volume ratio compared to adults.
- *Premenopausal women* are more likely to develop encephalopthy than postmenopausal females and males because of inhibitory effects of sex hormones and the effects of vasopressin on cerebral circulation resulting in vasoconstriction and hypoperfusion of brain.
- *Hypoxaemia* is a major risk factor for hyponatraemic encephalopathy. Patients with hyponatraemia, who develop hypoxia due to either non-cardiac pulmonary oedema or hypercapnic respiratory failure, have a high risk of mortality. Hypoxia is the strongest predictor of mortality in patients with symptomatic hyponatraemia.

Investigations

The cause of hyponatraemia with apparently normal extracellular volume requires investigation:

- Plasma and urine electrolytes and osmolalities. The plasma concentrations of sodium, chloride and urea are low, giving a low osmolality. The urine sodium concentration is usually high and the urine osmolality is typically higher than the plasma osmolality. However, maximal dilution (<50 mosmol/kg) is not always present.
- Further investigations to exclude Addison's disease, hypothyroidism, 'syndrome of inappropriate ADH secretion' (SIADH) and drug-induced water retention, e.g. chlorpropamide.

Remember, potassium and magnesium depletion potentiate ADH release and are causes of diuretic-associated hyponatraemia.

The syndrome of inappropriate ADH secretion is often over-diagnosed. Some causes are associated with a lower set-point for ADH release, rather than completely autonomous ADH release – an example is chronic alcohol abuse.

Treatment

The underlying cause should be corrected where possible.

- Most cases are simply managed by restriction of water intake (to 1000 or even 500 mL per day) with review of diuretic therapy. Magnesium and potassium deficiency must be corrected. In mild sodium deficiency, 0.9% saline given slowly (1 L over 12 hours) is sufficient.
- *Acute onset with symptoms*. Hypertonic saline (3%, 513 mmol/L) is used with extreme care and restricted only for patients in whom there are severe neurological signs, such as fits or coma. It should be given very slowly (not more than 70 mmol per hour), the aim being to increase the serum sodium by 8–10 mmol/L in the first 4 hours, but the absolute change should not exceed 15–20 mmol/L over 48 hours. In general, the plasma sodium should not be corrected to greater than 125–130 mmol/L. 1 mL/kg of 3% sodium chloride will raise the plasma sodium by 1 mmol/L, assuming that total body water comprises 50% of total bodyweight. In some cases furosemide can be used to prevent pulmonary congestion and to increase the rate of sodium correction. Hypertonic saline must **not** be given to patients who are already fluid overloaded because of the risk of acute heart failure; in this situation, 100 mL of 20% mannitol may be infused in an attempt to increase renal water excretion (see below).
- *Chronic/asymptomatic*. If hyponatraemia has developed slowly, as it does in the majority of patients, the brain will have adapted by decreasing intracellular osmolality and the hyponatraemia can be corrected slowly.

A rapid rise in extracellular osmolality, particularly if there is an 'overshoot' to high serum sodium and osmolality, will result in severe shrinking of brain cells and the syndrome of '*central pontine myelinolysis*', which may be fatal. This is diagnosed by the appearance of characteristic hypointense lesions on T1-weighted images and hyperintense on T2-weighted images on MRI; these take up to 2 weeks or longer to appear. The evidence to suggest that these demyelinating lesions are associated with the rate of correction of sodium in hyponatraemic states is controversial. Other factors associated with demyelination are pre-existing hypoxaemia, liver disease and CNS radiation (see above).

Vasopressin V_2 receptor antagonists (see p. 658) which produce a free water diuresis are being used in clinical trials for the treatment of hyponatraemic encephalopathy. Three oral agents, lixivaptan, tolvaptan, and satavaptan, are selective for the V_2 (antidiuretic) receptor, while conivaptan blocks both the V_{1A} and V_2 receptors.

These agents produce a selective water diuresis without affecting sodium and potassium excretion; they raise the plasma sodium concentration in patients with hyponatraemia caused by the SIADH, heart failure, and cirrhosis.

The efficacy of oral tolvaptan in ambulatory patients was demonstrated in a recent report in which patients with hyponatraemia (mean plasma sodium 129 mmol/L) caused by the SIADH, heart failure, or cirrhosis had a sustained rise in plasma sodium to 136 mmol/L for 4 weeks. Oral vasopressin antagonists are likely to be useful in the management of mild to moderate asymptomatic hyponatremia if approved for clinical use. At present, only intravenous conivaptan is available and is approved only for the treatment of euvolaemic hyponatraemia (i.e. SIADH) in some countries. Side-effects include an increase in thirst and urinary frequency but the need for water restriction may still be required during long-term therapy. Careful monitoring is required initially since there is a potential risk of rapid correction of the hyponatraemia if the serum sodium is very low.

Hyponatraemia with hypervolaemia

(see Table 12.9)

The common causes of hyponatraemia due to water excess are shown in Table 12.9. In all these conditions there is usually an element of reduced glomerular filtration rate with avid reabsorption of sodium and chloride in the proximal tubule. This leads to reduced delivery of chloride to the 'diluting' ascending limb of Henle's loop and a reduced ability to generate 'free water', with a consequent inability to excrete dilute urine. This is commonly compounded by the administration of diuretics that block chloride reabsorption and interfere with the dilution of filtrate either in Henle's loop (loop diuretics) or distally (thiazides).

Syndrome of inappropriate ADH secretion

This is described in Chapter 18. There is inappropriate secretion of ADH, causing water retention and hyponatraemia.

HYPERNATRAEMIA

This is much rarer than hyponatraemia and nearly always indicates a water deficit. This may be due to (Table 12.11):

- impaired thirst or impaired conscious state
- pituitary diabetes insipidus (see p. 1017) (failure of ADH secretion)
- nephrogenic diabetes insipidus (failure of response to ADH)
- osmotic diuresis, e.g. hyperglycaemic, hyperosmolar state (see p. 1048)
- excessive loss of water through the skin or lungs.

Excessive administration of sodium may also contribute, for example:

Table 12.11 Causes of hypernatraemia

ADH deficiency
Diabetes insipidus

Iatrogenic
Administration of hypertonic sodium solutions

Insensitivity to ADH (nephrogenic diabetes insipidus)
Lithium
Tetracyclines
Amphotericin B
Acute tubular necrosis

Osmotic diuresis
Total parenteral nutrition
Hyperosmolar hyperglycaemic state

PLUS
Deficient water intake

- excessive reliance on 0.9% (150 mmol/L) saline for volume replacement
- administration of drugs with a high sodium content (e.g. piperacillin)
- use of 8.4% sodium bicarbonate after cardiac arrest.

Hypernatraemia is always associated with increased plasma osmolality, which is a potent stimulus to thirst. None of the above cause hypernatraemia unless thirst sensation is abnormal or access to water limited. For instance, a patient with diabetes insipidus will maintain a normal serum sodium concentration by maintaining a high water intake until an intercurrent illness prevents this. Thirst is frequently deficient in elderly people, making them more prone to water depletion. Hypernatraemia may occur in the presence of normal, reduced or expanded extracellular volume, and does not necessarily imply that total body sodium is increased.

Clinical features

Symptoms of hypernatraemia are non-specific. Nausea, vomiting, fever and confusion may occur. A history of long-standing polyuria, polydipsia and thirst suggests diabetes insipidus. There may be clues to a pituitary cause. A drug history may reveal ingestion of nephrotoxic drugs. Assessment of extracellular volume status guides resuscitation. Mental state should be assessed. Convulsions occur in severe hypernatraemia.

Investigations

Simultaneous urine and plasma osmolality and sodium should be measured. Plasma osmolality is high in hypernatraemia. Passage of urine with an osmolality lower than that of plasma in this situation is clearly abnormal and indicates diabetes insipidus. In pituitary diabetes insipidus, urine osmolality will increase after administration of desmopressin; the drug (a vasopressin analogue) has no effect in nephrogenic diabetes insipidus. If urine osmolality is high this suggests either an osmotic diuresis due to an unmeasured solute (e.g. in parenteral feeding) or excessive extrarenal loss of water (e.g. heat stroke).

Treatment

Treatment is that of the underlying cause, for example:

- in ADH deficiency, replace ADH in the form of desmopressin, a stable non-pressor analogue of ADH
- remember to withdraw nephrogenic drugs where possible and replace water either orally or, if necessary, intravenously.

In *severe* (>170 mmol/L) hypernatraemia, 0.9% saline (150 mmol/L) should be used initially. Avoid too rapid a drop in serum sodium concentration; the aim is correction over 48 hours, as over-rapid correction may lead to cerebral oedema.

In *less severe* (e.g. >150 mmol/L) hypernatraemia, the treatment is 5% dextrose or 0.45% saline; the latter is obviously preferable in hyperosmolar diabetic coma. Very large volumes – 5 L a day or more – may need to be given in diabetes insipidus.

If there is clinical evidence of volume depletion (see p. 661, this implies that there is a sodium deficit as well as a water deficit. Treatment of this is discussed on page 662.

FURTHER READING

Decaux G, Soupart A, Vassart G. Non-peptide arginine-vasopressin antagonists: the vaptans. *Lancet* 2008; **371**: 1624–1632.

Ellison DH, Berl T. Clinical practice. The syndrome of inappropriate antidiuresis. *New England Journal of Medicine* 2007; **356:** 2064–2072.

Liamis G, Kalogirou M, Saugos V et al. Therapeutic approach in patients with dysnatraemias. *Nephrology Dialysis Transplantation* 2006; **21**: 1564–1569.

Liamis G, Milionis H, Elisaf M. A review of drug-induced hyponatremia. *American Journal of Kidney Diseases* 2008; **52**(1): 144–153.

DISORDERS OF POTASSIUM CONCENTRATION

Regulation of serum potassium concentration

The usual dietary intake varies between 80 and 150 mmol daily, depending upon fruit and vegetable intake. Most of the body's potassium (3500 mmol in an adult man) is intracellular. Serum potassium levels are controlled by:

- uptake of K^+ into cells
- renal excretion
- extrarenal losses (e.g. gastrointestinal).

Uptake of potassium into cells is governed by the activity of the Na^+/K^+-ATPase in the cell membrane and by H^+ concentration.

Uptake is *stimulated* by:

- insulin
- β-adrenergic stimulation
- theophyllines.

Uptake is decreased by:

- α-adrenergic stimulation
- acidosis – K^+ exchanged for H^+ across cell membrane
- cell damage or cell death – resulting in massive K^+ release.

Renal excretion of potassium is increased by aldosterone, which stimulates K^+ and H^+ secretion in exchange for Na^+ in the collecting duct (Fig. 12.8). Because H^+ and K^+ are interchangeable in the exchange mechanism, acidosis decreases and alkalosis increases the secretion of K^+. Aldosterone secretion is stimulated by hyperkalaemia and increased angiotensin II levels, as well as by some drugs, and this acts to protect the body against hyperkalaemia and against extracellular volume depletion. The body adapts to dietary deficiency of potassium by reducing aldosterone secretion. However, because aldosterone is also influenced by volume status, conservation of potassium is relatively inefficient, and significant potassium depletion may therefore result from prolonged dietary deficiency.

A number of drugs affect K^+ homeostasis by affecting aldosterone release (e.g. heparin, NSAIDs) or by directly affecting renal potassium handling (e.g. diuretics).

Normally only about 10% of daily potassium intake is excreted in the gastrointestinal tract. Vomit contains around

5–10 mmol/L of K^+, but prolonged vomiting causes hypokalaemia by inducing sodium depletion, stimulating aldosterone, which increases renal potassium excretion. Potassium is secreted by the colon, and diarrhoea contains 10–30 mmol/L of K^+; profuse diarrhoea can therefore induce marked hypokalaemia. Colorectal villous adenomas may rarely produce profuse diarrhoea and K^+ loss.

HYPOKALAEMIA

Causes

The most common causes of chronic hypokalaemia are diuretic treatment (particularly thiazides) and hyperaldosteronism. Acute hypokalaemia is often caused by intravenous fluids without potassium and redistribution into cells. The common causes are shown in Table 12.12.

Table 12.12 Causes of hypokalaemia

Increased renal excretion (urinary K^+ > 20 mmol/day)
Diuretics:
Thiazides
Loop diuretics

Increased aldosterone secretion
Liver failure
Heart failure
Nephrotic syndrome
Cushing's syndrome
Conn's syndrome
ACTH-producing tumours

Exogenous mineralocorticoid
Corticosteroids
Carbenoxolone
Liquorice (potentiates renal actions of cortisol)

Renal disease
Renal tubular acidosis
Types 1 and 2
Renal tubular damage (diuretic phase)
Acute leukaemia
Nephrotoxicity
Amphotericin
Aminoglycosides
Cytotoxic drugs
Release of urinary tract obstruction
Bartter's syndrome
Liddle's syndrome
Gitelman's syndrome

Reduced intake of K^+
Intravenous fluids without K^+
Dietary deficiency

Redistribution into cells
β-Adrenergic stimulation
Acute myocardial infarction
Beta-agonists: e.g. fenoterol, salbutamol
Insulin treatment, e.g. treatment of diabetic ketoacidosis
Correction of megaloblastic anaemia, e.g. B_{12} deficiency
Alkalosis
Hypokalaemic periodic paralysis

Gastrointestinal losses (urinary K^+ < 20 mmol/day)
Vomiting
Severe diarrhoea
Purgative abuse
Villous adenoma
Ileostomy or uterosigmoidostomy
Fistulae
Ileus/intestinal obstruction

Rare causes

These rare causes are discussed in detail because they show the mechanisms of how diuretics can affect the kidney.

Bartter's syndrome (similar to loop diuretics)

This consists of metabolic alkalosis, hypokalaemia, hypercalciuria, normal blood pressure, and an elevated plasma renin and aldosterone. The primary defect in this disorder is an impairment in sodium and chloride reabsorption in the thick ascending limb of the loop of Henle (Fig. 12.6). Mutation in the genes encoding either the sodium–potassium–2-chloride cotransporter (NKCC2), the ATP-regulated renal outer medullary potassium channel (ROMK) or kidney-specific basolateral chloride channels (ClC-Kb) – Bartter's types I, II and III respectively – causes loss of function of these channels, with consequent impairment of sodium and chloride reabsorption. There is also an increased intrarenal production of prostaglandin E_2 which is secondary to sodium and volume depletion, hypokalaemia and the consequent neurohumoral response rather than a primary defect. PE_2 causes vasodilatation and may explain why the blood pressure remains normal.

Barttin, a beta-subunit for ClC-Ka and ClC-Kb chloride channels, is encoded by the BSND (Bartter's syndrome with sensorineural deafness) gene. Loss of function mutations causes type IV Bartter's syndome associated with sensorineural deafness and renal failure. Barttin co-localizes with a subunit of the chloride channel in basolateral membranes of the renal tubule and inner ear epithelium. It appears to mediate chloride exit in the thick ascending limb (TAL) of the loop of Henle and chloride recycling in potassium-secreting strial marginal cells in the inner ear. A very rare variant of type IV is a disorder with an impairment of both chloride channels (ClC-Ka and ClC-Kb) producing the same phenotypic defects.

A gain of function mutation of the calcium sensing receptor (CaSR) which leads to autosomal dominant hypocalcaemia has also been recognized in Bartter's syndrome. In the kidney, the CaSR is expressed mainly in the basolateral membrane of cortical TAL. Activation of CaSR by high calcium or magnesium or by gain-of-function mutation triggers intracellular signalling, including release of arachidonic acid and inhibition of adenylate cyclase. Both actions result in inhibition of ROMK activity, which in turn leads to reduction in the lumen-positive electrical potential and transcellular absorption of calcium. This effect of CaSR explains why patients with mutations in this receptor may present with both hypocalcaemia, hypercalciuria and renal wasting of NaCl, resulting in a Bartter-like syndrome.

In summary, these defects in sodium chloride transport are thought to initiate the following sequence, which is almost identical to that seen with chronic ingestion of a loop diuretic. The initial salt loss leads to mild volume depletion, resulting in activation of the renin–angiotensin–aldosterone system. The combination of hyperaldosteronism and increased distal flow (owing to the reabsorptive defect) enhances potassium and hydrogen secretion at the secretory sites in the collecting tubules, leading to hypokalaemia and metabolic alkalosis.

Diagnostic pointers include high urinary potassium and chloride despite low serum values as well as increased plasma renin (NB: in primary aldosteronism, renin levels are low). Hyperplasia of the juxtaglomerular apparatus is seen on renal biopsy (careful exclusion of diuretic abuse is necessary). Hypercalciuria is a common feature but magnesium wasting, though rare, also occurs. *Treatment* is with combinations of potassium supplements, amiloride and indometacin.

Gitelman's syndrome (similar to thiazide diuretics)

Gitelman's syndrome is a phenotype variant of Bartter's syndrome characterized by hypokalaemia, metabolic alkalosis, hypocalciuria, hypomagnesaemia, normal blood pressure, and elevated plasma renin and aldosterone. There are striking similarities between the Gitelman syndrome and the

biochemical abnormalities induced by chronic thiazide diuretic administration. Thiazides act in the distal convoluted tubule to inhibit the function of the apical sodium–chloride cotransporter (NCCT) (Fig. 12.7). Analysis of the gene encoding the NCCT has identified loss of function mutations in Gitelman's syndrome.

Like Bartter's syndrome, defective NCCT function leads to increased solute delivery to the collecting duct, with resultant solute wasting, volume contraction and an aldosterone-mediated increase in potassium and hydrogen secretion. Unlike Bartter's syndrome, the degree of volume depletion and hypokalaemia is not sufficient to stimulate prostaglandin E_2 production. Impaired function of NCCT is predicted to cause hypocalciuria, as does thiazide administration. Impaired sodium reabsorption across the apical membrane, coupled with continued intracellular chloride efflux across the basolateral membrane, causes the cell to become hyperpolarized. This in turn stimulates calcium reabsorption via apical, voltage-activated calcium channels. Decreased intracellular sodium also facilitates calcium efflux via the basolateral sodium–calcium exchanger. The mechanism for urinary magnesium losses is not known. *Treatment* consists of potassium and magnesium supplementation ($MgCl_2$) and a potassium-sparing diuretic. Volume resuscitation is usually not necessary, because patients are not dehydrated. Elevated prostaglandin E_2 does not occur (see above) and, therefore, NSAIDs are not indicated in this disorder.

Liddle's syndrome

This is characterized by potassium wasting, hypokalaemia and alkalosis, but is associated with low renin and aldosterone production, and high blood pressure. There is a mutation in the gene encoding for the amiloride-sensitive epithelial sodium channel in the distal tubule/collecting duct. This leads to constitutive activation of the epithelial sodium channel, resulting in excessive sodium reabsorption with coupled potassium and hydrogen secretion. Unregulated sodium reabsorption across the collecting tubule results in volume expansion, inhibition of renin and aldosterone secretion and development of low renin hypertension (Fig. 12.8).

Therapy consists of sodium restriction along with amiloride or triamterene administration. Both are potassium-sparing diuretics which directly close the sodium channels. The mineralocorticoid antagonist spironolactone is ineffective, since the increase in sodium-channel activity is not mediated by aldosterone.

Hypokalaemic periodic paralysis (see p 1182)

This condition may be precipitated by carbohydrate intake, suggesting that insulin-mediated potassium influx into cells may be responsible. This syndrome also occurs in association with hyperthyroidism in Chinese patients.

Clinical features

Hypokalaemia is usually asymptomatic, but severe hypokalaemia (<2.5 mmol) causes muscle weakness. Potassium depletion may also cause symptomatic hyponatraemia (see p. 661).

Hypokalaemia is associated with an increased frequency of atrial and ventricular ectopic beats. This association may not always be causal, because adrenergic activation (for instance after myocardial infarction) causes both hypokalaemia and increased cardiac irritability. Hypokalaemia in patients without cardiac disease is unlikely to lead to serious arrhythmias.

Hypokalaemia seriously increases the risk of digoxin toxicity by increasing binding of digoxin to cardiac cells, potentiating its action, and decreasing its clearance.

Chronic hypokalaemia is associated with interstitial renal disease, but the pathogenesis is not completely understood.

Treatment

The underlying cause should be identified and treated where possible. Table 12.13 shows some examples.

Acute hypokalaemia may correct spontaneously. In most cases, withdrawal of oral diuretics or purgatives, accompanied by the oral administration of potassium supplements in the form of slow-release potassium or effervescent potassium, is all that is required. Intravenous potassium replacement is required only in conditions such as cardiac arrhythmias, muscle weakness or severe diabetic ketoacidosis. When using intravenous therapy in the presence of poor renal function, replacement rates <2 mmol per hour should be used only, with hourly monitoring of serum potassium and ECG changes. Ampoules of potassium should be thoroughly mixed in 0.9% saline; do not use a glucose solution as this would make hypokalaemia worse.

The treatment of adrenal disorders is described on page 1013.

Failure to correct hypokalaemia may be due to concurrent hypomagnesaemia. Serum magnesium should be measured and any deficiency corrected.

Hyperkalaemia

Causes

Acute self-limiting hyperkalaemia occurs normally after vigorous exercise and is of no pathological significance. Hyperkalaemia in all other situations is due either to increased release from cells or to failure of excretion (Table 12.14). The most common causes are renal impairment and drug interference with potassium excretion. The combination of ACE inhibitors with potassium-sparing diuretics or NSAIDs is particularly dangerous.

Table 12.13 Treatment of hypokalaemia

Cause	Treatment
Dietary deficiency	Increase intake of fresh fruit/vegetables or oral potassium supplements (20–40 mmol daily). (Potassium supplements can cause gastrointestinal irritation)
Hyperaldosteronism, e.g. cirrhosis, thiazide therapy	Spironolactone/eplerenone Co-prescription of a potassium-sparing diuretic with a similar onset and duration of action
Intravenous fluid replacement	Add 20 mmol of K^+ per litre of fluid with monitoring

Table 12.14 Causes of hyperkalaemia

Decreased excretion
Renal failure*
Drugs:*
- Amiloride
- Triamterene
- Spironolactone/eplerenone
- ACE inhibitors/ACE blockers*
- NSAIDs

Ciclosporin treatment
Heparin treatment
Aldosterone deficiency
Hyporeninaemic hypoaldosteronism (RTA type 4)
Addison's disease
Acidosis*
Gordon's syndrome

Increased release from cells (decreased Na^+/K^+-ATPase activity)
Acidosis
Diabetic ketoacidosis
Rhabdomyolysis/tissue damage
Tumour lysis
Succinylcholine (amplified by muscle denervation)
Digoxin poisoning
Vigorous exercise (α-adrenergic; transient)

Increased extraneous load
Potassium chloride
Salt substitutes
Transfusion of stored blood*

Spurious
Increased in vitro release from abnormal cells
- Leukaemia
- Infectious mononucleosis
- Thrombocytosis

Familial pseudohyperkalaemia, e.g. haemolysis in syringe
Increased release from muscles
- Vigorous fist clenching during phlebotomy

*Common causes

Rare causes

Hyporeninaemic hypoaldosteronism

This is also known as type 4 renal tubular acidosis (see p. 676). Hyperkalaemia occurs because of acidosis and hypoaldosteronism.

Pseudo-hypoaldosteronism type 1 (autosomal recessive and dominant types)

This is a disease of infancy, apparently due to resistance to the action of aldosterone. It is characterized by hyperkalaemia and evidence of sodium wasting (hyponatraemia, extracellular volume depletion). Autosomal recessive forms result from loss of function because of mutations in the gene for epithelial sodium channel activity (opposite to Liddle's syndrome). This disorder involves multiple organ systems and is especially marked in the neonatal period. With aggressive salt replacement and control of hyperkalaemia, these children can survive and the disorder appears to become less severe with age. The autosomal dominant type is due to mutations affecting the mineralocorticoid receptor (Fig. 12.8). These patients present with salt wasting and hyperkalaemia but do not have other organ-system involvement.

Hyperkalaemic periodic paralysis (see p. 1182)

This is precipitated by exercise, and is caused by an autosomal dominant mutation of the skeletal muscle sodium channel gene.

Gordon's syndrome (familial hyperkalaemic hypertension, pseudohypoaldosteronism type 2)

This appears to be a mirror image of Gitelman's syndrome (see p. 665), in which primary renal retention of sodium causes hypertension, volume expansion, low renin/aldosterone, hyperkalaemia and metabolic acidosis. There is also an increased sensitivity of sodium reabsorption to thiazide diuretics, suggesting that the thiazide-sensitive sodium–chloride cotransporter (NCCT) is involved. Genetic analyses, however, have excluded abnormalities in NCCT. The involvement of two loci on chromosomes 1 and 12 and further genetic heterogeneity has also been found. These genes do not correspond to ionic transporters but to unexpected proteins, WNK (With No lysine Kinase) 1 and WNK 4, which are two closely related members of a novel serine – threonine kinase family. WNK 4 normally inhibits NCCT by preventing its membrane translocation from the cytoplasm. Loss of function mutation in WNK 4 results in escape of NCCT from normal inhibition and its overactivity as seen from the patient's phenotype. WNK 1 is an inhibitor of WNK 4 and in some patients with Gordon's syndrome, gain of function mutation in WNK 1 results in functional deficiency of WNK 4 and overactivity of NCCT.

Suxamethonium and other depolarizing muscle relaxants

These cause release of potassium from cells. Induction of muscle paralysis during general anaesthesia may result in a rise of plasma potassium of up to 1 mmol/L. This is not usually a problem unless there is pre-existing hyperkalaemia.

Clinical features

Serum potassium of greater than 7.0 mmol/L is a medical emergency and is associated with ECG changes (Fig. 12.9). Severe hyperkalaemia may be asymptomatic and may predispose to sudden death from asystolic cardiac arrest. Muscle weakness is often the only symptom, unless (as is commonly the case) the hyperkalaemia is associated with metabolic acidosis, causing Kussmaul respiration. Hyperka-

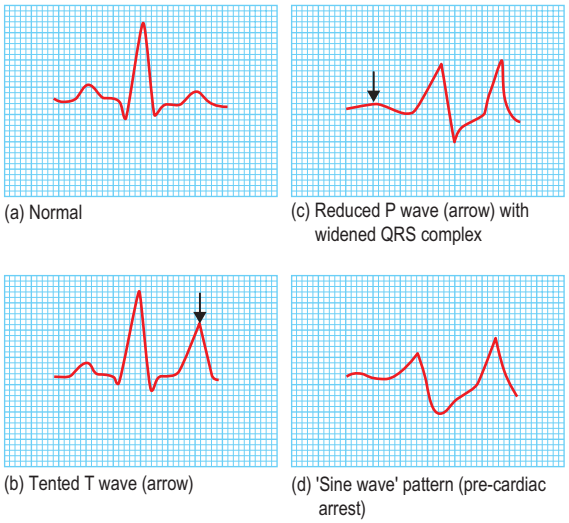


Fig. 12.9 Progressive ECG changes with increasing hyperkalaemia.

laemia causes depolarization of cell membranes, leading to decreased cardiac excitability, hypotension, bradycardia, and eventual asystole.

Treatment

Treatment for severe hyperkalaemia requires both urgent measures to save lives and maintenance therapy to keep potassium down, as summarized in Emergency box 12.1. The cause of the hyperkalaemia should be found and treated.

High potassium levels are cardiotoxic as they inactivate sodium channels. Divalent cations, e.g. calcium, restore the voltage dependability of the channels. *Calcium ions* protect the cell membranes from the effects of hyperkalaemia but do not alter the potassium concentration.

Supraphysiological insulin (20 units) drives potassium into the cell and lowers plasma potassium by 1 mmol in 60 minutes, but must be accompanied by glucose to avoid hypoglycaemia. Regular measurements of blood glucose for at least 6 hours after use of insulin should be performed and extra glucose must be available for immediate use. The use of glucose alone in non-diabetic patients, to stimulate endogenous insulin release, does not produce the high levels of insulin required and therefore is not recommended.

Intravenous or nebulized salbutamol (10–20 mg) has not yet found widespread acceptance and may cause disturbing muscle tremors at the doses required.

Correction of acidosis with hypertonic (8.4%) sodium bicarbonate causes volume expansion and should not be used; 1.26% is used with severe acidosis (pH < 6.9) (see p. 678). Gastric aspiration will remove potassium and leads to alkalosis.

Ion-exchange resins (polystyrene sulphonate resins) are used as maintenance therapy to keep potassium down after emergency treatment. They make use of the ion fluxes which occur in the gut to remove potassium from the body, and are the only way short of dialysis of removing potassium from the body. They may cause fluid overload (resonium contains Na^+) or hypercalcaemia (calcium resonium). Resins do not appear to significantly enhance the excretion of potassium beyond the effect of diarrhoea induced by osmotic or secretory cathartics.

In *general*, all of these measures are simply ways of buying time either to correct the underlying disorder or to arrange removal of potassium by dialysis, which is the definitive treatment for hyperkalaemia in renal failure.

Emergency Box 12.1

Correction of severe hyperkalaemia

IMMEDIATE

ECG monitor and i.v. access

Protect myocardium

10 mL of 10% calcium gluconate i.v. over 5 min
Effect is temporary but dose can be repeated after 15 min

Drive K^+ into cells

Insulin 10 units + 50 mL of 50% glucose i.v. over 10–15 min followed by regular checks of blood glucose and plasma K^+
Repeat as necessary:
and/or correction of severe acidosis (pH < 6.9) – infuse $NaHCO_3$ (1.26%)
and/or salbutamol 0.5 mg in 100 mL of 5% glucose over 15 min (rarely used)

LATER
Deplete body K^+ (to decrease plasma K^+ over next 24 h)

Polystyrene sulphonate resins:
15 g orally up to three times daily with laxatives
30 g rectally followed 9 hours later by an enema
Haemodialysis or peritoneal dialysis if the above fails

FURTHER READING

Alfonzo AV, Isles C, Geddes C et al. Potassium disorders – clinical spectrum and emergency management. *Resuscitation* 2006; **70**: 10–25.

Evans KJ, Greenberg A. Hyperkalemia: a review. *Journal of Intensive Care Medicine* 2005; **20**: 272–290.

Flatman PW. Contransporters, WNKs and hypertension: an update. *Current Opinion in Nephrology and Hypertension* 2008; **17**(2): 186–192.

Kramer BK, Bergler T, Stoelcker B et al. Mechanisms of disease: the kidney-specific chloride channels CICKA and CICKB, the Barttin subunit, and their clinical relevance. *Nature Clinical Practice. Nephrology* 2008; **4**(1): 38–46.

Lin SH, Halperin ML. Hypokalemia: a practical approach to diagnosis and its genetic basis. *Current Medicinal Chemistry* 2007; **14**: 1551–1565.

DISORDERS OF MAGNESIUM CONCENTRATION

Plasma magnesium levels are normally maintained within the range 0.7–1.1 mmol/L (1.4–2.2 meq/L). The average daily magnesium intake is 15 mmol, one-third of which is absorbed, principally in the small bowel. In the healthy adult, there is no net gain or loss of magnesium from bone, so that balance is achieved by the urinary excretion of the net magnesium absorbed.

Primary disturbance of magnesium balance is uncommon; hypo- or hypermagnesaemia usually developing on a background of more obvious fluid and electrolyte disturbances. Disturbance in magnesium balance should always be suspected in association with other fluid and electrolyte disturbances when the patient develops unexpected neurological signs or symptoms.

Renal handling of magnesium

Magnesium transport differs from that of most other ions in that the proximal tubule is not the major site of reabsorption with only 15–25% of the filtered magnesium being reabsorbed here passively. The major site of magnesium transport is the cortical thick ascending limb (cTAL) of the loop of Henle, where 60–70% of the filtered load is reabsorbed (Fig. 12.6). This transport is passive and paracellular, driven by the lumen positive electrochemical gradient, characteristic of this segment. Conversely, magnesium reabsorption in the distal convoluted tubule (DCT) is transcellular and active (Fig. 12.7), mediated by TRPM6, a member of the transient receptor potential channel family. TRPM6 is a Mg^{2+}-permeable channel that is also expressed in the luminal membrane of the intestinal epithelium. Inactivating mutations of TRPM6 thus causes a combination of impaired gut absorption of Mg^{2+} and renal wasting as seen in hypomagnesemia with secondary hypocalcemia (HSH). HSH is a rare autosomal recessive disease typified by low serum Mg^{2+} levels and high urinary fractional Mg^{2+} excretion. The reabsorption rate in the DCT (5–15%) is much lower than in the cTAL, but it defines the final urinary excretion, as there is no significant reabsorption in the collecting duct. Three to five per cent of filtered magnesium is finally excreted in the urine. cTAL magnesium reabsorption varies with changes in the plasma magnesium concentration, which is the main physiological regulator of urinary magnesium excretion. Hypermagnesaemia inhibits loop transport, while hypomagnesaemia stimulates magnesium transport. Hypercalcaemia inhibits magnesium loop transport by an unknown mechanism.

Another factor that can influence loop magnesium transport is the rate of sodium chloride reabsorption. This effect is relatively unimportant in normal subjects, but decreased reabsorption and magnesium wasting can be induced by the administration of a loop diuretic (Fig. 12.6).

HYPOMAGNESAEMIA

This most often develops as a result of deficient intake, defective gut absorption, or excessive gut or urinary loss (Table 12.15). It can also occur with acute pancreatitis, possibly owing to the formation of magnesium soaps in the areas of fat necrosis. Calcium deficiency usually develops with hypomagnesaemia. The serum magnesium is usually <0.7 mmol/L (1.4 mEq/L).

Isolated dominant hypomagnesaemia (IDH)

IDH follows an autosomal dominant mode of inheritance, presenting as hypomagnesaemia and hypocalciuria. It has similarities with Gitelman's syndrome but lacks hypokalaemia and metabolic alkalosis. The FXYD2 gene, encoding the gamma-subunit of Na^+/K^+-ATPase which is mainly expressed in TAL and DCT, is mutated, resulting in the inhibition of Na^+/K^+-ATPase activity and limiting the amount of potassium entry into the cell. Closing of potassium-sensitive apical magnesium channels causes reduced magnesium absorption and hypomagnesaemia in patients with IDH.

Familial hypomagnesaemia, hypercalciuria and nephrocalcinosis (FHHNC)

This disorder is characterized by excessive renal magnesium and calcium wasting. In addition, affected individuals develop bilateral nephrocalcinosis and progressive renal failure. Patients also have elevated PTH levels, which precedes any reduction in GFR. A substantial proportion of patients show incomplete distal renal tubular acidosis, hypocitraturia and hyperuricaemia. Extrarenal involvement such as myopia, nystagmus, chorioretinitis has been reported. Based on clinical observation and clearance studies, the main defect in magnesium and calcium reabsorption lies in cTAL. Ten different mutations have been identified in a novel gene which encodes for paracellin-1 and claudin 19, members of the claudin family of tight junction proteins (see p. 25).

Table 12.15 Causes of hypomagnesaemia

Decreased magnesium absorption	**Gut losses**
Malabsorption (severe)	Prolonged nasogastric suction
Malnutrition	Excessive purgation
Alcohol excess	Gastrointestinal/biliary fistulae
Increased renal excretion	Severe diarrhoea
Drugs:	**Miscellaneous**
loop diuretics	Acute pancreatitis
thiazide diuretics	Inherited tubular wasting
digoxin	Isolated dominant hypomagnesaemia
Diabetic ketoacidosis	Familial hypomagnesaemia, hypercalciuria and nephrocalcinosis
Gitelman's syndrome (p. 665)	Hypomagnesaemia with secondary hypocalcaemia (HSH)
Hyperaldosteronism	
SIADH	
Alcohol excess	
Hypercalciuria	
1,25-(OH)-vitamin D_3 deficiency	
Drug toxicity:	
amphotericin	
aminoglycosides	
cisplatin	
ciclosporin	

SIADH, syndrome of inappropriate antidiuretic hormone secretion

Clinical features

Symptoms and signs (indicating a deficit of 0.5–1 mmol/kg) include irritability, tremor, ataxia, carpopedal spasm, hyperreflexia, confusional and hallucinatory states, and epileptiform convulsions. An ECG may show a prolonged QT interval, broad flattened T waves, and occasional shortening of the ST segment.

Treatment

This involves the withdrawal of precipitating agents such as diuretics or purgatives. If symptomatic (or with hypocalcaemia), give a parenteral infusion of 50 mmol of magnesium chloride in 1 L of 5% dextrose or other isotonic fluid over 12–24 hours. This should be repeated daily and continued for 2 days after normal plasma levels have been achieved.

HYPERMAGNESAEMIA

This primarily occurs in patients with acute or chronic renal failure given magnesium-containing laxatives or antacids. It can also be induced by magnesium-containing enemas. Mild hypermagnesaemia may occur in patients with adrenal insufficiency. Causes are given in Table 12.16.

Clinical features

Symptoms and signs relate to neurological and cardiovascular depression, and include weakness with hyporeflexia proceeding to narcosis, respiratory paralysis and cardiac conduction defects. Symptoms usually develop when the plasma magnesium level exceeds 2 mmol/L (4 mEq/L).

Treatment

Treatment requires withdrawal of any magnesium therapy. An intravenous injection of 10 mL of calcium gluconate 10% (2.25 mmol calcium) is given to antagonize the effects of hypermagnesaemia, along with dextrose and insulin (as for hyperkalaemia) to lower the plasma magnesium level. Dialysis may be required in patients with severe renal failure.

Table 12.16 Causes of hypermagnesaemia

Impaired renal excretion
Chronic renal failure
Acute renal failure
Increased magnesium intake
Purgatives, e.g. magnesium sulphate
Antacids, e.g. magnesium trisilicate
Haemodialysis with high [Mg^{2+}] dialysate

DISORDERS OF PHOSPHATE CONCENTRATION

Phosphate forms an essential part of most biochemical systems, from nucleic acids downwards. The regulation of plasma phosphate level is closely linked to calcium.

About 80% of all body phosphorus is within bone, plasma phosphate normally ranging from 0.80 to 1.40 mmol/L (2.5 to 4.5 mg/dL). Phosphate reabsorption from the kidney (85% from the proximal tubule) is decreased by parathyroid hormone (PTH), mediated by a cyclic AMP-dependent mechanism; thus hyperparathyroidism is associated with low plasma levels of phosphate. Other factors that are known to control phosphate reabsorption in the proximal tubule are 1,25-dihydroxyvitamin D_3, sodium delivery to the proximal tubule, serum concentrations of calcium, bicarbonate, carbon dioxide tension, glucose, alanine, serotonin, dopamine and sympathetic activity. Osteoblast secreted phosphaturic factors (phosphatonins) such as fibroblast growth factor 23 (FGF-23), matrix extracellular phosphoglycoprotein (MEPG) and frizzled related protein 4 (FRP-4) are being investigated for a normal physiological role in phosphate homeostasis. FGF 23 binds to its receptors in the kidney and causes phosphaturia and also regulates vitamin D activation by 1 alpha hydroxylase (CYP27B1) and 24 hydroxylase (CYP24A1) enzymes. Moreover, FGF 23 requires *Klotho* (130 kDa anti-ageing protein predominantly expressed in the kidney) as a cofactor for its activity. Loss of function mutation in either FGF 23 or Klotho results in a similar phenotype of shortened life span, premature ageing (p. 37) including hyperphosphataemia and paradoxically increased 1-25 vitamin D levels.

HYPOPHOSPHATAEMIA

Significant hypophosphataemia (<0.4 mmol/L or <1.25 mg/dL) occurs in a number of clinical situations, owing to redistribution into cells, to renal losses, or to decreased intake (Table 12.17). Clinical features include:

- muscle weakness, e.g. diaphragmatic weakness, decreased cardiac contractility, skeletal muscle rhabdomyolysis
- a left-shift in the oxyhaemoglobin dissociation curve (reduced 2,3-bisphosphoglycerate (2,3-BPG)) and rarely haemolysis
- confusion, hallucinations, and convulsions.

Mild hypophosphataemia often resolves without specific treatment. However, diaphragmatic weakness may be severe in acute hypophosphataemia, and may impede weaning a patient from a ventilator. Interestingly, chronic hypophosphataemia (in X-linked hypophosphataemia) is associated with normal muscle power.

Table 12.17 Causes of hypophosphataemia

Redistribution	**Renal losses**
Respiratory alkalosis	Hyperparathyroidism
Treatment of diabetic ketoacidosis (insulin drives phosphate into cells)	Renal tubular defects
	Diuretics
	Hypophosphataemic rickets
Refeeding (particularly with carbohydrate) after fasting or starvation	Vitamin D-dependent rickets Types I and II
	Tumour induced
After parathyroidectomy (hungry bone disease)	Autosomal dominant
	X-linked dominant
Decreased intake/absorption	Dent's disease
Dietary	
Malabsorption	
Vomiting	
Gut phosphate binders, e.g. sevelamer	
Vitamin D deficiency	
Alcohol withdrawal	

Causes

Hypophosphataemia can be part of osteomalacia and rickets due to vitamin D deficiency either dietary or genetic and is usually accompanied by hypocalcaemia (calcipenic).

Vitamin D-dependent rickets type I

Also known as pseudovitamin D-deficient rickets, this is caused by 1-alpha hydroxylase deficiency due to inactivating mutations in its gene. It can be corrected with high daily doses of vitamin D. This condition manifests clinically in the first year of life with severe hypocalcaemia often complicated by tetany, moderate hypophospataemia and enamel hypoplasia. The characteristic biochemical findings are normal serum levels of 25 hydroxyvitamin D and low values of 1-25 hydroxyvitamin D. The treatment of choice is replacement therapy with calcitriol. The aim of therapy is to maintain serum levels of calcium, phosphorus, and alkaline phosphatase within normal limits. Close monitoring is mandatory because of possible side-effects of calcitriol therapy including hypercalciuria, hypercalcaemia, nephrocalcinosis, and intraocular calcifications.

Vitamin D-dependent rickets type II

This is a form of vitamin D resistance and is now known as hereditary vitamin D-resistant rickets. It is an autosomal recessive disorder and is usually caused by loss of function mutations in the gene encoding the vitamin D receptor. The clinical manifestations vary widely, depending upon the type of mutation within the vitamin D receptor and the amount of residual vitamin D receptor activity. Affected children usually develop rickets within the first 2 years of life with alopecia in two-thirds of cases which is due to lack of vitamin D receptor action within keratinocytes. The treatment involves a therapeutic trial of calcitriol and calcium supplementation. The individual response is difficult to predict because the severity of the receptor defect varies among patients. Long-term infusion of calcium into a central vein is a possible alternative for severely resistant patients. This infusion must be maintained for many months. Oral calcium therapy may be sufficient once radiographic healing has been observed.

Decreased reabsorption of phosphate

This condition also occurs in patients with *tumour-induced osteomalacia* (TIO), *X-linked dominant hypophosphataemic rickets* (XLR), and *autosomal-dominant hypophosphataemic rickets* (ADHR). These syndromes have similar biochemical and osseous phenotypes. Patients have osteomalacia or rickets, reduced tubular phosphate reabsorption, hypophos-

phataemia, normal or low serum calcium, normal PTH and PTH-related protein concentrations, and normal or low 1,25-dihydroxyvitamin D_3. Urinary cyclic AMP levels are generally in the normal range.

In *TIO*, there is excessive production of phosphatonins (which are normally produced by osteoblasts and act as hormones by acting on kidneys and regulating phosphate absorption and vitamin D activation), e.g. fibroblast growth factor 23 (FGF-23), matrix extracellular phosphoglycoprotein (MEPG), and frizzled related protein 4 (FRP-4). These cannot be degraded by normal concentrations of PHEX (phosphate regulating gene with homologies to endopeptidases on the X chromosome). This results in net excess of inhibitors of the sodium–phosphate cotransporter in the proximal tubule, and phosphaturia.

In *ADHR*, FGF-23 is mutated so that it is resistant to PHEX proteolysis. In XLR, mutations in PHEX prevent binding to FGF-23 and FRP-4, resulting in a net relative excess of phosphatonins.

Normal adaptive increases in 1,25-dihydroxyvitamin D_3 synthesis in response to low phosphate levels do not occur in TIO, ADHR and XLR, aggravating phosphaturia (Fig. 12.10).

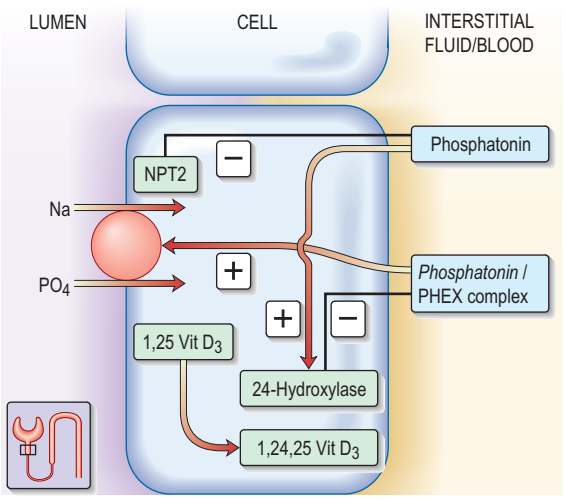


Fig. 12.10 **PHEX regulation of phosphate transport and vitamin D metabolism in the proximal tubule.** The hormone phosphatonin is an inhibitor of sodium–phosphate cotransporter (NPT2). When bound to PHEX (phosphate-regulating gene with homologies to endopeptidase on the X chromosome) it is bioinactive and is thought to stimulate phosphate reabsorption via the sodium–phosphate cotransporter (NPT2) and to retard the degradation of 1,25-dihydroxyvitamin D by downregulating 24-hydroxylase activity.

Over-production of phosphatonins occurs in tumour-induced osteomalacia, gain of function mutation as in autosomal dominant hypophosphataemic rickets and loss of function mutation in PHEX as X-linked hypophosphataemia. This causes renal phosphate wasting, as well as increased 1,25-dihydroxyvitamin D degradation. Modified from Rowe PSN. The role of the PHEX gene (PEX) in families with X-linked hypophosphataemic rickets. *Current Opinion in Nephrology and Hypertension* 1998; **7**(4): 367–376 with permission from Lippincott Williams and Wilkins.

Dent's disease

Dent's disease is the generally accepted name for a group of hereditary tubular disorders including X-linked recessive nephrolithiasis with renal failure, X-linked recessive hypophosphataemic rickets, and idiopathic low-molecular-weight proteinuria. Dent's disease is characterized by low-molecular-weight proteinuria, hypercalciuria, hyperphosphaturia, nephrocalcinosis, kidney stones and eventual renal failure, with some patients developing rickets or osteomalacia. It is caused by loss of function mutation of a proximal tubular endosomal chloride channel, Cl^-C5. This chloride channel, along with the proton pump, is essential for acidification of proximal tubular endosomes. The process is linked with normal endocytosis, degradation and recycling of absorbed proteins, vitamins and hormones. Defective endosomal acidification (owing to the mutated Cl^-C5 gene) results in impaired endosomal degradation and recycling of endocytosed hormones such as PTH. High PTH concentration activates apical receptors, increases endocytosis of the sodium–phosphate cotransporter and produces increased phosphaturia. It also results in increased 1α-hydoxylase activity, causing excess production of 1,25-$(OH)_2$ D_3, hypercalciuria, renal stones and nephrocalcinosis. Moreover, the receptors (megalin and cubilin) for reabsorption of low-molecular-weight proteins and albumin in the proximal tubules are decreased in Dent's disease. This is secondary to the defective acidification of endosomes, which can explain low-molecular-weight proteinuria and excessive urinary leaks of cytokines, hormones and chemokines. This urinary profile is associated with progressive renal fibrosis and more rapid decline in renal function.

Treatment

Treatment for hypophosphataemia includes combined therapy with phosphate supplementation and calcitriol (1,25-dihydroxyvitamin D) administration.

Treatment of acute hypophosphataemia, if warranted, is with intravenous phosphate at a maximum rate of 9 mmol every 12 hours, with repeated measurements of calcium and phosphate, as over-rapid administration of phosphate may lead to severe hypocalcaemia, particularly in the presence of alkalosis. *Chronic hypophosphataemia* can be corrected, if warranted, with oral effervescent sodium phosphate.

HYPERPHOSPHATAEMIA

Hyperphosphataemia is common in patients with chronic renal failure (see p. 628 and Table 12.18). Hyperphosphataemia is usually asymptomatic but may result in precipitation of calcium phosphate, particularly in the presence of a normal or raised calcium or of alkalosis. Uraemic itching may be caused by a raised calcium phosphate product. Prolonged

Table 12.18 Causes of hyperphosphataemia

Chronic renal failure
Phosphate-containing enemas
Tumour lysis
Myeloma-abnormal phosphate-binding protein
Rhabdomyolysis

FURTHER READING

Alexander RT, Hoenderop JG, Bindels RJ. Molecular determinants of magnesium homeostasis: insights from human disease. *Journal of the American Society of Nephrology* 2008; **19**(8): 1451–1458.

Hoenderop JG, Bindels RJ. Epithelial Ca^{2+} and Mg^{2+} channels in health and disease. *Journal of the American Society of Nephrology* 2005; **16**: 15–26.

Negri AL. Hereditary hypophosphatemias: new genes in the bone–kidney axis. *Nephrology (Carlton)* 2007; **12**(4): 317–320.

Veizis IE, Cotton CU. Role of kidney chloride channels in health and disease. *Pediatric Nephrology* 2007; **22**(6): 770–777.

Yuan B, Takaiwa M, Clemens TL et al. Aberrant Phex function in osteoblasts and osteocytes alone underlies murine X-linked hypophosphatemia. *Journal of Clinical Investigation* 2008; **118**: 722–734.

hyperphosphataemia causes hyperparathyroidism, and periarticular and vascular calcification.

Usually no treatment is required for acute hyperphosphataemia, as the causes are self-limiting. Treatment of chronic hyperphosphataemia is with gut phosphate binders and dialysis (see p. 633).

ACID–BASE DISORDERS

The concentration of hydrogen ions in both extracellular and intracellular compartments is extremely tightly controlled, and very small changes lead to major cell dysfunction. The blood pH is tightly regulated and is normally maintained at between 7.38 and 7.42. Any deviation from this range indicates a change in the hydrogen ion concentration $[H^+]$ because blood pH is the negative logarithm of $[H^+]$ (Table 12.19). The $[H^+]$ at a physiological blood pH of 7.40 is 40 nmol/L. An increase in the $[H^+]$ – a fall in pH – is termed acidaemia. A decrease in $[H^+]$ – a rise in the blood pH – is termed alkalaemia. The disorders that cause these changes in the blood pH are acidosis and alkalosis, respectively.

Normal acid–base physiology

The normal adult diet contains 70–100 mmol of acid. Throughout the body, there are buffers that minimize any changes in blood pH that these ingested hydrogen ions might cause. Such buffers include intracellular proteins (e.g. haemoglobin) and tissue components (e.g. the calcium carbonate and calcium phosphate in bone) as well as the bicarbonate–carbonic acid buffer pair generated by the hydration of carbon dioxide. This buffer pair is clinically most relevant, in part because its contribution can be measured and because alterations in this buffer pair reveal changes in all other buffer systems. Bicarbonate ions $[HCO_3^-]$ and carbonic acid (H_2CO_3) exist in equilibrium; and in the presence of carbonic anhydrase, carbonic acid dissociates to carbon dioxide and water, as expressed in the following equation:

$$H^+ + HCO_3^- \rightleftharpoons H_2CO_3 \overset{\text{carbonic anhydrase}}{\rightleftharpoons} CO_2 + H_2O$$

The addition of hydrogen ions drives the reaction to the right, decreasing the plasma bicarbonate concentration $[HCO_3^-]$ and increasing the arterial carbon dioxide pressure (P_{aCO_2}). As shown in the following Henderson– Hasselbalch equation, a fall in the plasma $[HCO_3^-]$ increases $[H^+]$ and thus lowers blood pH:

$$[H^+] = 181 \times P_{aCO_2} [HCO_3^-]$$

where $[H^+]$ is expressed in nmol per litre, P_{aCO_2} in kilopascals, $[HCO_3^-]$ in mmol per litre, and 181 is the dissociation coefficient of carbonic acid. Alternatively the equation can be expressed as:

$$pH = pK + \log[HCO_3^-]/[H_2CO_3]$$

where pK = 6.1. Thus, the bicarbonate used in the buffering process must be regenerated to maintain normal acid–base balance.

Although the acidaemia stimulates an increase in ventilation, which blunts this change in pH, increased ventilation does not regenerate the bicarbonate used in the buffering process. Consequently, the kidney must excrete hydrogen ions to return the plasma $[HCO_3^-]$ to normal. Maintenance of a normal plasma $[HCO_3^-]$ under physiological conditions depends not only on daily regeneration of bicarbonate but also on reabsorption of all bicarbonate filtered across the glomerular capillaries.

Renal reabsorption of bicarbonate

The plasma $[HCO_3^-]$ is normally maintained at approximately 25 mmol/L. In individuals with a normal glomerular filtration rate (120 mL/min), about 4500 mmol of bicarbonate is filtered each day. If this filtered bicarbonate were not reabsorbed, the plasma $[HCO_3^-]$ would fall, along with blood pH. Thus, maintenance of a normal plasma $[HCO_3^-]$ requires that essentially all of the bicarbonate in the glomerular filtrate be reabsorbed (Fig. 12.11).

The proximal convoluted tubule reclaims 85–90% of filtered bicarbonate; by contrast, the distal nephron reclaims very little. This difference is caused by the greater quantity

Table 12.19 Relationship between $[H^+]$ and pH

pH	$[H^+]$ (nmol/L)
6.9	126
7.0	100
7.1	79
7.2	63
7.3	50
7.4	40
7.5	32
7.6	25

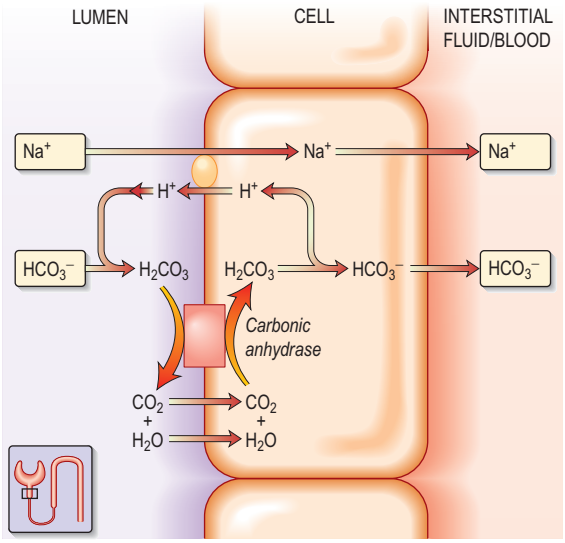


Fig. 12.11 **Resorption of sodium bicarbonate in the renal (mainly proximal) tubule.** Bicarbonate is reclaimed by the secretion of H^+ in exchange for Na^+ into the tubule. This results in the formation of H_2CO_3, which is then broken down to CO_2. This is reabsorbed and converted back to H_2CO_3, which now dissociates into H^+ and HCO_3^-. The net result is reabsorption of Na^+ and HCO_3^-. This process is dependent on carbonic anhydrase within the cells and on the luminal surface of the tubular cell.

of luminal (brush border) carbonic anhydrase in the proximal tubule than in the distal nephron. As a result of these quantitative differences, bicarbonate that escapes reabsorption in the proximal tubule is excreted in the urine.

Proximal tubular bicarbonate reabsorption is catalysed by the Na^+/K^+-ATPase pump located in the basolateral cell membrane. By exchanging peritubular potassium ions for intracellular sodium ions, the pump keeps the intracellular sodium concentration low, allowing sodium ions to enter the cell by moving down the sodium concentration gradient from the tubule lumen to the cell interior. Hydrogen ions are transported in the opposite direction (at the Na^+–H^+ antiporter), thereby maintaining electroneutrality. Before bicarbonate enters the proximal tubule, it combines with secreted hydrogen ions, forming carbonic acid. In the presence of luminal carbonic anhydrase (CA-IV) carbonic acid rapidly dissociates into carbon dioxide and water, which can then rapidly enter the proximal tubular cell. In the cell, carbon dioxide is hydrated by cytosolic carbonic anhydrase (CA-II), ultimately forming bicarbonate, which is then transported down an electrical gradient from the cell interior, across the membrane into the peritubular fluid, and into the blood. In this process, each hydrogen ion secreted into the proximal tubule lumen is reabsorbed and can be resecreted; there is no net loss of hydrogen ions or net gain of bicarbonate ions.

Renal excretion of $[H^+]$ (Fig. 12.12)

More acid is secreted into the proximal tubule (up to 4500 nmol of hydrogen ions each day) than into any other nephron segment. However, the hydrogen ions secreted into the proximal tubule are almost completely reabsorbed with bicarbonate; consequently, proximal tubular hydrogen ion secretion does not contribute significantly to hydrogen ion elimination from the body. The excretion of the daily acid load requires hydrogen ion secretion in more distal nephron segments.

Most dietary hydrogen ions come from sulphur-containing amino acids that are metabolized to sulphuric acid (H_2SO_4), which then reacts with sodium bicarbonate as follows:

$$H_2SO_4 + 2NaHCO_3 \rightarrow Na_2SO_4 + 2CO_2 + 2H_2O.$$

Excess sulphate is excreted in the urine, whereas excess hydrogen ions are buffered by bicarbonate and lower the plasma $[HCO_3^-]$. This fall in plasma $[HCO_3^-]$ leads to a slight decrease in the blood pH, although a smaller decrease in the blood pH than would have occurred if buffer were unavailable. The subsequent excretion of hydrogen ions takes place primarily in the collecting duct and results in the regeneration of 1 mmol of bicarbonate for every mmol of hydrogen ions excreted in the urine.

The *collecting duct* has three types of cells:

- The *principal cell* with an aldosterone-sensitive Na^+ absorption site. These cells reabsorb Na^+ and H_2O and secrete K^+ under the influence of aldosterone.
- The *α-intercalated cell*, which possesses the proton pump for the active secretion of hydrogen ions in exchange for reabsorption of K^+ ions. Aldosterone increases H^+ ion secretion.
- The *β-intercalated cells* are mirror images of α-intercalated cells where the H^+-ATPase pump is located in the basolateral rather than the apical membrane whereby H^+ ions are secreted into the peritubular capillary. The HCO_3^- ions, on the other hand, are secreted into the tubular lumen by an anion exchanger in the apical membrane. The identity of this transporter is uncertain, however, as it does not appear to represent the same Cl^--HCO_3^- exchanger that is present in the basolateral membrane of the H^+-secreting intercalated cells.

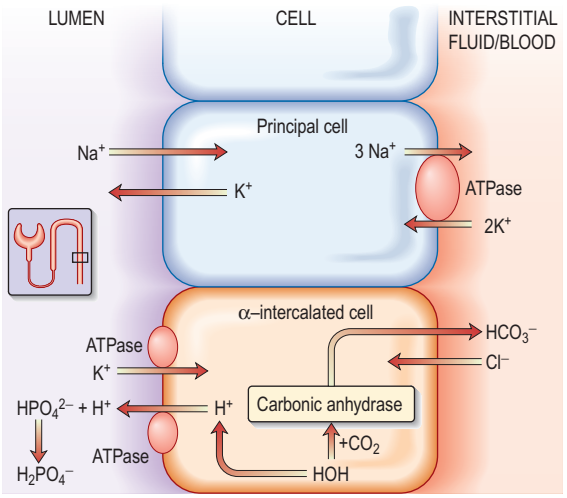


Fig. 12.12 **Renal excretion of H^+.** Secretion of H^+ from the cortical collecting ducts is indirectly linked to Na^+ reabsorption. Intracellular potassium is exchanged for sodium in the principal cell.

Aldosterone stimulates H^+ secretion by entering the principal cell, where it opens Na^+ channels in the luminal membrane and increases Na^+/K^+-ATPase activity. The movement of cationic Na^+ into the principal cells then creates a negative charge within the tubule lumen. K^+ moves from the electrochemical gradient and into the lumen.

Aldosterone apparently also stimulates the H^+-ATPase directly in the intercalated cell, further enhancing H^+ secretion. When the urinary pH falls to 4.0–4.5, further H^+ secretion by the α-intercalated cells ceases. The filtration of titratable acids (e.g. phosphoric acid, H_2PO_4) raises the intraluminal pH and permits this process to continue. Secreted H^+ binds to the conjugate anion of a titratable acid (HPO_4^{2-} in this case) and is excreted in the urine. The H^+ to be secreted arises from the reassociation of H_2O and CO_2 in the presence of carbonic anhydrase; thus, a bicarbonate molecule is regenerated each time an H^+ is eliminated in the urine.

Secretion of hydrogen ions from the cortical collecting duct is indirectly linked to sodium reabsorption. Aldosterone has several facilitating effects on hydrogen ion secretion. Aldosterone opens sodium channels in the luminal membrane of the principal cell and increases Na^+/K^+-ATPase activity. The subsequent movement of cationic sodium into the principal cell creates a negative charge within the tubule lumen. Potassium ions from the principal cells and hydrogen ions from the α-intercalated cells move out from the cells down the electrochemical gradient and into the lumen. Aldosterone also stimulates directly the H^+-ATPase in the α-intercalated cell, further enhancing hydrogen ion secretion.

When hydrogen ions are secreted into the lumen of the collecting tubule, a tiny, but physiologically critical, fraction

of these excess hydrogen ions remains in solution. Here, they increase the urinary [H^+] and lower urinary pH below 4.0. Nevertheless, below this urine pH, inhibition of proton-secreting pumps such as H^+-ATPase severely restricts kidney secretion of more hydrogen ions. Consequently, secretion of hydrogen ions depends on the presence of buffers in the urine that maintain the urine pH at a level higher than 4.0.

In the presence of alkali excess, the homeostatic needs are reversed. Although the kidney can excrete excess alkaline load by reducing reabsorption of filtered bicarbonate in the proximal and distal tubule, the collecting ducts also contribute by secreting bicarbonate brought about by switching to β-intercalated cells. This switch enables kidneys to secrete bicarbonate and conserve H^+ ions.

Buffer systems in acid excretion

Two buffer systems are involved in acid excretion: the titratable acids such as phosphate, and the ammonia system. Each system is responsible for excreting about half of the daily acid load of 50–100 mmol under physiological conditions (Fig. 12.12).

Titratable acid

A titratable acid is a filtered buffer substance having a conjugate anion that can be titrated within the pH range occurring physiologically in the urine. Phosphoric acid (pK_a 6.8) is the usual titratable urinary buffer. Hydrogen ions bind to the conjugate anions of the titratable acids and are excreted in the urine. For each hydrogen ion excreted in this form, a bicarbonate ion is regenerated within the cell and returned to the blood (Fig. 12.12).

Ammonium ($NH4^+$)

In the setting of metabolic acidosis, titratable acids cannot increase significantly because the availability of titratable acid is fixed by the plasma concentration of the buffer and by the GFR. The ammonia buffer system, by contrast, can increase several hundred fold when necessary. Consequently, impaired renal excretion of hydrogen ions is always associated with a defect in ammonium excretion (Fig. 12.13).

All ammonia used to buffer urinary hydrogen ions in the collecting tubule is synthesized in the proximal convoluted tubule. Glutamine is the primary source of ammonia. It undergoes deamination catalysed by glutaminase, resulting in α-ketoglutaric acid (Fig. 12.13) and ammonia. Once formed, ammonia can diffuse into the proximal tubule lumen and become acidified, forming ammonium. Once in the proximal tubule lumen, ammonium flows along the tubule to the thick ascending limb of Henle's loop. Here, it is transported out of the tubule into the medullary interstitium. Ammonium then dissociates to ammonia, leading to a high interstitial ammonia concentration. Ammonia diffuses down its concentration gradient into the lumen of the collecting tubule. Here, it reacts with the hydrogen ions secreted by the collecting tubular cells to form ammonium. Because ammonium (NH_4) is not lipid-soluble, it is trapped in the lumen and excreted in the urine as ammonium chloride. Two conditions predominantly promote ammonia synthesis by the proximal tubular cell: systemic acidosis and hypokalaemia.

Whether glutamine metabolism and ureagenesis in the liver play a role in acid–base homeostasis is controversial. Ureagenesis in the liver consumes up to 1000 mmol of bicarbonate/day in humans as a result of $2NH_4 + 2HCO_3$ = urea + $CO_2 + H_2O$. The liver was believed to contribute to regulation of acid–base balance by controlling the rate of ureagenesis and therefore bicarbonate consumption in response to changes in plasma acidity. Studies in human volunteers have shown that metabolic acidosis increases ureagenesis and potentially makes it worse by consuming more bicarbonate. Similarly, metabolic alkalosis decreases ureagenesis whereby less bicarbonate is consumed, resulting in an increase of alkali load in the face of alkalosis. It is therefore concluded that ureagenesis is a maladaptive process for acid–base regulation and that ureagenesis has no discernible homeostatic effect on acid–base equilibrium in humans.

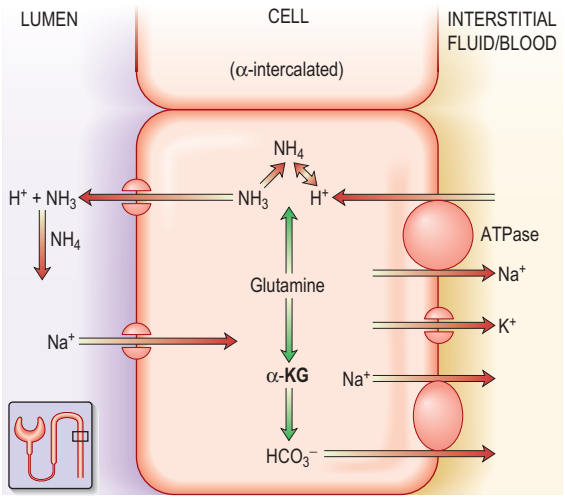


Fig. 12.13 **The ammonia buffering system in the kidney.** All ammonia used to buffer H^+ in the collecting duct is synthesized in the proximal convoluted tubule, and glutamine is the main source of this ammonia. As glutamine is metabolized, α-ketoglutarate (α-KG) is formed, which ultimately breaks down to bicarbonate that is then secreted into the peritubular fluid at an Na^+–HCO_3^- cotransporter.

Causes of acid–base disturbance

Acid–base disturbance may be caused by:

- abnormal CO_2 removal in the lungs ('respiratory' acidosis and alkalosis)
- abnormalities in the regulation of bicarbonate and other buffers in the blood ('metabolic' acidosis and alkalosis).

Both may, and usually do, coexist. For instance, metabolic acidosis causes hyperventilation (via medullary chemoreceptors, see p. 815), leading to increased removal of CO_2 in the lungs and partial compensation for the acidosis. Conversely, respiratory acidosis is accompanied by renal bicarbonate retention, which could be mistaken for primary metabolic alkalosis. The situation is even more complex if a patient has both respiratory disease and a metabolic disturbance.

Diagnosis

Clinical history and examination usually point to the correct diagnosis. Table 12.20 shows the typical blood changes, but in complicated patients the acid–base nomogram (Fig. 12.14) is invaluable. The [H^+] and P_aCO_2 are measured in arterial

Table 12.20 Changes in arterial blood gases

	pH	P_aCO_2	HCO_3^-
Respiratory acidosis	N or ↓	↑↑	↑ (compensated)
Respiratory alkalosis	N or ↑	↓↓	↓ (slight)
Metabolic acidosis	N or ↓	↓	↓↓
Metabolic alkalosis	N or ↑	↑ (slight)	↑↑

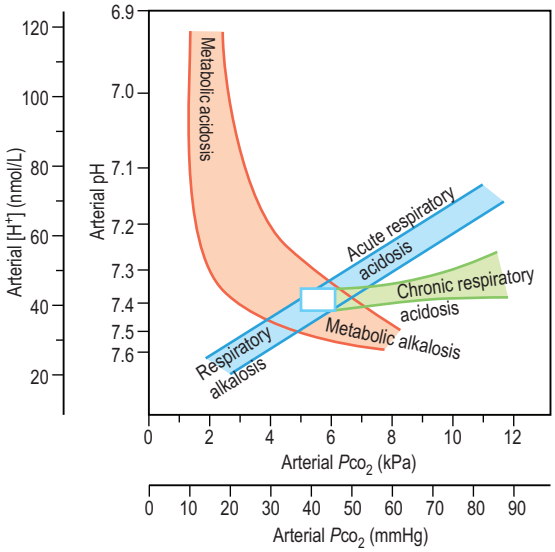


Fig. 12.14 **The Flenley acid–base nomogram.** This was derived from a large number of observations in patients with 'pure' respiratory or metabolic disturbances. The bands show the 95% confidence limits representing the individual varieties of acid–base disturbance. The central white box shows the approximate limits of arterial pH and P_{CO_2} in normal individuals.

blood (for precautions see p. 672) as well as the bicarbonate. If the values from a patient lie in one of the bands in the diagram, it is likely that only one abnormality is present. If the $[H^+]$ is high (pH low) but the P_aCO_2 is normal, the intercept lies between two bands: the patient has respiratory dysfunction, leading to failure of CO_2 elimination, but this is partly compensated for by metabolic acidosis, stimulating respiration and CO_2 removal (this is the most common 'combined' abnormality in practice).

RESPIRATORY ACIDOSIS AND ALKALOSIS

Respiratory acidosis

This is caused by retention of CO_2. The P_aCO_2 and $[H^+]$ rise. Renal retention of bicarbonate may partly compensate, returning the $[H^+]$ towards normal (see p. 893).

Respiratory alkalosis

Increased removal of CO_2 is caused by hyperventilation, so there is a fall in P_aCO_2 and $[H^+]$ (see p. 893).

METABOLIC ACIDOSIS

This is due to the accumulation of any acid other than carbonic acid, and there is a primary decrease in the plasma $[HCO_3^-]$. Several disorders can lead to metabolic acidosis: acid administration, acid generation (e.g. lactic acidosis during shock or cardiac arrest), impaired acid excretion by the kidneys, or bicarbonate losses from the gastrointestinal tract or kidneys. From a diagnostic viewpoint, calculation of the plasma anion gap is extremely useful in narrowing this differential diagnosis.

The anion gap

The first step is to identify whether the acidosis is due to retention of H^+Cl^- or to another acid. This is achieved by calculation of the anion gap. The principles underlying this calculation are straightforward:

- The normal cations present in plasma are Na^+, K^+, Ca^{2+}, Mg^{2+}.
- The normal anions present in plasma are Cl^-, HCO_3^-, negative charges present on albumin, phosphate, sulphate, lactate, and other organic acids.
- The sums of the positive and negative charges are equal.
- Measurement of plasma $[Na^+]$, $[K^+]$, $[Cl^-]$ and $[HCO_3^-]$ is usually easily available.

$$\text{ANION GAP} = \{[Na^+] + [K^+]\} - \{[HCO_3^-] + [Cl^-]\}$$

Because there are more unmeasured anions than cations, the normal anion gap is 10–18 mmol/L, although calculations with more sensitive methods place this at 6–12 mmol/L. Albumin normally makes up the largest portion of these unmeasured anions. As a result, a fall in the plasma albumin concentration from the normal value of about 40 g/L to 20 g/L may reduce the anion gap by as much as 6 mmol/L, because each 1 g/L of albumin has a negative charge of 0.2–0.28 mmol/L.

Metabolic acidosis with a normal anion gap

If the anion gap is normal in the presence of acidosis, this suggests that H^+Cl^- is being retained or that $Na^+HCO_3^-$ is being lost. Causes of a normal-anion-gap acidosis are given in Table 12.21. In these conditions, plasma bicarbonate decreases and is replaced by chloride to maintain electroneutrality. Consequently, these disorders are sometimes referred to collectively as hyperchloraemic acidoses.

Table 12.21 Causes of metabolic acidosis with a normal anion gap

Increased gastrointestinal bicarbonate loss
Diarrhoea
Ileostomy
Ureterosigmoidostomy

Increased renal bicarbonate loss
Acetazolamide therapy
Proximal (type 2) renal tubular acidosis
Hyperparathyroidism
Tubular damage, e.g. drugs, heavy metals, paraproteins

Decreased renal hydrogen ion excretion
Distal (type 1) renal tubular acidosis
Type 4 renal tubular acidosis (aldosterone deficiency)

Increased HCl production
Ammonium chloride ingestion
Increased catabolism of lysine, arginine

Ask the authors
www.studentconsult.com

Renal tubular acidosis (RTA)

This term refers to systemic acidosis caused by impairment of the ability of the renal tubules to maintain acid–base balance. This group of disorders is uncommon and only rarely a cause of significant clinical disease.

Type 4 renal tubular acidosis

Also called 'hyporeninaemic hypoaldosteronism', this is probably the most common of these disorders. The cardinal features are hyperkalaemia and acidosis occurring in a patient with mild chronic renal insufficiency, usually caused by tubulo-interstitial disease (e.g. reflux nephropathy) or diabetes (Gordon's syndrome (see p. 667) shares biochemical abnormalities but differs in having normal GFR and hypertension). Plasma renin and aldosterone are found to be low, even after measures which would normally stimulate their secretion. The features for the diagnosis are shown in Table 12.22. An identical syndrome is caused by chronic ingestion of NSAIDs, which impair renin and aldosterone secretion. In the presence of acidosis, urine pH may be low. Treatment is with fludrocortisone, sodium bicarbonate, diuretics, or ion exchange resins to remove potassium, or a combination of these. Dietary potassium restriction alone is ineffective.

Type 3 renal tubular acidosis (RTA)

This condition is vanishingly rare, and represents a combination of type 1 and type 2. Inherited type 3 RTA is caused by mutations resulting in carbonic anhydrase type II deficiency, which is characterized by osteopetrosis, RTA of mixed type, cerebral calcification, and mental retardation. Typical radiographic features of osteopetrosis are present and histopathological study of the iliac crest reveals unreabsorbed calcified primary spongia.

Type 2 ('proximal') renal tubular acidosis

This is very rare in adult practice. It is caused by failure of sodium bicarbonate reabsorption in the proximal tubule. The cardinal features are acidosis, hypokalaemia, an inability to lower the urine pH below 5.5 despite systemic acidosis, and the appearance of bicarbonate in the urine despite a subnormal plasma bicarbonate. This disorder normally occurs as part of a generalized tubular defect, together with other features such as glycosuria and amino-aciduria. Inherited forms of isolated type 2 RTA are described as both autosomal dominant and recessive patterns of inheritance, where putative mutations are in the Na^+–H^+ antiporter in the apical membrane and Na^+–HCO_3^- cotransporter in the basolateral membrane of proximal tubular cells respectively (see Fig. 12.11). Treatment is with sodium bicarbonate: massive doses may be required to overcome the renal 'leak'.

Type 1 ('distal') renal tubular acidosis

This is due to a failure of H^+ excretion in the distal tubule (Table 12.23). It consists of:

- acidosis
- hypokalaemia (few exceptions)
- inability to lower the urine pH below 5.3 despite systemic acidosis
- low urinary ammonium production.

These features may be present only in the face of increased acid production; hence the need for an acid load test in diagnosis (Practical box 12.1). Other features include:

- low urinary citrate (owing to increased citrate absorption in the proximal tubule where it can be converted to bicarbonate)
- hypercalciuria.

These abnormalities result in osteomalacia, renal stone formation and recurrent urinary infections. *Osteomalacia* is

Table 12.22 Features of hyporeninaemic hypoaldosteronism (type 4 renal tubular acidosis)

- **Hyperkalaemia**
  (In the absence of drugs known to cause hyperkalaemia)
- **Low plasma bicarbonate and hyperchloraemia**
- **Normal ACTH stimulation test (p. 1010)**
- **Low basal 24-hour urinary aldosterone**
- **Subnormal response of plasma renin and plasma aldosterone to stimulation**
  Samples taken over 2 hours supine and again after 40 mg furosemide (80 mg if creatinine > 120 µmol/L) and 4 hours upright posture
- **Correction of hyperkalaemia by fludrocortisone 0.1 mg daily**

Table 12.23 Causes of distal renal tubular acidosis (type 1 RTA)

Primary
Idiopathic
Genetic
Marfan's syndrome
Ehlers–Danlos syndrome
Sickle cell anaemia

Nephrocalcinosis
Chronic hypercalcaemia
Medullary sponge kidney

Hypergammaglobulinaemic states
Amyloidosis*
Cryoglobulinaemia
Chronic liver disease

Drugs and toxins
Amphotericin B
Lithium carbonate
NSAIDs
Lead

Autoimmune diseases
Sjögren's syndrome*
Thyroiditis
Autoimmune hepatitis
Primary biliary cirrhosis
Systemic lupus erythematosus

Renal transplant rejection*

*May also cause proximal renal tubular acidosis

Practical Box 12.1 Diagnosis of renal tubular acidosis

Plasma HCO_3^- < 21 mmol/L, urine pH > 5.3 = renal tubular acidosis
To differentiate between proximal (very rare) and distal (rare) requires bicarbonate infusion test
Plasma HCO_3^- > 21 mmol/L but suspicion of partial renal tubular acidosis (e.g. nephrocalcinosis-associated diseases): **acid load test** required as follows:

- Give 100 mg/kg ammonium chloride by mouth
- Check urine pH hourly and plasma HCO_3^- at 3 hours
- Plasma HCO_3^- should drop below 21 mmol/L unless the patient vomits (in which case the test should be repeated with an antiemetic)
- If urine pH remains > 5.3 despite a plasma HCO_3^- of 21 mmol/L, the diagnosis is confirmed

caused by buffering of H^+ by Ca_2^+ in bone, resulting in depletion of calcium from bone. *Renal stone formation* is caused by hypercalciuria, hypocitraturia (citrate inhibits calcium phosphate precipitation), and alkaline urine (which favours precipitation of calcium phosphate). *Recurrent urinary infections* are caused by renal stones.

Both autosomal dominant and recessive inheritance patterns have been reported in primary distal RTA. Patients in South East Asia with an autosomal dominant type have a milder phenotype and may have an associated ovalocytosis. In this form the Cl^-–HCO_3^- exchanger in the basolateral surface of intercalated cells is mutated. However, in the autosomal recessive distal RTA, a substantial proportion of patients have sensorineural deafness, and this is associated with a loss of function mutation in the H^+-ATPase at the apical surface of intercalated cells.

Treatment is with sodium bicarbonate, potassium supplements and citrate. Thiazide diuretics are useful by causing volume contraction and increased proximal sodium bicarbonate reabsorption.

Urinary anion gap

Another useful tool in the evaluation of metabolic acidosis with a normal anion gap is the urinary anion gap:

$$\text{URINARY ANION GAP} = \{\text{urinary } [Na^+] + \text{urinary } [K^+]\} - \text{urinary } [Cl^-].$$

This calculation can be used to distinguish the normal-anion-gap acidosis caused by diarrhoea (or other gastrointestinal alkali loss) from that caused by distal renal tubular acidosis. In both disorders, the plasma $[K^+]$ is characteristically low. In patients with renal tubular acidosis, urinary pH is always greater than 5.3.

Although excretion of urinary hydrogen ions in the patient with diarrhoea should acidify the urine, hypokalaemia leads to enhanced ammonia synthesis by the proximal tubular cells. Despite acidaemia, the excess urinary buffer increases the urine pH to a value above 5.3 in some patients with diarrhoea.

Whenever urinary acid is excreted as ammonium chloride, the increase in urinary chloride excretion decreases the urinary anion gap. Thus, the urinary anion gap should be negative in the patient with diarrhoea, regardless of the urine pH. On the other hand, although hypokalaemia may result in enhanced proximal tubular ammonia synthesis in distal renal tubular acidosis, the inability to secrete hydrogen ions into the collecting duct in this condition limits ammonium chloride formation and excretion; thus, the urinary anion gap is positive in distal renal tubular acidosis.

Metabolic acidosis with a high anion gap

If the anion gap is increased, there is an unmeasured anion present in increased quantities. This is either one of the acids normally present in small, but unmeasured quantities, such as lactate, or an exogenous acid. Causes of a high-anion-gap acidosis are given in Table 12.24.

Lactic acidosis

Increased lactic acid production occurs when cellular respiration is abnormal, because of either a lack of oxygen in the tissues ('type A') or a metabolic abnormality, such as drug-induced ('type B') (Table 12.24). The most common cause in clinical practice is type A lactic acidosis, occurring in septic or cardiogenic shock. Significant acidosis can occur despite a normal blood pressure and P_aCO_2, owing to splanchnic and peripheral vasoconstriction. Acidosis worsens cardiac function and vasoconstriction further, contributing to a downward spiral and fulminant production of lactic acid.

Table 12.24 Causes of metabolic acidosis with an increased anion gap

Renal failure (sulphate, phosphate)

Accumulation of organic acids

Lactic acidosis
L-lactic
- Type A – anaerobic metabolism in tissues
  - Hypotension/cardiac arrest
  - Sepsis
- Poisoning – e.g. ethylene glycol, methanol
- Type B – decreased hepatic lactate metabolism
  - Insulin deficiency (decreased pyruvate dehydrogenase activity)
  - Metformin accumulation (chronic kidney disease)
  - Haematological malignancies
  - Rare inherited enzyme defects

D-lactic (fermentation of glucose in bowel by abnormal bowel flora, complicating abnormal small bowel anatomy, e.g. blind loops)

Ketoacidosis
Insulin deficiency
Alcohol excess
Starvation

Exogenous acids
Salicylate

Diabetic ketoacidosis (see p. 1046)

There is a high-anion-gap acidosis due to the accumulation of acetoacetic and hydroxybutyric acids, owing to increased production and some reduced peripheral utilization.

Uraemic acidosis

Kidney disease causes acidosis in several ways. Reduction in the number of functioning nephrons decreases the capacity to excrete ammonia and H^+ in the urine. In addition, tubular disease may cause bicarbonate wasting. Acidosis is a particular feature of those types of chronic renal failure in which the tubules are particularly affected, such as reflux nephropathy and chronic obstructive uropathy.

Chronic acidosis is most often caused by chronic kidney disease, where there is a failure to excrete fixed acid. Up to 40 mmol of hydrogen ions may accumulate daily. These are buffered by bone, in exchange for calcium. Chronic acidosis is therefore a major risk factor for renal osteodystrophy and hypercalciuria.

Chronic acidosis has also been shown to be a risk factor for muscle wasting in renal failure, and may also contribute to the inexorable progression of some types of renal disease.

Uraemic acidosis should be corrected because of these effects on growth, muscle turnover and bones. Oral sodium bicarbonate 2–3 mmol/kg daily is usually enough to maintain serum bicarbonate above 20 mmol/L, but may contribute to sodium overload. Calcium carbonate improves acidosis and also acts as a phosphate binder and calcium supplement, and is commonly used. Acidosis in end-stage renal failure is usually fully corrected by adequate dialysis.

Mixed metabolic acidosis

Both types of acidosis may coexist. For instance, cholera would be expected to cause a normal-anion-gap acidosis owing to massive gastrointestinal losses of bicarbonate, but the anion gap is often increased owing to renal failure and lactic acidosis as a result of hypovolaemia.

Clinical features

Clinically the most obvious effect is stimulation of respiration, leading to the clinical sign of 'air hunger', or Kussmaul respiration. Interestingly, patients with profound hyperventilation may not complain of breathlessness, although in others it may be a presenting complaint.

Acidosis increases delivery of oxygen to the tissues by shifting the oxyhaemoglobin dissociation curve to the right, but it also leads to inhibition of 2,3-BPG production, which returns the curve towards normal (see p. 892). Cardiovascular dysfunction is common in acidotic patients, although it is often difficult to dissociate the numerous possible causes of this. Acidosis is negatively inotropic. Severe acidosis also causes venoconstriction, resulting in redistribution of blood from the peripheries to the central circulation, and increased systemic venous pressure, which may worsen pulmonary oedema caused by myocardial depression. Arteriolar vasodilatation also occurs, further contributing to hypotension.

Cerebral dysfunction is variable. Severe acidosis is often associated with confusion and fits, but numerous other possible causes are usually present.

As mentioned earlier, acidosis stimulates potassium loss from cells, which may lead to potassium deficiency if renal function is normal, or to hyperkalaemia if renal potassium excretion is impaired.

General treatment of acidosis

Treatment should be aimed at correcting the primary cause. In lactic acidosis caused by poor tissue perfusion ('type A'), treatment should be aimed at maximizing oxygen delivery to the tissues by protecting the airway, improving breathing and circulation. This usually requires inotropic agents, mechanical ventilation and invasive monitoring. In 'type B' lactic acidosis, treatment is that of the underlying disorder; e.g.:

- insulin in diabetic ketoacidosis
- treatment of methanol and ethylene glycol poisoning with ethanol
- removal of salicylate by dialysis.

The question of whether severe acidosis should be treated with bicarbonate is extremely controversial. Severe acidosis ($[H^+] > 100$ nmol/L, pH < 7.0) is associated with a very high mortality, which makes many doctors keen to correct it. Since acidosis is known to impair cardiac contractility, it would seem sensible to correct acidosis with bicarbonate in a sick patient. However:

- Rapid correction of acidosis may result in tetany and fits owing to a rapid decrease in ionized calcium.
- Administration of sodium bicarbonate (8.4%) provides 1 mmol/mL of sodium, which may lead to extracellular volume expansion, exacerbating pulmonary oedema.
- Bicarbonate therapy increases CO_2 production and will therefore correct acidosis only if ventilation can be increased to remove the added CO_2 load.
- The increased amounts of CO_2 generated may diffuse more readily into cells than bicarbonate, worsening intracellular acidosis.

Administration of sodium bicarbonate (50 mmol, as 50 mL of 8.4% sodium bicarbonate intravenously) is still occasionally given during cardiac arrest and is often necessary before arrhythmias can be corrected. Correction of hyperkalaemia associated with acidosis is also of undoubted benefit. In other situations there is no clinical evidence to show that correction of acidosis improves outcome, but it is standard practice to administer sodium bicarbonate when $[H^+]$ is above 126 nmol/L (pH < 6.9), using intravenous 1.26% (150 mmol/L) bicarbonate infused over 2–3 hours with electrolyte and pH monitoring. Intravenous sodium lactate should never be given.

METABOLIC ALKALOSIS

Metabolic alkalosis is common, comprising half of all the acid–base disorders in hospitalized patients. This observation should not be surprising since vomiting, the use of diuretics, and nasogastric suction are common among hospitalized patients. The mortality associated with metabolic alkalosis is substantial; the mortality rate is 45% in patients with an arterial pH of 7.55 and 80% when the pH is greater than 7.65. Although this relationship is not necessarily causal, severe alkalosis should be viewed with concern.

Classification and definitions

Metabolic alkalosis has been classified on the basis of underlying pathophysiology (Table 12.25).

The most common group is due to chloride depletion which can be corrected without potassium repletion. The other major grouping is that due to potassium depletion,

Table 12.25 Causes of metabolic alkalosis

Chloride depletion
Gastric losses: vomiting, mechanical drainage, bulimia
Chloruretic diuretics: e.g. bumetanide, furosemide, chlorothiazide, metolazone
Diarrhoeal states: villous adenoma, congenital chloridorrhoea
Cystic fibrosis (high sweat chloride)

Potassium depletion/mineralocorticoid excess
Primary aldosteronism
Secondary aldosteronism
Apparent mineralocorticoid excess
Primary deoxycorticosterone excess: 11α- and 17α-hydroxylase deficiencies
Drugs: liquorice (glycyrrhizic acid) as a confection or flavouring, carbenoxolone
Liddle's syndrome
Bartter's and Gitelman's syndromes and their variants
Laxative abuse, clay ingestion

Hypercalcaemic states
Hypercalcaemia of malignancy
Acute or chronic milk–alkali syndrome

Others
Ampicillin, penicillin therapy
Bicarbonate ingestion: massive or with renal insufficiency
Recovery from starvation
Hypoalbuminaemia

usually with mineralocorticoid excess. Metabolic alkalosis due to both potassium and chloride depletion also occurs.

Chloride may be lost from the gut, kidney or skin. The loss of gastric fluid rich in acid results in alkalosis because bicarbonate generated during the production of gastric acid returns to the circulation. In Zollinger–Ellison syndrome (see p. 385) or gastric outflow obstruction these losses can be massive. Although sodium and potassium loss in the gastric juice is variable, the obligate urinary loss of these cations is intensified by bicarbonaturia, which occurs during disequilibrium.

Chloruretic agents all directly produce loss of chloride, sodium and fluid in the urine. These losses in turn promote metabolic alkalosis by several mechanisms:

- diuretic-induced increases in sodium delivery to the distal nephron enhance potassium and hydrogen ion secretion
- extracellular volume contraction stimulates renin and aldosterone secretion, which blunts sodium losses but accelerates potassium and hydrogen ion secretion
- potassium depletion augments bicarbonate reabsorption in the proximal tubule and
- stimulates ammonia production which in turn will increase urinary net acid excretion.

Urinary losses of chloride exceed those for sodium and are associated with alkalosis even when potassium depletion is prevented. The cessation of events that generate alkalosis is not necessarily accompanied by resolution of the alkalosis. A widely accepted hypothesis for the maintenance of alkalosis is chloride depletion rather than volume depletion. Although normal functioning of the proximal tubule is essential for bicarbonate absorption, the collecting duct appears to be the major nephron site for altered electrolyte and proton transport in both maintenance and recovery from metabolic alkalosis. During maintenance, the α-intercalated cells in the cortical collecting duct do not secrete bicarbonate because insufficient chloride is available for bicarbonate exchange. When chloride is administered and luminal or cellular chloride concentration increases, bicarbonate is promptly excreted and alkalosis is corrected.

Metabolic alkalosis in *hypokalaemia* is generated primarily by an increased intracellular shift of hydrogen ion causing intracellular acidosis. Potassium depletion is also associated with enhanced ammonia production with increased obligate net acid excretion. However, the role of intracellular acidosis is supported by the correction of the alkalosis by infusion of potassium without any suppression of renal net excretion. The correction is assumed to occur by the movement of potassium into and hydrogen ion out of the cell, which titrates extracellular fluid bicarbonate.

Milk–alkali syndrome in which both bicarbonate and calcium are ingested produces alkalosis by vomiting, calcium-induced bicarbonate absorption and reduced GFR. Cationic antibiotics in high doses can cause alkalosis by obligatory bicarbonate loss in the urine.

Clinical features

The symptoms of metabolic alkalosis per se are difficult to separate from those of chloride, volume or potassium depletion. Tetany (see p. 1022), apathy, confusion, drowsiness, cardiac arrhythmias and neuromuscular irritability are common when alkalosis is severe. The oxyhaemoglobin dissociation curve is shifted to the left. Respiration may be depressed.

Treatment

Chloride-responsive metabolic alkalosis

Although replacement of the chloride deficit is essential in chloride depletion states, selection of the accompanying cation – sodium, potassium or proton – is dependent on the assessment of extracellular fluid volume status (see p. 659), the presence or absence of associated potassium depletion, and the degree and reversibility of any depression of GFR. If kidney function is normal, bicarbonate and base equivalents will be excreted with sodium or potassium, and metabolic alkalosis will be rapidly corrected as chloride is made available.

If chloride and extracellular depletion coexist then isotonic saline solution is appropriate therapy.

In the clinical settings of fluid overload, saline is contraindicated. In such situations, intravenous use of hydrochloride acid or ammonium chloride can be given. If GFR is adequate, acetazolamide, which causes bicarbonate diuresis by inhibiting carbonic anhydrase, can also be used. When the kidney is incapable of responding to chloride repletion, dialysis is necessary.

Chloride-resistant metabolic alkalosis

Metabolic alkalosis due to potassium depletion is managed by the correction of the underlying cause (see hypokalaemia). Mild to moderate alkalosis requires oral potassium chloride administration. However, the presence of cardiac arrhythmia or generalized weakness requires intravenous potassium chloride.

QUESTIONS AND CASE STUDIES

Case 12.1

A 44-year-old patient with known chronic liver disease secondary to hepatitis B presented with marked exertional dyspnoea. Examination revealed: blood pressure 130/80, elevated JVP at 8 cm, bilateral pleural effusions, ascites and marked pitting pedal oedema. Investigations revealed: plasma sodium 136 mmol/L, potassium 3.6 mmol/L, bicarbonate 32 mmol/L, chloride 96 mmol/L, urinary sodium <10 mmol/day, urinary potassium 60 mmol/day.

Question

(a) Explain the persistent retention of sodium in the face of salt and water overload?

Answer

(a) In cirrhosis, arterial vasodilatation due to nitric oxide overactivity leads to arterial underfilling. This is perceived by the pressure and volume receptors as hypovolaemia, with consequent activation of the sympathetic system, non-osmotic release of ADH and activation of the renin–angiotensin–aldosterone system. These mediators lead to salt and water reduction (see Fig. 12.3a).

FURTHER READING

Ahya SN, Jose Soler M, Levitsky J. Acid–base and potassium disorders in liver disease. *Seminars in Nephrology* 2006; **26**: 466–470.

Fry AC, Karet FE. Inherited renal acidoses. *Physiology (Bethesda)* 2007; **22**: 202–211.

Kellum JA. Acid–base disorders and strong ion gap. *Contributions to Nephrology* 2007; **156**: 158–166.

Kraut JA, Madias NE. Serum anion gap: its uses and limitations in clinical medicine. *Clinical Journal of the American Society of Nephrology* 2007; **2**(1):162–174.

Miltiadous G, Christidis D, Kalogirou M et al. Causes and mechanisms of acid-base and electrolyte abnormalities in cancer patients. *European Journal of Internal Medicine* 2008; **19**(1): 1–7.

Case 12.2

A 32-year-old patient was referred for investigation of refractory hypertension. His blood pressure was elevated at 210/110 with grade 3 hypertensive retinopathy despite taking beta-blockers, a calcium antagonist and an ACE inhibitor. The rest of the examination was normal. Investigations revealed: sodium 148 mmol/L, potassium 3 mmol/L, bicarbonate 32 mmol/L, urea 4 mmol/L, glucose 4 mmol/L, urine dipstick negative for proteins.

Questions

(a) What is the likely diagnosis?
(b) What is the cause of absence of oedema in this patient?

Answers

(a) Conn's syndrome (see p. 1023).
(b) This is due to the escape from the action of aldosterone. The escape is dependent on an increase in delivery of sodium to the site of aldosterone in the collecting duct. The increase in sodium delivery is achieved by high extracellular volume-mediated arterial overfilling. This suppresses sympathetic activity and angiotensin II generation, and increases cardiac release of ANP with resultant increase in renal perfusion pressure and GFR. The net result of these events is reduced sodium absorption in the proximal tubules and increased sodium delivery, which overwhelms the sodium-retaining actions of aldosterone (see Fig. 12.3b).

Case 12.3

A 62-year-old woman is found lying on the floor of her house. On examination she is bradycardic and peripherally shut down. ECG confirms sinus bradycardia with changes consistent with inferior myocardial infarction. Emergency electrolyte measurements revealed: sodium 136 mmol/L, potassium 4 mmol/L, bicarbonate 10 mmol/L, chloride 102 mmol/L, urea 7.5 mmol/L.

Question

(a) Comment on the biochemistry and how you will confirm your diagnosis?

Answer

(a) The biochemical picture is one of high-anionic-gap metabolic acidosis. In the presence of myocardial infarction and hypotension, the likely cause would be lactic acidosis. Measurement of plasma lactate (normally <2 mmol/L) would confirm the diagnosis. Correction of the underlying cause will correct the acidosis by itself.

Case 12.4

The plasma biochemistry of a patient who presented with severe loin pain is as follows: sodium 138 mmol/L, potassium 2.5 mmol/L, urea 3.8 mmol/L, chloride 114 mmol/L, bicarbonate 14.5 mmol/L, urinary pH 6.5.

Questions

(a) What is the cause of this metabolic abnormality?
(b) What is the cause of the loin pain?

Answers

(a) Distal renal tubular acidosis (see the section on 'Renal tubular acidosis').
(b) Nephrocalcinosis and renal colic due to renal stones are characteristic features of renal tubular acidosis.

Case 12.5

An 18-year-old boy was referred for the investigation of chronic fatigue syndrome. His mother comments that he is easily tired by sport and is not doing well at school. He is normotensive. Investigations revealed: sodium 145 mmol/L, potassium 2.8 mmol/L, bicarbonate 35 mmol/L, chloride 80 mmol/L, magnesium 0.6 mmol/L (normal range > 0.8 mmol/L), urea 5 mmol/L, glucose 5.2 mmol/L, urinary sodium excretion of 60 mmol and potassium excretion of 60 mmol/day.

Question

(a) What is the likely diagnosis?

Answer

(a) This patient has classic Gitelman's syndrome. It is similar to Bartter's but hypomagnesaemia favours Gitelman's syndrome.

Case 12.6

A 36-year-old woman with a past medical history of peptic ulceration presented with a history of 3 days of vomiting. She looked unwell and investigations revealed: haemoglobin 16.3 mmol/L, sodium 138 mmol/L, potassium 2.8 mmol/L, urea 14.3 mmol/L, pH 7.52, bicarbonate 36 mmol/L, chloride 75 mmol/L.

Questions

(a) What is the likely metabolic abnormality?
(b) What is the likely diagnosis?
(c) What is the pH of the patient's urine?

Answers

(a) This patient suffers from hypochloraemic metabolic alkalosis.
(b) The underlying cause is gastric outflow obstruction.
(c) The patient's urine is paradoxically acid, despite alkalosis.

Case 12.7

A 72-year-old man was admitted with a 3-week history of feeling unwell, poor appetite and diarrhoea. On examination, he was clearly dehydrated and had acidotic breathing. His blood pressure was 90/70; the postural drop to 60/30. He was frail. The rest of the general and system examination was unremarkable. The initial biochemistry revealed: sodium 121 mmol/L, potassium 1.9 mmol/L, chloride 83 mmol/L, bicarbonate 4 mmol/L, urea 90 mmol/L, creatinine 960 µmol/L, pH 7.1, P_aO_2 15.7 kPa, P_aCO_2 2.2 kPa, serum osmolality 334 mOsm/kg, urine osmolality 404 mOsm/kg, urinary sodium 2 mmol/L and potassium 28 mmol/L. He continued to pass between 50 and 80 mL of urine per hour.

Questions

(a) How would you describe the metabolic abnormality?
(b) What was the likely cause?
(c) Why did this patient remain polyuric in the face of severe dehydration?

Answers

(a) This patient suffers from volume depletion syndrome characterized by dehydration, hypokalaemia, and hypochloraemic metabolic acidosis with high ionic gap. He has severe acute prerenal failure with normal urine output and dilute urine.
(b) The only single diagnosis which can explain this abnormality is a colorectal secretory villous adenoma.
(c) His polyuria in the face of severe dehydration is due to nephrogenic diabetes insipidus secondary to chronic hypokalaemia. Hypokalaemia downregulates aquaporin II, which is essential for ADH-dependent water absorption from the collecting duct.

13 Cardiovascular disease

Essential anatomy, physiology and embryology of the heart 681
Symptoms of heart disease 688
Examination of the cardiovascular system 690
General examination 690
The arterial pulse 690
The blood pressure 691
Jugular venous pressure 691
Examination of the precordium 692
Cardiac investigations 692
Therapeutic procedures 709
Cardiac arrhythmias 715
Sinus node function 715
Bradycardias and heart block 717
Supraventricular tachycardias 720
Ventricular tachyarrythmias 726
Long-term management of cardiac tachyarrhythmias 729
Heart failure 733
Pathophysiology 733
Clinical syndromes of heart failure 736
Investigations in heart failure 737
Treatment of heart failure 737
Acute heart failure 742
Ischaemic heart disease 743
Angina 748
Acute coronary syndromes 752
ST elevation myocardial infarction (STEMI) 755
Sudden cardiac death 759
Valvular heart disease 760
Mitral stenosis 760
Mitral regurgitation 763
Aortic stenosis 765
Aortic regurgitation 766
Tricuspid stenosis 767
Tricuspid regurgitation 768
Pulmonary stenosis 768
Pulmonary regurgitation 769
Prosthetic valves 769
Infective endocarditis 769
Congenital heart disease 773
Marfan's syndrome 780
Pulmonary heart disease 781
Myocardial and endocardial disease 787
Atrial myxoma 787
Myocardial disease 787
Cardiomyopathy 788
Pericardial disease 794
Heart disease in the elderly 797
Systemic hypertension 798
Peripheral vascular disease 805
Peripheral arterial disease 805
Peripheral venous disease 809

ESSENTIAL ANATOMY, PHYSIOLOGY AND EMBRYOLOGY OF THE HEART

Introduction

Myocardial cells constitute 75% of the heart mass but only about 25% of the cell number. They are designed to perform two fundamental functions: initiation and conduction of electrical impulses and contraction. Although most myocardial cells are able to perform both these functions, the vast majority are predominantly contractile cells (myocytes) and a small number are specifically designed as electrical cells. The latter, collectively known as the conducting system of the heart, are not nervous tissue but modified myocytes lacking in myofibril components. They have the ability to generate electrical impulses which are then conducted to the myocytes, leading to contraction by a process known as excitation–contraction coupling. The rate of electrical impulse generation and the force of myocardial contraction are modified by numerous factors including autonomic input and stretch.

Three epicardial coronary arteries supply blood to the myocardium, and a more complex network of veins is responsible for drainage. In the face of continuous arterial pressure fluctuations, blood vessels, especially in the cerebral circulation, maintain constant tissue perfusion by a process known as '*autoregulation*'; blood vessel control is, however, complex involving additional local and central mechanisms.

The conduction system of the heart

The sinus node

The sinus node is a complex spindle-shaped structure that lies in the lateral and epicardial aspects of the junction between the superior vena cava and the right atrium (Fig. 13.1). Physiologically, it generates impulses automatically by spontaneous depolarization of its membrane at a rate quicker than any other cardiac cell type. It is therefore the natural pacemaker of the heart.

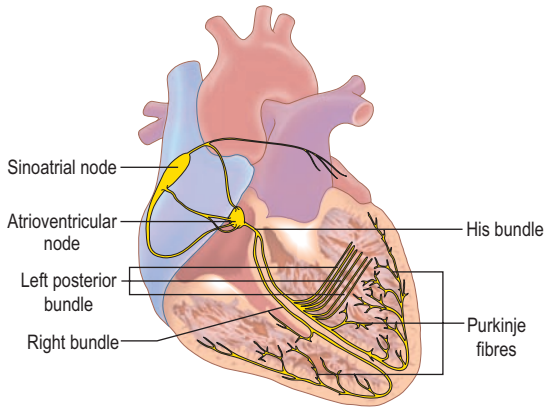


Fig. 13.1 **The conducting system of the heart.**

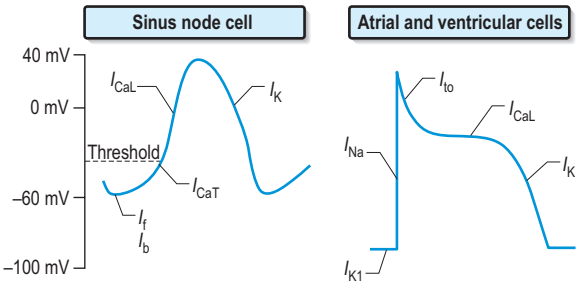


Fig. 13.2 **Myocardial action potentials.** I_b, background inward sodium current; I_f, 'funny' current; I_{CaT}, transient (or 'T' type) and I_{CaL}, long-lasting (or 'L' type) calcium channels; I_{Na}, inward sodium current; I_{to}, transient outward potassium current; I_{K1}, inwards rectifier current; I_K delayed rectifier potassium current.

A number of factors are responsible for the spontaneous decay of the sinus node cell membrane potential ('the pacemaker potential'), the most significant of which is a small influx of sodium ions into the cells. This small sodium current has two components: the background inward current (I_b) and the 'funny' (I_f) current (or pacemaker current) (Fig. 13.2). The term 'funny' current denotes ionic flow through channels activated in *hyperpolarized* cells (–60 mV or greater), unlike other time- and voltage-dependent channels activated by *depolarization*. The rate of depolarization of the sinus node membrane potential is modulated by autonomic tone (i.e. sympathetic and parasympathetic input), stretch, temperature, hypoxia, blood pH and in response to other hormonal influences (e.g. tri-iodothyronine and serotonin).

Atrial and ventricular myocyte action potentials

Action potentials in the sinus node trigger depolarization of the atrial and subsequently the ventricular myocytes. These cells have a different action potential from that of sinus node cells (see Fig. 13.2). Their resting membrane potential is a consequence of a small flow of potassium ions into the cells through open 'inward rectifier' channels; at this stage sodium and calcium channels are closed. The arrival of adjacent action potentials triggers the opening of voltage-gated, fast, self-inactivating sodium channels, resulting in a sharp depolarization spike. This is followed by a partial repolarization of the membrane due to activation of 'transient outward' potassium channels.

The plateau phase which follows is unique to myocytes and results from a small, but sustained inward calcium current through L-type calcium channels lasting 200–400 ms. This calcium influx is caused by a combined increase in permeability of the cell, especially the sarcolemmal membranes to calcium (Fig. 13.3). This plateau (or refractory) phase in myocyte action potential prevents early reactivation of the myocytes and directly determines the strength of contraction. The gradual inactivation of the calcium channels activates delayed rectifier potassium channels repolarizing the membrane. Atrial tissue is activated like a 'forest fire', but the activation peters out when the insulating layer between the atrium and the ventricle – the annulus fibrosus – is reached. Controversy still exists about whether impulses from the sinoatrial (SA) node travel over specialized conducting 'pathways' or over ordinary atrial myocardium.

The atrioventricular node, His bundle and Purkinje fibres

The depolarization continues to conduct slowly through the atrioventricular (AV) node. This is a small, bean-shaped structure that lies beneath the right atrial endocardium within the lower interatrial septum. The AV node continues as the His bundle, which penetrates the annulus fibrosus and conducts the cardiac impulse rapidly towards the ventricle. The His bundle reaches the crest of the interventricular septum and divides into the right bundle branch and the main left bundle branch.

The *right bundle branch* continues down the right side of the interventricular septum to the apex, from where it radiates and divides to form the Purkinje network, which spreads throughout the subendocardial surface of the right ventricle. The *main left bundle branch* is a short structure, which fans out into many strands on the left side of the interventricular septum. These strands can be grouped into an anterior superior division (the anterior hemi-bundle) and a posterior inferior division (the posterior hemi-bundle). The anterior hemi-bundle supplies the subendocardial Purkinje network of the anterior and superior surfaces of the left ventricle, and the inferior hemi-bundle supplies the inferior and posterior surfaces. Impulse conduction through the AV node is slow and depends on action potentials largely produced by slow transmembrane calcium flux. In the atria, ventricles and His–Purkinje system, conduction is rapid and is due to action potentials generated by rapid transmembrane sodium diffusion.

The cellular basis of myocardial contraction – excitation–contraction coupling

Each myocyte, approximately 100 μm long, branches and interdigitates with adjacent cells. An intercalated disc permits electrical conduction to adjacent cells. Myocytes contain bundles of parallel myofibrils. Each myofibril is made up of a series of sarcomeres (Fig. 13.4). A sarcomere (which is the basic unit of contraction) is bound by two transverse Z lines, to each of which is attached a perpendicular filament of the protein actin. The actin filaments from each of the two Z bands overlap with thicker parallel protein filaments known as myosin. Actin and myosin filaments are attached to each other by cross-bridges that contain ATPase, which breaks down adenosine triphosphate (ATP) to provide the energy for contraction.

Two chains of actin molecules form a helical structure, with another molecule, tropomyosin, in the grooves of the

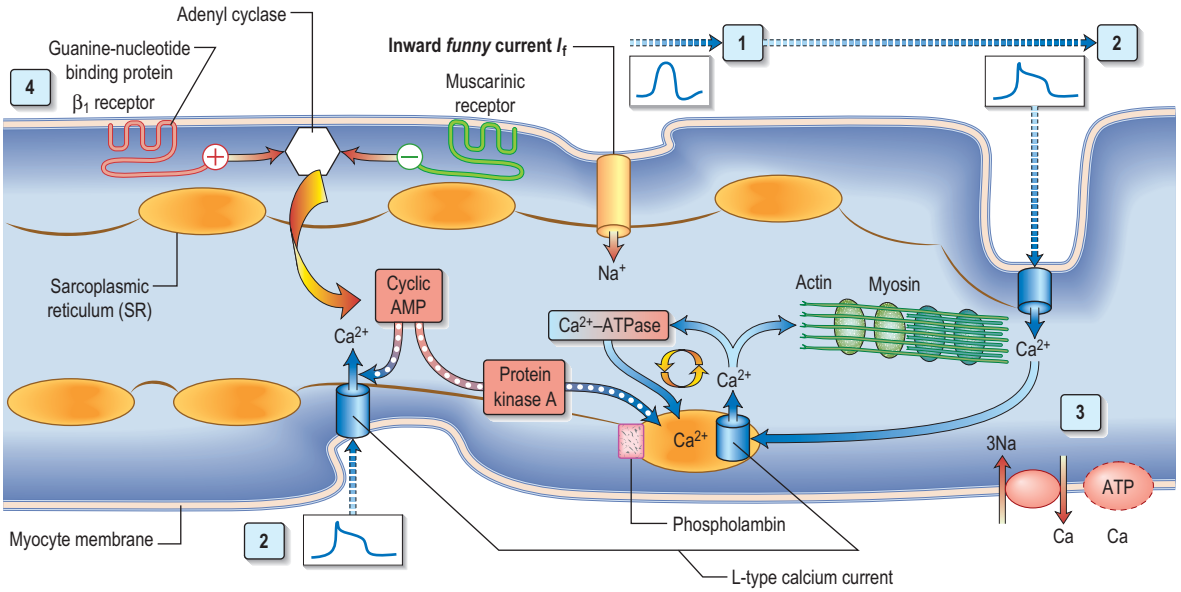


Fig. 13.3 **The 'complete' cardiac cell. 1.** Spontaneous depolarization in sinus node cells due to sodium (Na) influx through the 'funny' current generates the 'pacemaker' potential. **2.** This activates other atrial and ventricular myocytes, triggering action potentials and activating L-type calcium (Ca) channels in the surface and transverse tubule membranes (at top and bottom of figure). **3.** The resulting Ca influx acts on Ca-induced Ca release channels (RyR2) on the sarcoplasmic reticulum (SR), resulting in release of stored Ca, which acts on actin and myosin fibrils, resulting in contraction. Ca reuptake pumps in the SR, regulated by phospholamban, replenish the stores; various exchange pumps also expel Ca from the cell. **4.** Autonomic input has either a positive chronotropic/inotropic effect (β_1 receptors) or a negative chronotropic/inotropic effect (muscarinic receptors).

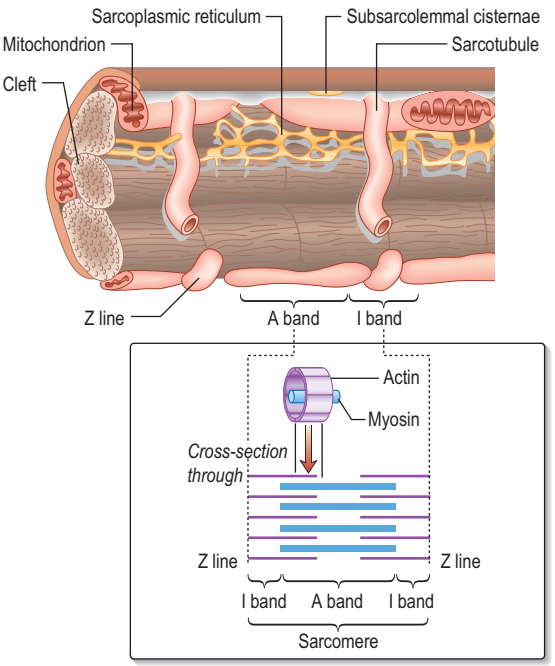


Fig. 13.4 **Schematic diagram showing the structure of a myofibril within a myocyte.** The myofibrils are made up of a series of sarcomeres joined at the Z line.

actin helix, and a further molecule, troponin, is attached to every seven actin molecules. During cardiac contraction the length of the actin and myosin monofilaments does not change. Rather, the actin filaments slide between the myosin filaments when ATPase splits a high-energy bond of ATP. To supply the ATP, the myocyte (which cannot stop for a rest) has a very high mitochondrial density (35% of the cell volume). As calcium ions bind to troponin C, the activity of troponin I is inhibited, which induces a conformational change in tropomyosin. This event unlocks the active site between actin and myosin, enabling contraction to proceed.

Calcium is made available during the plateau phase of the action potential by calcium ions entering the cell and by being mobilized from the sarcoplasmic reticulum through the ryanodine receptor (RyR2) calcium-release channel. RyR2 activity is regulated by the protein calstabin 2 (see p. 791) and nitric oxide. The force of cardiac muscle contraction ('inotropic state') is thus regulated by the influx of calcium ions into the cell through calcium channels (Fig. 13.3). T (transient) calcium channels open when the muscle is more depolarized, whereas L (long-lasting) calcium channels require less depolarization. The extent to which the sarcomere can shorten determines the stroke volume of the ventricle. It is maximally shortened in response to powerful inotropic drugs or severe exercise.

Starling's law of the heart

The contractile function of an isolated strip of cardiac tissue can be described by the relationship between the velocity of muscle contraction, the load that may be moved by the contracting muscle, and the extent to which the muscle is stretched before contracting. As with all other types of muscle, the velocity of contraction of myocardial tissue is reduced by increasing the load against which the tissue must contract. However, in the non-failing heart, pre-stretching of cardiac muscle improves the relationship between the force and velocity of contraction (Fig. 13.5).

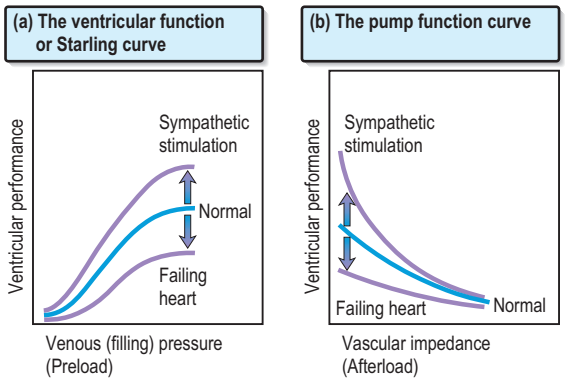


Fig. 13.5 **The Frank–Starling mechanism,** showing the effect on ventricular contraction of alteration in filling pressures and outflow impedance in the normal, failing and sympathetically stimulated ventricle.

This phenomenon was described in the intact heart as an increase of stroke volume (ventricular performance) with an enlargement of the diastolic volume (preload), and is known as 'Starling's law of the heart' or the 'Frank–Starling relationship'. It has been transcribed into more clinically relevant indices. Thus, stroke work (aortic pressure × stroke volume) is increased as ventricular end-diastolic volume is raised. Alternatively, within certain limits, cardiac output rises as pulmonary capillary wedge pressure increases. This clinical relationship is described by the ventricular function curve (Fig. 13.5), which also shows the effect of sympathetic stimulation.

Nerve supply of the myocardium

Adrenergic nerves supply atrial and ventricular muscle fibres as well as the conduction system. β_1-Receptors predominate in the heart with both epinephrine (adrenaline) and norepinephrine (noradrenaline) having positive inotropic and chronotropic effects. Cholinergic nerves from the vagus supply mainly the SA and AV nodes via M_2 muscarinic receptors. The ventricular myocardium is sparsely innervated by the vagus. Under basal conditions, vagal inhibitory effects predominate over the sympathetic excitatory effects, resulting in a slow heart rate.

Adrenergic stimulation and cellular signalling

β_1-Adrenergic stimulation enhances Ca^{2+} flux in the myocyte and thereby strengthens the force of contraction (Fig. 13.3). Binding of catecholamines, e.g. norepinephrine (noradrenaline), to the myocyte β_1-adrenergic receptor stimulates membrane-bound adenylate kinases. These enzymes enhance production of cyclic adenosine monophosphate (cAMP) that activates intracellular protein kinases, which in turn phosphorylate cellular proteins, including L-type calcium channels within the cell membrane. β_1-Adrenergic stimulation of the myocyte also enhances myocyte relaxation.

The return of calcium from the cytosol to the sarcoplasmic reticulum (SR) is regulated by phospholamban (PL), a low-molecular-weight protein in the SR membrane. In its dephosphorylated state, PL inhibits Ca^{2+} uptake by the SR ATPase pump (Fig. 13.3). However, β_1-adrenergic activation of protein kinase phosphorylates PL, and blunts its inhibitory effect. The subsequently greater uptake of Ca^{2+} ions by the SR hastens Ca^{2+} removal from the cytosol and promotes myocyte relaxation.

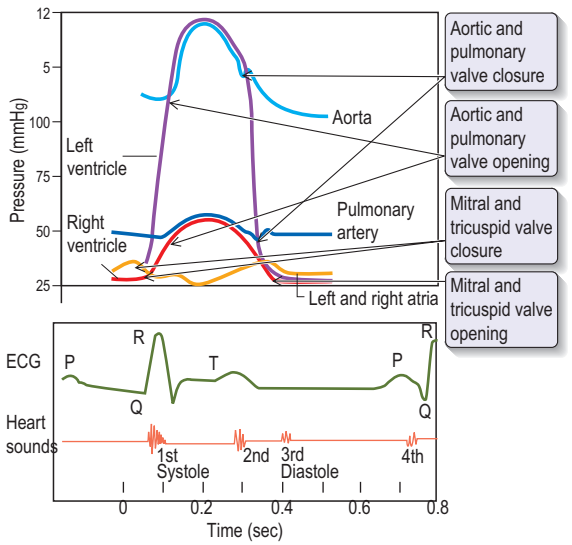


Fig. 13.6 **The cardiac cycle.**

The increased cAMP activity also results in phosphorylation of troponin I, an action that inhibits actin–myosin interaction, and further enhances myocyte relaxation.

The cardiac cycle

The cardiac cycle (Fig. 13.6) consists of precisely timed rhythmic electrical and mechanical events that propel blood into the systemic and pulmonary circulations. The first event in the cardiac cycle is atrial depolarization (a P wave on the surface ECG) followed by right atrial and then left atrial contraction. Ventricular activation (the QRS complex on the ECG) follows after a short interval (the PR interval). Left ventricular contraction starts and shortly thereafter right ventricular contraction begins. The increased ventricular pressures exceed the atrial pressures, and close first the mitral and then the tricuspid valves.

Until the aortic and pulmonary valves open, the ventricles contract with no change of volume (isovolumetric contraction). When ventricular pressures rise above the aortic and pulmonary artery pressures, the pulmonary valve and then the aortic valve open and ventricular ejection occurs. As the ventricles begin to relax, their pressures fall below the aortic and pulmonary arterial pressures, and aortic valve closure is followed by pulmonary valve closure. Isovolumetric relaxation then occurs. After the ventricular pressures have fallen below the right atrial and left atrial pressures, the tricuspid and mitral valves open. The cardiac cycle can be graphically depicted as the relationship between the pressure and volume of the ventricle. This is shown in Figure 13.7, which illustrates the changing pressure–volume relationships in response to increased contractility and to exercise.

The coronary circulation

The coronary arterial system (Fig. 13.8) consists of the right and left coronary arteries. These arteries branch from the aorta, arising immediately above two cusps of the aortic valve. These arteries are unique in that they fill during diastole, when not occluded by valve cusps and when not squeezed by myocardial contraction. The right coronary artery arises from the right coronary sinus and courses through the right side of the AV groove, giving off vessels

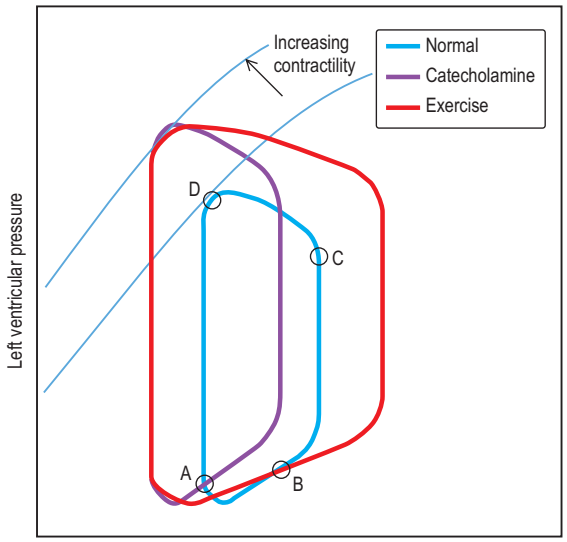


Fig. 13.7 **Pressure–volume loop.** AB, diastole – ventricular filling; BC systole – isovolumetric ventricular contraction; CD, systole – ventricular emptying; DA, diastole – isovolumetric ventricular relaxation.

that supply the right atrium and the right ventricle. The vessel usually continues as the posterior descending coronary artery, which runs in the posterior interventricular groove and supplies the posterior part of the interventricular septum and the posterior left ventricular wall.

Within 2.5 cm of its origin from the left coronary sinus, the left main coronary divides into the left anterior descending artery and the circumflex artery. The left anterior descending artery runs in the anterior interventricular groove and supplies the anterior septum and the anterior left ventricular wall. The left circumflex artery travels along the left AV groove and gives off branches to the left atrium and the left ventricle (marginal branches).

The sinus node and the AV node are supplied by the right coronary artery in about 60% and 90% of people, respectively. Therefore, disease in this artery may cause sinus bradycardia and AV nodal block. The majority of the left ventricle is supplied by the left coronary artery, so that stenosis in the left main artery is extremely dangerous; total obstruction of this vessel is rarely compatible with life.

Some blood from the capillary beds in the wall of the heart drains directly into the cavities of the heart by tiny veins, but the majority returns by veins which accompany the arteries, to empty into the right atrium via the coronary sinus. An extensive lymphatic system drains into vessels that travel along the coronary vessels and then into the thoracic duct.

Blood vessel control and functions of the vascular endothelium

In functional terms, the tunica intima with the vascular endothelium and the smooth-muscle-cell-containing tunica media are the main constituents of blood vessels. These two structures are closely interlinked by a variety of mechanisms to regulate vascular tone. The central control of blood vessels is achieved via the neuroendocrine system. Sympathetic vasoconstrictor and parasympathetic vasodilator nerves regulate vascular tone in response to daily activity. Where

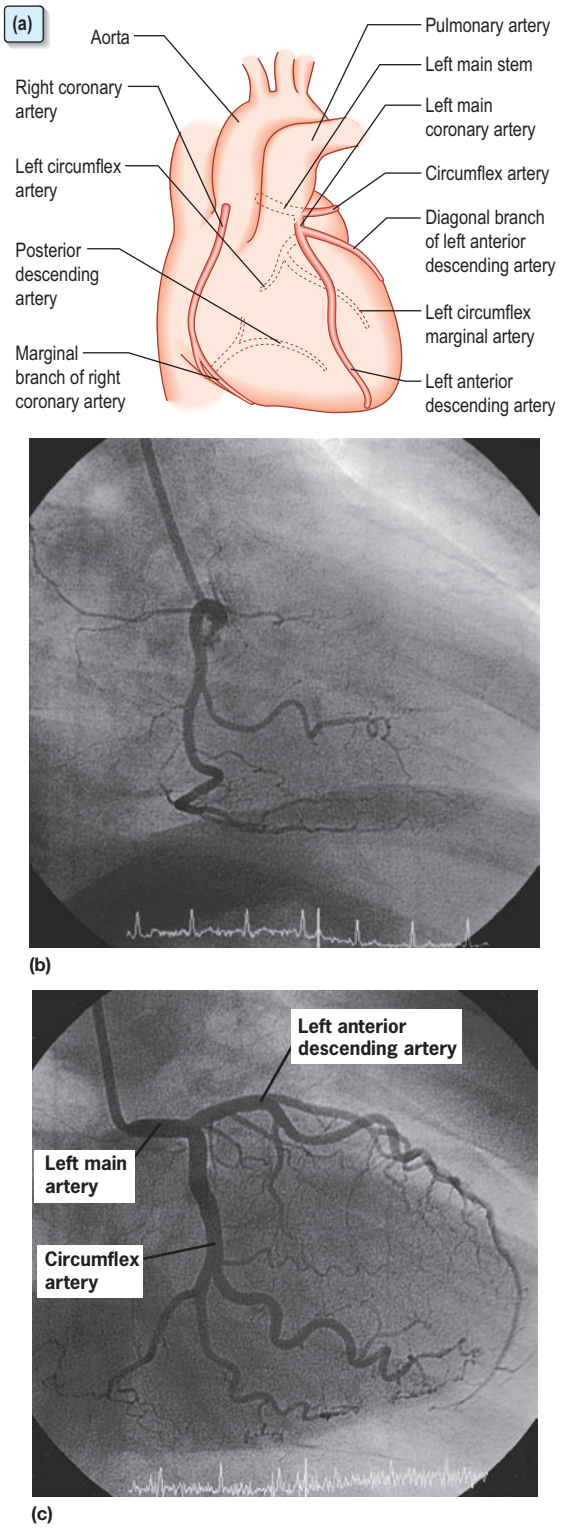


Fig. 13.8 **(a) Diagram of the normal coronary arterial anatomy. (b) Angiogram of non-dominant right coronary system. (c) Angiogram of dominant left coronary system from the same patient.** Right anterior oblique projection.

neural control is impaired, or in various pathological states, e.g. haemorrhage, endocrine control of blood vessels mediated through epinephrine (adrenaline), angiotensin and vasopressin takes over.

At a local level, tissue perfusion is maintained automatically and by the effect of various factors synthesized and/or

released in the immediate vicinity. In the face of fluctuating arterial pressures, blood vessels vasoconstrict independently of nervous input when blood pressure drops and vice versa. This process of *autoregulation* is a consequence of:

- the *Bayliss myogenic* response – the ability of blood vessels to constrict when distended
- the *vasodilator washout* effect – the vasoconstriction triggered by a decrease in the concentration of tissue metabolites.

The vascular endothelium is a cardiovascular endocrine organ, which occupies a strategic interface between blood and other tissues. It produces various compounds (e.g. nitrous oxide (NO), prostacyclin (PGI_2), endothelin, endothelial-derived hyperpolarizing factor (ERHF), adhesion molecules, vascular endothelial growth factor (VEGF)) and has enzymes located on the surface controlling the levels of circulating compounds such as angiotensin, bradykinin and serotonin. It has many regulatory roles:

Vasomotor control

NO is a diffusible gas with a very short half-life, produced in endothelial cells from the amino acid L-arginine via the action of the enzyme NO synthase (NOS), which is controlled by cytoplasmic calcium/calmodulin (Fig. 13.9). It is produced in response to various stimuli (Table. 13.1), triggering vascular smooth muscle relaxation through activation of guanylate cyclase, leading to an increase in the intracellular levels of cyclic 3,5-guanine monophosphate (cGMP). Its cardiovascular effects protect against atherosclerosis, high blood

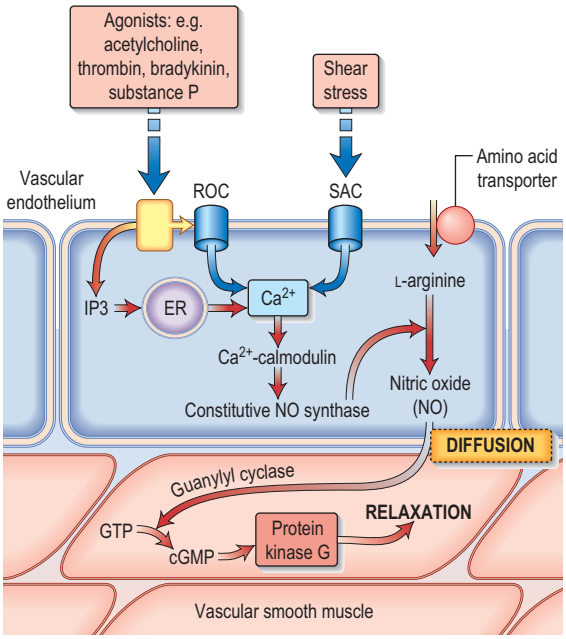


Fig. 13.9 **Nitric oxide (NO): the stimulus for production and function of NO.** Various stimuli lead to the production of NO via cytoplasmic calcium/calmodulin. NO triggers smooth muscle relaxation via the activation of guanylyl cyclase. ROC, receptor-operated Ca^{2+} channel; SAC, stretch-activated Ca^{2+} channel; IP3, inositol triphosphate; ER, endoplasmic reticulum; GTP, guanine triphosphate; cGMP, cyclic guanine monophosphate.

Table 13.1 Some of the products and functions of the vascular endothelium

Endothelial product	Function(s)	Stimulus
Nitrous oxide	Vasodilation Inhibits platelet aggregation Inhibits transcription of adhesion molecules Inhibits vascular smooth muscle proliferation	Shear stress, e.g. induced by exercise Agonists: thrombin, acetylcholine, endothelin, bradykinin, serotonin, substance P Inflammation/endotoxin shock
Prostacyclin (PGI_2)	Vasodilation Inhibits platelet aggregation	Agonist: thrombin Inflammation
Prostanoids	Vasoconstriction	Hypoxia
Endothelin	Vasoconstriction	Thrombin, angiotensin II, vasopressin Hypoxia N.B.: Inhibited by shear stress
Endothelial-derived hyperpolarizing factor	Vasodilation	Agonists: bradykinin, acetylcholine
Angiotensin-converting enzyme	Vasoconstriction	Expressed naturally
von Willebrand factor	Promotes platelet aggregation Stabilizes factor VIII	Agonists: thrombin, epinephrine (adrenaline)
Adhesion molecules		
P, L, E selectins ICAM, VCAM, PECAM	Margination of while blood cells (WBCs) Binding and diapedesis of WBCs into vessel wall	Inflammatory mediators: histamine, thrombin, TNF, IL-6
Vascular endothelial growth factor (VEGF)	Angiogenesis Vasodilation Increased vascular permeability	Pregnancy Hypoxia Inflammation: rheumatoid arthritis, trauma, tumours

ICAM, intracellular adhesion molecule; VCAM, vascular cell adhesion molecule, PECAM, platelet/endothelial cell adhesion molecule; TNF, tumour necrosis factor; IL, interleukin.

pressure, heart failure and thrombosis. NO is also the neurotransmitter in various 'nitrergic' nerves in the central and peripheral nervous systems and may play a role in the central regulation of vascular tone. The new class of drugs used to treat erectile dysfunction, the phosphodiesterase (PDE_5) inhibitors, prevent the breakdown of cGMP and promote vasodilatation.

PGI_2 is synergistic to NO and also plays a role in the local regulation of vasomotor tone.

Endothelin is a 21-amino-acid peptide that counteracts the effects of NO. Its production is inhibited by shear stress, i.e. the stress exerted on the vessel wall by the flowing blood, and it causes profound vasoconstriction and vascular smooth muscle hypertrophy. It is thought to play a role in the genesis of hypertension and atheroma.

Angiotensin-converting enzyme located on the endothelial cell membrane converts circulating angiotensin I (synthesized by the action of renin on angiotensinogen) to angiotensin II which has vasoconstrictor properties and leads to aldosterone release (Fig. 18.27). Aldosterone promotes sodium absorption from the kidney and together with the angiotensin-induced vasoconstriction provides haemodynamic stability.

Other factors which influence vasomotor tone include histamine (released by mast cells), bradykinin (synthesized from kininogen by the action of coagulation factor XIIa) and serotonin released by platelets.

Anti- and pro-thrombotic mechanisms

PGI_2, produced from arachidonic acid in the endothelial cell membrane by the action of the enzyme cyclo-oxygenase (Fig. 14.32), inhibits platelet aggregation. Low-dose aspirin prevents activation of the cyclo-oxygenase pathway in platelets but only to a degree that does not affect PGI_2 synthesis, unlike higher doses. Other antithrombotic agents such as clopidogrel (ADP receptor antagonist) and glycoprotein IIb/IIIa inhibitors achieve their effects by acting directly on platelet receptors. The antithrombotic effect of PGI_2 is aided by NO, affecting platelets via activation of guanylate cyclase. The endothelial cell membrane also produces other anticoagulant molecules such as thrombomodulin, heparin sulphate and various fibrinolytic factors. Clinically used, fast-acting, heparin preparations are identical to this naturally occurring molecule.

In addition to their ability to prevent clotting, endothelial cells also aid thrombosis. They are responsible for the production of von Willebrand factor through a unique organelle called the Weibel–Palade body, which not only acts as a carrier for factor VIII but also promotes platelet adhesion by binding to exposed collagen (p. 432).

Modulation of immune responses

In response to various inflammatory mediators, the vascular endothelium expresses various so-called 'adhesion molecules' which promote leucocyte attraction, adhesion and infiltration into the blood vessel wall (p. 67).

Regulation of vascular cell growth

The endothelial cells are also responsible for the development of new blood vessels ('angiogenesis') in the placenta, wound healing, tissue repair and tumour growth. This process is facilitated by VEGF.

Fetal circulation

In utero, the pulmonary circulation is largely unnecessary because fetal blood is oxygenated by placental blood flow, a parallel and integral element in the systemic circulation. In the fetus, systemic venous blood returning to the right atrium is partly deflected through the foramen ovale to the left atrium. Blood that passes through the right ventricle is diverted from the pulmonary artery to the aorta through the ductus arteriosus (Fig. 13.10a). Thus, the systemic

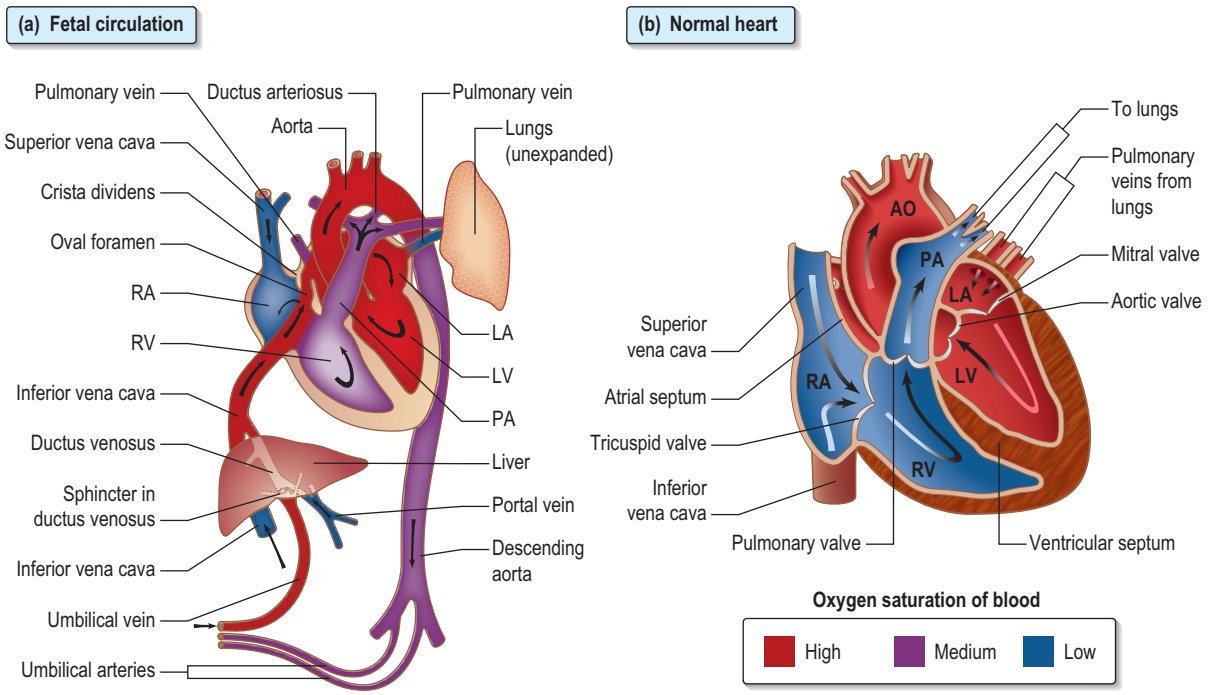


Fig. 13.10 **Anatomy showing circulation: (a) fetal; (b) after birth (normal heart).** AO, aorta; LA, left atrium; LV, left ventricle; PA, pulmonary artery; RA, right atrium; RV, right ventricle.

FURTHER READING

Levick JR. *An Introduction to Cardiovascular Physiology*, 4th edn. London: Arnold, 2003.

Nabel EG. Cardiovascular disease. *New England Journal of Medicine* 2003; **349**: 60–72.

venous return, which is a mixture of oxygenated and deoxygenated blood, is mostly returned to the systemic arterial system.

At birth, inspiration dilates the pulmonary arterioles, resulting in a dramatic reduction of pulmonary vascular resistance. Blood therefore flows through the pulmonary circulation. The increased oxygen tension and reduced levels of prostaglandins trigger closure of the ductus arteriosus, and the reduced right atrial pressure and increasing left atrial pressure tend to close the foramen ovale. Thus, the circulation is divided into two separate circuits connected in series (Fig. 13.10b).

In the fetus the left and right heart both propel blood from the systemic veins to the systemic arteries; thus, severe abnormalities of the heart may not compromise fetal blood flow.

SYMPTOMS OF HEART DISEASE

The following symptoms occur with heart disease:

- chest pain
- dyspnoea
- palpitations
- syncope
- fatigue
- peripheral oedema.

The severity of cardiac symptoms or fatigue may be classified according to the New York Heart Association (NYHA) grading of 'cardiac status' (Table 13.2). The differential diagnosis of chest pain is given in Table 13.3.

Central chest pain

This is the most common symptom associated with heart disease. The pain of angina pectoris and myocardial infarction is due to myocardial hypoxia. Types of pain include:

- retrosternal heavy or gripping sensation with radiation to the left arm or neck that is provoked by exertion and eased with rest or nitrates – *angina* (p. 748)
- similar pain at rest – *acute coronary syndrome* (p. 752)
- severe tearing chest pain radiating through to the back – *aortic dissection* (p. 808)
- sharp central chest pain that is worse with movement or respiration but relieved with sitting forward – *pericarditis pain* (p. 795)
- sharp stabbing left submammary pain associated with anxiety – *Da Costa's syndrome.*

Table 13.2 The New York Heart Association grading of 'cardiac status' (modified)

Grade 1	Uncompromised (no breathlessness)
Grade 2	Slightly compromised (on severe exertion)
Grade 3	Moderately compromised (on mild exertion)
Grade 4	Severely compromised (breathless at rest)

Dyspnoea

Left ventricular failure causes dyspnoea due to oedema of the pulmonary interstitium and alveoli. This makes the lungs stiff (less compliant), thus increasing the respiratory effort required to ventilate the lungs. *Tachypnoea* (increased respiratory rate) is often present owing to stimulation of pulmonary stretch receptors.

Orthopnoea refers to breathlessness on lying flat. Blood is redistributed from the legs to the torso, leading to an increase in central and pulmonary blood volume. The patient uses an increasing number of pillows to sleep.

Paroxysmal nocturnal dyspnoea (PND) is when a patient is woken from sleep fighting for breath. It is due to the same mechanisms as orthopnoea. However, as sensory awareness is reduced whilst asleep, the pulmonary oedema can become quite severe before the patient is awoken.

Hyperventilation with alternating episodes of apnoea (*Cheyne–Stokes respiration*) occurs in severe heart failure.

Central sleep apnoea syndrome (CSAS). If hypopnoea occurs rather than apnoea, the phenomenon is termed periodic breathing, but the two variations are known together as CSAS. This occurs due to malfunctioning of the respiratory centre in the brain, caused by poor cardiac output with concurrent cerebrovascular disease.The symptoms of CSAS, such as daytime somnolence and fatigue, are similar to those of obstructive sleep apnoea syndrome (OSAS, p. 842) and there is considerable overlap with the symptoms of heart failure. CSAS is believed to lead to myocardial hypertrophy and fibrosis, deterioration in cardiac function and complex arrhythmias, including non-sustained ventricular tachycardia, hypertension and stroke. Patients with CSAS have a worse prognosis compared to similar patients without CSAS.

Palpitations

These represent an increased awareness of the normal heart beat or the sensation of slow, rapid or irregular heart rhythms.

Table 13.3 Differential diagnosis of chest pain

Central	Lateral/peripheral
Cardiac	**Pulmonary**
Ischaemic heart disease (infarction or angina)	Infarction
Coronary artery spasm	Pneumonia
Pericarditis/myocarditis	Pneumothorax
Mitral valve prolapse	Lung cancer
Aortic aneurysm/dissection	Mesothelioma
Non-cardiac	**Non-pulmonary**
Pulmonary embolism	Bornholm disease (epidemic myalgia)
Oesophageal disease (see Box 6.2)	Herpes zoster
Mediastinitis	Trauma (ribs/muscular)
Costochondritis (Tietze's disease)	
Trauma (soft tissue, rib)	

The most common arrhythmias felt as palpitations are premature ectopic beats and paroxysmal tachycardias. A useful trick is to ask patients to tap out the rate and rhythm of their palpitations, as the different arrhythmias have different characteristics:

- *Premature beats* are felt by the patient as a pause followed by a forceful beat. This is because premature beats are usually followed by a pause before the next normal beat, as the heart resets itself. The next beat is more forceful as the heart has had a longer diastolic period and therefore is filled with more blood before this beat.
- *Paroxysmal tachycardias* (see p. 722) are felt as a sudden racing heart beat.
- *Bradycardias* (p. 717) may be appreciated as slow, regular, heavy or forceful beats. Most often, however, they are simply not sensed. All palpitations may be graded by the NYHA cardiac status (Table 13.2).

Syncope

This is a transient loss of consciousness due to inadequate cerebral blood flow. The cardiovascular causes are listed in Table 13.4.

Vascular:

- A *vasovagal attack* (p. 1144) is a simple faint and is the most common cause of syncope. The mechanism begins with peripheral vasodilatation and venous pooling of blood, leading to a reduction in the amount of blood returned to the heart. The near-empty heart responds by contracting vigorously, which in turn stimulates mechanoreceptors (stretch receptors) in the inferoposterior wall of the left ventricle. These in turn trigger reflexes via the central nervous system, which act to reduce ventricular stretch (i.e. further vasodilatation and sometimes profound bradycardia), but this causes a drop in blood pressure and therefore syncope. These episodes are usually associated with a prodrome of dizziness, nausea, sweating, tinnitus, yawning and a sinking feeling. Recovery occurs within a few seconds, especially if the patient lies down.
- *Postural (orthostatic) hypotension* is a drop in systolic blood pressure of 20 mmHg or more on standing from a sitting or lying position. Usually, reflex vasoconstriction prevents a drop in pressure but if this is absent or the patient is fluid depleted or on *vasodilating or diuretic drugs*, hypotension occurs.
- *Postprandial hypotension* is a drop in systolic blood pressure of 20 mmHg or more or the systolic blood pressure drops from above 100 mmHg to under 90 mmHg within 2 hours of eating. The mechanism is unknown but may be due to pooling of blood in the splanchnic vessels. In normal subjects, this elicits a homeostatic response via activation of baroreceptors and the sympathetic system, peripheral vasoconstriction and an increase in cardiac output.
- *Micturition syncope* refers to loss of consciousness whilst passing urine.
- *Carotid sinus syncope* occurs when there is an exaggerated vagal response to carotid sinus stimulation, provoked by wearing a tight collar, looking upwards or turning the head.

Obstructive. The obstructive cardiac causes listed in Table 13.4 all lead to syncope due to restriction of blood flow from the heart into the rest of the circulation, or between the different chambers of the heart.

Arrhythmias. *Stokes–Adams attacks* (p. 719) are a sudden loss of consciousness unrelated to posture and due to intermittent high-grade AV block, profound bradycardia or ventricular standstill. The patient falls to the ground without warning, is pale and deeply unconscious. The pulse is usually very slow or absent. After a few seconds the patient flushes brightly and recovers consciousness as the pulse quickens. Often there are no sequelae, but patients may injure themselves during falls. Occasionally a generalized convulsion may occur if the period of cerebral hypoxia is prolonged, leading to a misdiagnosis of epilepsy.

Table 13.4 Cardiovascular causes of syncope

Vascular
Neurocardiogenic (vasovagal)
Postural hypotension
Postprandial hypotension
Micturition syncope
Carotid sinus syncope
Obstructive
Aortic stenosis
Hypertrophic cardiomyopathy
Pulmonary stenosis
Tetralogy of Fallot
Pulmonary hypertension/embolism
Atrial myxoma/thrombus
Defective prosthetic valve
Arrhythmias
Rapid tachycardias
Profound bradycardias (Stokes–Adams)
Significant pauses (in rhythm)
Artificial pacemaker failure

Fatigue

Fatigue may be a symptom of inadequate systemic perfusion in heart failure. It is due to poor sleep, a direct side-effect of medication, particularly beta-blockers, electrolyte imbalance due to diuretic therapy and as a systemic manifestation of infection such as endocarditis.

Peripheral oedema

Heart failure results in salt and water retention due to renal underperfusion and consequent activation of the renin–angiotensin–aldosterone system (p. 1023). This leads to dependent pitting oedema.

FURTHER READING

Freeman R. Neurogenic orthostatic hypotension. *New England Journal of Medicine* 2008; **358**: 615–624.

EXAMINATION OF THE CARDIOVASCULAR SYSTEM

GENERAL EXAMINATION

General features of the patient's well-being should be noted as well as the presence of conjunctival pallor, obesity, jaundice and cachexia.

- **Clubbing** (p. 820) is seen in congenital cyanotic heart disease, particularly Fallot's tetralogy and also in 10% of patients with subacute infective endocarditis.
- **Splinter haemorrhages.** These small, subungual linear haemorrhages are frequently due to trauma, but are also seen in infective endocarditis.
- **Cyanosis** is a dusky blue discoloration of the skin (particularly at the extremities) or of the mucous membranes when the capillary oxygen saturation is less than 85%. *Central cyanosis* (p. 820) is seen with shunting of deoxygenated venous blood into the systemic circulation, as in the presence of a right-to-left heart shunt. *Peripheral cyanosis* is seen in the hands and feet, which are cold. It occurs in conditions associated with peripheral vasoconstriction and stasis of blood in the extremities leading to increased peripheral oxygen extraction. Such conditions include congestive heart failure, circulatory shock, exposure to cold temperatures and abnormalities of the peripheral circulation.

THE ARTERIAL PULSE

The first pulse to be examined is the right radial pulse. A delayed femoral pulsation occurs because of a proximal stenosis, particularly of the aorta (coarctation).

Pulse rate

The pulse rate should be between 60 and 80 beats per minute (b.p.m.) when an adult patient is lying quietly in bed.

Rhythm

The rhythm is regular except for a slight quickening in early inspiration and a slowing in expiration (sinus arrhythmia).

Premature beats occur as occasional or repeated irregularities superimposed on a regular pulse rhythm. Similarly, intermittent heart block is revealed by occasional beats dropped from an otherwise regular rhythm.

Atrial fibrillation produces an irregularly irregular pulse. This irregular pattern persists when the pulse quickens in response to exercise, in contrast to pulse irregularity due to ectopic beats, which usually disappears on exercise.

Character of pulse (Fig. 13.11)

Carotid pulsations are not normally apparent on inspection of the neck but may be visible (Corrigan's sign) in conditions associated with a large-volume pulse, including high output states (such as thyrotoxicosis, anaemia or fever) and in aortic regurgitation.

A *'collapsing'* or *'water hammer'* pulse is a large-volume pulse characterized by a short duration with a brisk rise and fall. This is best appreciated by palpating the radial artery with the palmer aspect of four fingers whilst elevating the patient's arm above the level of the heart. A collapsing pulse is characteristic of aortic valvular regurgitation or a persistent ductus arteriosus.

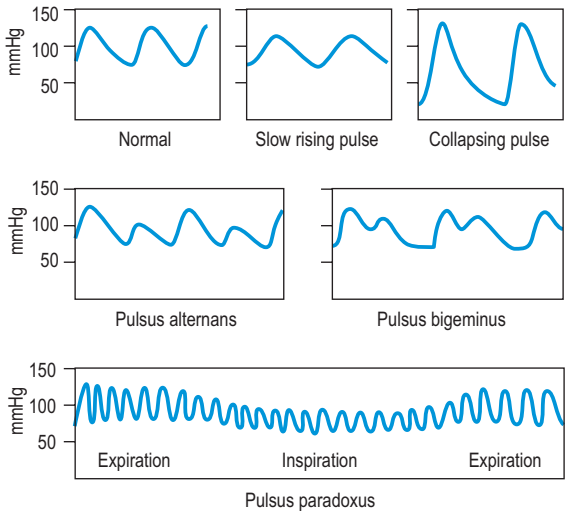


Fig. 13.11 **Arterial waveforms.**

A *small-volume pulse* is seen in cardiac failure, shock and obstructive valvular or vascular disease. It may also be present during tachyarrhythmias.

A *plateau pulse* is small in volume and slow in rising to a peak due to aortic stenosis (Fig. 13.11).

Paradoxical pulse

Paradoxical pulse is a misnomer, as it is actually an exaggeration of the normal pattern (Fig. 13.11). In normal subjects, the systolic pressure and the pulse pressure (the difference between the systolic and diastolic blood pressures) fall during inspiration. The normal fall of systolic pressure is less than 10 mmHg, and this can be measured using a sphygmomanometer. It is due to increased pulmonary intravascular volume during inspiration. In severe airflow limitation (especially severe asthma) there is an increased negative intrathoracic pressure on inspiration which enhances the normal fall in blood pressure. In patients with cardiac tamponade, the fluid in the pericardium increases the intrapericardial pressure, thereby impeding diastolic filling of the heart. The normal inspiratory increase in venous return to the right ventricle is at the expense of the left ventricle, as both ventricles are confined by the accumulated pericardial fluid within the pericardial space. Paradox can occur through a similar mechanism in constrictive pericarditis but is less common.

Alternating pulse (pulsus alternans)

This is characterized by regular alternate beats that are weak and strong. It is a feature of severe myocardial failure and is due to the prolonged recovery time of damaged myocardium; it indicates a very poor prognosis. It is easily noticed when taking the blood pressure because the systolic pressure may vary from beat to beat by as much as 50 mmHg (Fig. 13.11).

Bigeminal pulse (pulsus bigeminus)

This is due to a premature ectopic beat following every sinus beat. The rhythm is not regular (Fig. 13.11) because every weak pulse is premature.

Ask the authors
www.studentconsult.com

Practical Box 13.1 Taking the blood pressure

Use a properly calibrated machine.

1. The blood pressure is taken in the (right) arm with the patient relaxed and comfortable.
2. The sphygmomanometer cuff is wrapped around the upper arm with the inflation bag placed over the brachial artery.
3. The cuff is inflated until the pressure exceeds the arterial pressure – when the radial pulse is no longer palpable.
4. The diaphragm of the stethoscope is positioned over the brachial artery just below the cuff.
5. The cuff pressure is slowly reduced until sounds (Korotkoff sounds) can be heard (*phase 1*). This is the **systolic pressure**.
6. The pressure is allowed to fall further until the Korotkoff sounds become suddenly muffled (*phase 4*).
7. The pressure is allowed to fall still further until they disappear (*phase 5*).

The **diastolic pressure** is usually taken as *phase 5* because this phase is more reproducible and nearer to the intravascular diastolic pressure. The Korotkoff sounds may disappear (*phase 2*) and reappear (*phase 3*) between the systolic and diastolic pressures. Do not mistake *phase 2* for the diastolic pressure or *phase 3* for the systolic pressure.

Pulsus bisferiens

This is a pulse that is found in hypertrophic cardiomyopathy and in mixed aortic valve disease (regurgitation combined with stenosis). The first systolic wave is the 'percussion' wave produced by the transmission of the left ventricular pressure in early systole. The second peak is the 'tidal' wave caused by recoil of the vascular bed. This normally happens in diastole (the dicrotic wave), but when the left ventricle empties slowly or is obstructed from emptying completely, the tidal wave occurs in late systole. The result is a palpable double pulse (Fig. 13.11).

THE BLOOD PRESSURE

The peak systemic arterial blood pressure is produced by transmission of left ventricular systolic pressure. Vascular tone and an intact aortic valve maintain the diastolic blood pressure. How to take the blood pressure is outlined in Practical box 13.1.

JUGULAR VENOUS PRESSURE

There are no valves between the internal jugular vein and the right atrium. Observation of the column of blood in the internal jugular system is therefore a good measure of right atrial pressure. The external jugular cannot be relied upon because of its valves and because it may be obstructed by the fascial and muscular layers through which it passes; it can only be used if typical venous pulsation is seen, indicating no obstruction to flow.

Measurement of jugular venous pressure (JVP) (Practical box 13.2)

Elevation of the jugular venous pressure occurs in *heart failure*. It is also produced by an elevated jugular venous pressure which occurs in constrictive pericarditis and cardiac tamponade (increases in inspiration – *Kussmaul's sign*), renal disease with salt and water retention, overtransfusion or excessive infusion of fluids, congestive cardiac failure and superior vena cava obstruction. A reduced jugular venous pressure occurs in hypovolaemia.

Practical Box 13.2 Measurement of jugular venous pressure

The patient is positioned at about 45° to the horizontal (between 30° and 60°), wherever the top of the venous pulsation can be seen in a good light.

The jugular venous pressure is measured as the vertical distance between the manubriosternal angle and the top of the venous column.

The normal jugular venous pressure is usually less than 3 cmH_2O, which is equivalent to a right atrial pressure of 8 cmH_2O when measured with reference to a point midway between the anterior and posterior surfaces of the chest.

The venous pulsations are not usually palpable (except for the forceful venous distension associated with tricuspid regurgitation).

Compression of the right upper abdomen causes a temporary increase in venous pressure and makes the JVP more visible (hepatojugular reflux).

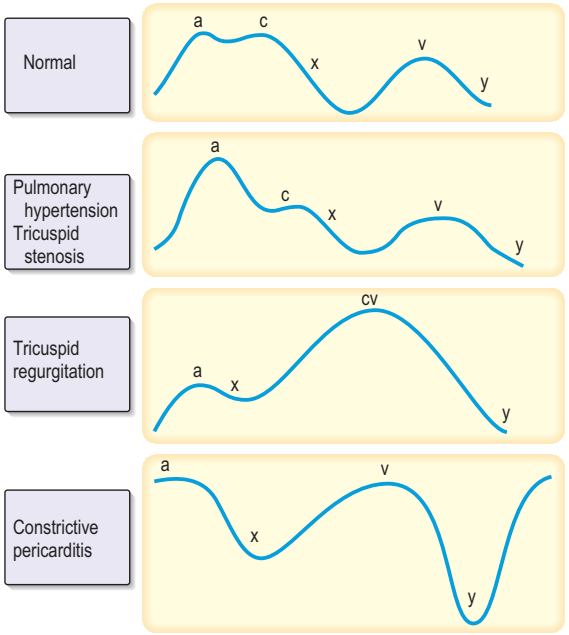


Fig. 13.12 **Jugular venous waveforms.**

The jugular venous pressure wave

This consists of three peaks and two troughs (Fig. 13.12). The peaks are described as *a*, *c* and *v* waves and the troughs are known as *x* and *y* descents:

- The ***a* wave** is produced by atrial systole and is increased with right ventricular hypertrophy secondary to pulmonary hypertension or pulmonary stenosis. Giant cannon waves occur in complete heart block and ventricular tachycardia.
- The ***x* descent** occurs when atrial contraction finishes.
- The ***c* wave** occurs during the *x* descent and is due to transmission of right ventricular systolic pressure before the tricuspid valve closes.
- The ***v* wave** occurs with venous return filling the right atrium. Giant *v* waves occur in tricuspid regurgitation.

- The ***y* descent** follows the *v* wave when the tricuspid valve opens. A steep *y* descent is seen in constrictive pericarditis and tricuspid incompetence.

EXAMINATION OF THE PRECORDIUM

- With the patient at 45°, the cardiac apex is located in the 5th intercostal space mid-clavicular line. Left ventricular dilatation will displace the apex downwards and laterally. It may be impalpable in patients with emphysema, obesity, pericardial or pleural effusions.
- A *tapping apex* is a palpable first sound and occurs in mitral stenosis.
- A *vigorous apex* may be present in diseases with volume overload, e.g. aortic regurgitation. A *heaving apex* may occur with left ventricular hypertrophy – aortic stenosis, systemic hypertension and hypertrophic cardiomyopathy.
- A *double pulsation* may occur in hypertrophic cardiomyopathy.
- A *sustained left parasternal heave* occurs with right ventricular hypertrophy or left atrial enlargement.
- A *palpable thrill* may be felt overlying an abnormal cardiac valve, e.g. systolic thrill with aortic stenosis.

Auscultation

The *bell* of the stethoscope is used for low-pitched sounds (heart sounds and mid-diastolic murmur in mitral stenosis).

The *diaphragm* is used for high-pitched sounds (systolic murmurs, aortic regurgitation, ejection clicks and opening snaps).

First heart sound (S1)

This is due to mitral and tricuspid valve closure. A loud S1 occurs in thin people, hyperdynamic circulation, tachycardias and mild–moderate mitral stenosis. A soft S1 occurs in obesity, emphysema, pericardial effusion, severe calcific mitral stenosis, mitral or tricuspid regurgitation, heart failure, shock, bradycardias and first-degree block.

Second heart sound (S2)

This is due to aortic and pulmonary valve closure. Physiological splitting of S2 occurs during inspiration in children and young adults.

Third and fourth heart sounds

These are pathological.

- A *third heart sound* is due to rapid ventricular filling and is present in heart failure.
- A *fourth heart sound* occurs in late diastole and is associated with atrial contraction.

Either, singly or together, will produce a gallop rhythm.

Heart murmurs

These are due to turbulent blood flow and occur in hyperdynamic states or with abnormal valves. Innocent or flow murmurs are soft, early systolic, short and non-radiating. They are heard frequently in children and young adults. The individual murmurs are discussed under valve disease, p. 760.

- aortic area is the 2nd intercostal space right sternal edge
- pulmonary area is the 2nd intercostal space left sternal edge
- tricuspid area is the 3rd–5th intercostal space left sternal edge
- mitral area is the 5th intercostals space mid-clavicular line.

CARDIAC INVESTIGATIONS

Chest X-ray

Ideally, this is taken in the postero-anterior (PA) direction at maximum inspiration with the heart close to the X-ray film to minimize magnification with respect to the thorax. A lateral may give additional information if the PA is abnormal. The cardiac structures and great vessels that can be seen on these X-rays are indicated in Figure 13.13. An antero-posterior (AP) view is only taken in an emergency.

Heart size

Heart size can be reliably assessed only from the PA chest film. The maximum transverse diameter of the heart is compared with the maximum transverse diameter of the thorax measured from the inside of the ribs (the cardiothoracic ratio).

The cardiothoracic ratio (CTR) is usually less than 50%, except in neonates, infants, athletes and patients with skeletal abnormalities such as scoliosis and funnel chest. A transverse cardiac diameter of more than 15.5 cm is abnormal. Pericardial effusion or cardiac dilatation causes an increase in the ratio.

A *pericardial effusion* produces a globular, sharp-edged shadow (Fig. 13.120, p. 796). This enlargement may occur quite suddenly and, unlike in heart failure, there is no associated change in the pulmonary vasculature. The echocardiogram is more specific (p. 704).

Certain patterns of specific chamber enlargement may be seen on the chest X-ray:

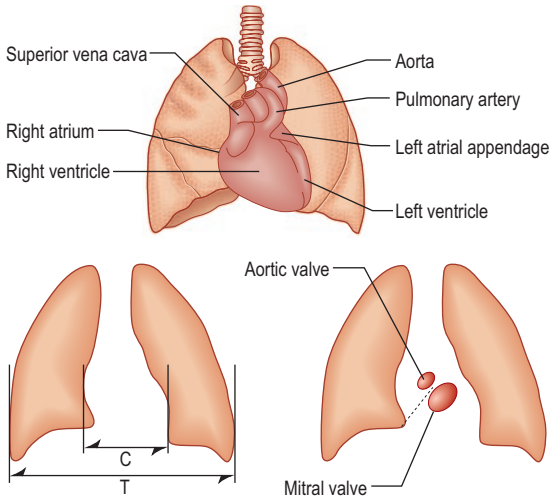


Fig. 13.13 **Diagrams to show the heart silhouette on the chest X-ray,** measurements of the cardiothoracic ratio (CTR) and the location of the cardiac valves. The dotted line is an arbitrary line from the left hilum to the right cardiophrenic angle: a calcified aortic valve is seen above this line. CTR = (C/T) × 100%; normal CTR <50%.

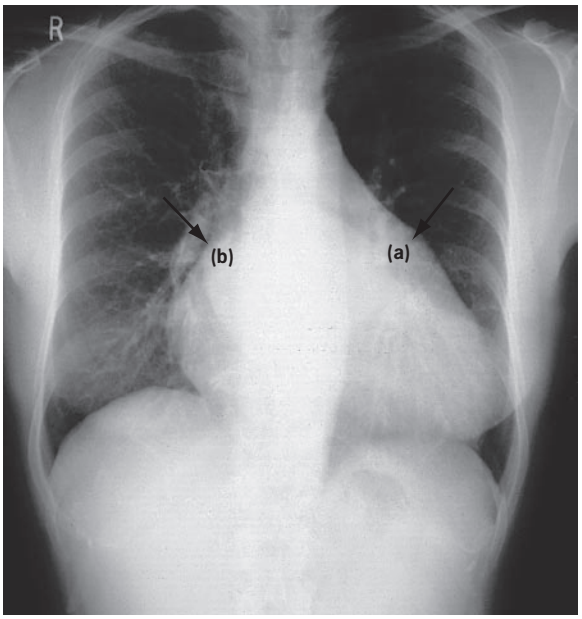


Fig. 13.14 **Plain PA chest X-ray of a patient with mixed mitral valve disease.** The left atrium is markedly enlarged (arrow a). Note the large bulge on the left heart border (left atrium) and the 'double shadow' (border of the right and left atria) (arrow b) on the right side of the heart. There is cardiac (left ventricular) enlargement due to mitral regurgitation.

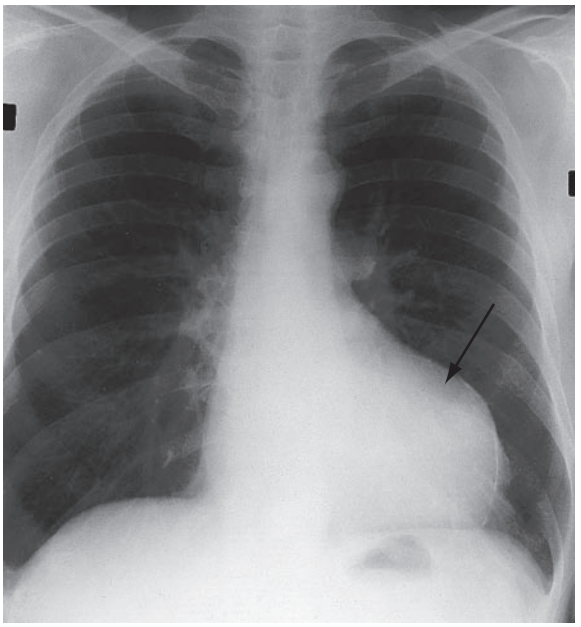


Fig. 13.15 **Left ventricular aneurysm: plain PA chest X-ray demonstrating a cardiac silhouette with a 'bulge' (arrow) on the left lateral border.** This bulge is due to aneurysm formation of many years following a myocardial infarction. A thin line of calcification can be seen along the edge of this bulge.

Left atrial dilatation. This results in prominence of the left atrial appendage and a straightening or convex bulging of the upper left heart border, a double atrial shadow to the right of the sternum, and splaying of the carina because a large left atrium elevates the left main bronchus (Fig. 13.14). On a lateral chest X-ray an enlarged left atrium bulges backwards, impinging on the oesophagus.

Left ventricular enlargement. This results in an increase in the CTR and a smooth elongation and increased convexity of the left heart border. A left ventricular aneurysm may produce a distinct bulge or distortion of the left heart border (Fig. 13.15).

Right atrial enlargement. This results in the right border of the heart projecting into the right lower lung field.

Right ventricular enlargement. This results in an increase of the CTR and an upward displacement of the apex of the heart because the enlarging right ventricle pushes the left ventricle leftwards, upwards and eventually backwards. Differentiation of left from right ventricular enlargement may be difficult from the shape of the left heart border alone, but the lateral view shows enlargement anteriorly for the right ventricle and posteriorly for the left ventricle.

Ascending aortic dilatation or enlargement. This is seen as a prominence of the aortic shadow to the right of the mediastinum between the right atrium and superior vena cava.

Dissection of the ascending aorta. This is seen as a widening of the mediastinum on chest X-ray but an ultrasound/magnetic resonance imaging (MRI) should be performed.

Enlargement of the pulmonary artery. Enlargement of the pulmonary artery in pulmonary hypertension, pulmonary artery stenosis and left-to-right shunts produces a prominent bulge on the left-hand border of the mediastinum below the aortic knuckle.

Calcification

Calcification in the cardiovascular system occurs because of tissue degeneration. Calcification is visible on a lateral or a penetrated PA film, but is best studied with computed tomography (CT) scanning. Various types of calcification can occur.

Pericardial calcification is plaque-like opacities over the surface of the heart (particularly concentrated in the atrioventricular groove) often result from tuberculous pericarditis and can cause pericardial constriction (see Fig. 13.121, p. 797).

Valvular calcification results from long-standing rheumatic or bicuspid aortic valve disease.

Myocardial calcification occurs after myocardial infarction, especially in association with a left ventricular aneurysm (Fig. 13.15).

Calcification of the aorta is a common, normal finding in patients over the age of 40 years and appears as a curvilinear opacity around the circumference of the aortic knuckle. Calcification in the ascending aorta usually denotes syphilitic aortitis, whereas in the descending aorta it is due to atheroma or, in the younger patient, to non-specific aortitis.

Coronary arterial calcification, especially of the proximal left coronary artery, is associated with coronary atheroma and is best seen on CT with multidetector CT (MDCT) scanners (see p. 705).

Lung fields

Pulmonary plethora results from left-to-right shunts (e.g. atrial or ventricular septal defects). It is seen as a general increase in the vascularity of the lung fields and as an

increase in the size of hilar vessels (e.g. in the right lower lobe artery), which normally should not exceed 16 mm diameter.

Pulmonary oligaemia is a paucity of vascular markings and a reduction in the width of the arteries. It occurs in situations where there is reduced pulmonary blood flow, such as pulmonary embolism, severe pulmonary stenosis and Fallot's tetralogy.

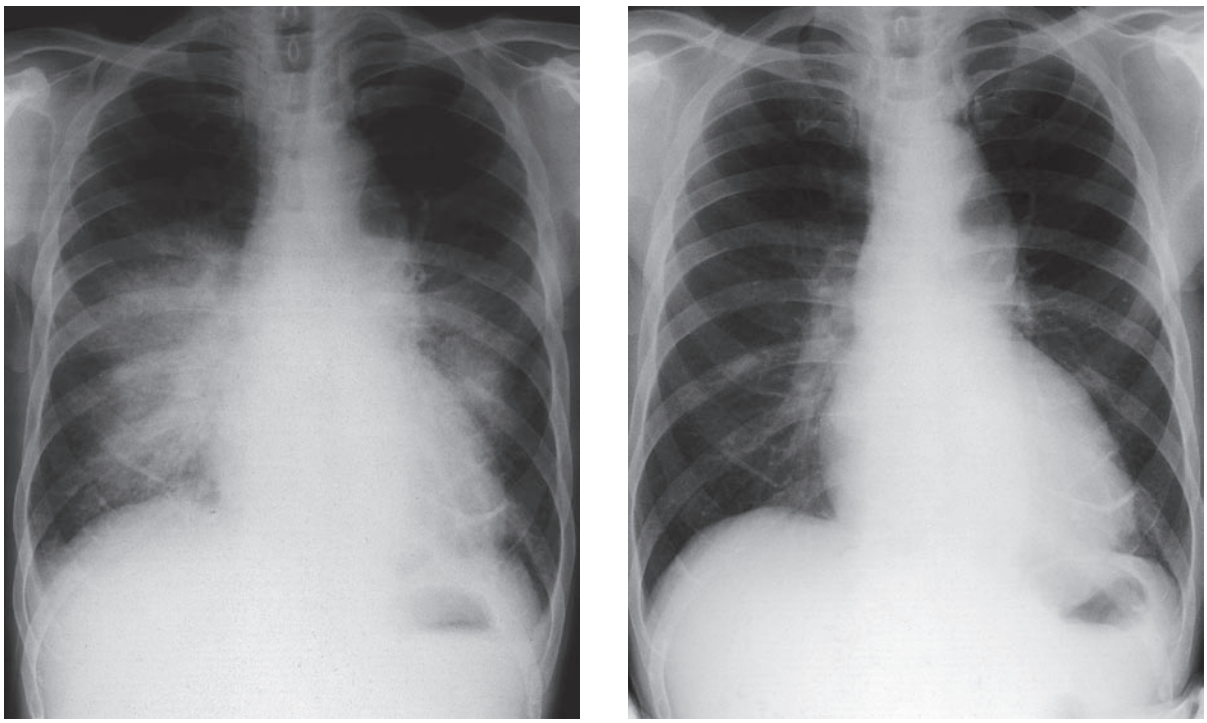
Pulmonary arterial hypertension may result from pulmonary embolism, chronic lung disease or chronic left heart disease, such as shunts due to a ventricular septal defect or mitral valve stenosis. In addition to X-ray features of these conditions, the pulmonary arteries are prominent close to the hila but are reduced in size (pruned) in the peripheral lung fields. This pattern is usually symmetrical.

Pulmonary venous hypertension occurs in left ventricular failure or mitral valve disease. Normal pulmonary venous pressure is 5–14 mmHg at rest. Mild pulmonary venous hypertension (15–20 mmHg) produces isolated dilatation of the upper zone vessels. Interstitial oedema occurs when the pressure is between 21 and 30 mmHg. This manifests as fluid collections in the interlobar fissures, interlobular septa (Kerley B lines) and pleural spaces. This gives rise to indistinctness of the hilar regions and haziness of the lung fields. Alveolar oedema occurs when the pressure exceeds 30 mmHg, appearing as areas of consolidation and mottling of the lung fields (Fig. 13.16) and pleural effusions. Patients with long-standing elevation of the pulmonary venous pressure have reactive thickening of the pulmonary arteriolar intima, which protects the alveoli from pulmonary oedema. Thus, in these patients the pulmonary venous pressure may increase to well above 30 mmHg before frank pulmonary oedema develops.

Electrocardiography

The electrocardiogram (ECG) is a recording of the electrical activity of the heart. It is the vector sum of the depolarization and repolarization potentials of all myocardial cells (see Fig. 13.2, p. 682). At the body surface these generate potential differences of about 1 mV, and the fluctuations of these potentials create the familiar P-QRS-T pattern. At rest the intracellular voltage of the myocardium is polarized at –90 mV compared with that of the extracellular space. This diastolic voltage difference occurs because of the high intracellular potassium concentration, which is maintained by the sodium potassium pump despite the free membrane permeability to potassium. Depolarization of cardiac cells occurs when there is a sudden increase in the permeability of the membrane to sodium. Sodium rushes into the cell and the negative resting voltage is lost (phase 0 in Fig. 13.42, p. 716). The depolarization of a myocardial cell causes the depolarization of adjacent cells and, in the healthy heart, the entire myocardium is depolarized in a coordinated fashion. During repolarization, cellular electrolyte balance is slowly restored (phases 1, 2 and 3). Slow diastolic depolarization (phase 4) follows until the threshold potential is reached. Another action potential then follows.

The ECG is recorded from two or more simultaneous points of skin contact (electrodes). When cardiac activation proceeds towards the positive contact, an upward deflection is produced on the ECG. Correct representation of a three-



(a) (b)

Fig. 13.16 **Acute pulmonary oedema.** This pair of chest X-rays were taken from a patient before **(a)** and after **(b)** treatment of acute pulmonary oedema. The X-ray taken when the oedema was present demonstrates hilar haziness, Kerley B lines, upper lobe venous engorgement and fluid in the right horizontal interlobar fissure. These abnormalities are resolved on the film taken after successful treatment.

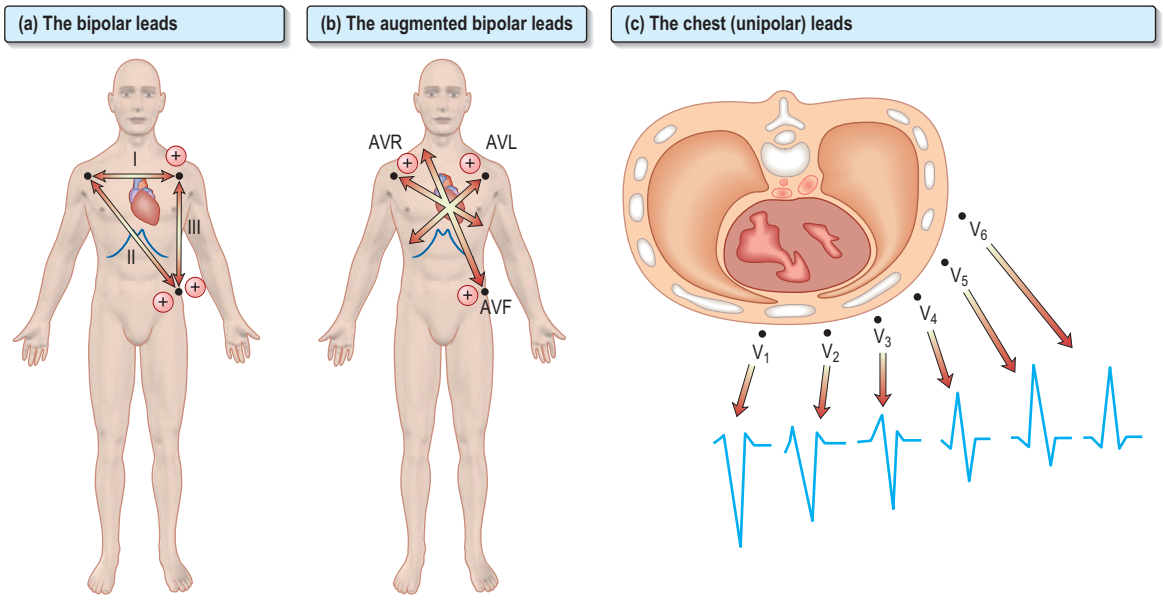


Fig. 13.17 **The connections or directions that comprise the 12-lead electrocardiogram.**

dimensional spatial vector requires recordings from three mutually perpendicular (orthogonal) axes. The shape of the human torso does not make this easy, so the practical ECG records 12 projections of the vector, called 'leads' (Fig. 13.17).

Six of the leads are obtained by recording voltages from the limbs (I, II, III, AVR, AVL and AVF). The other six leads record potentials between points on the chest surface and an average of the three limbs: RA, LA and LL. These are designated V_1–V_6 and aim to select activity from the right ventricle (V_1–V_2), interventricular septum (V_3–V_4) and left ventricle (V_5–V_6). Note that leads AVR and V_1 are oriented towards the cavity of the heart, leads II, III and AVF face the inferior surface, and leads I, AVL and V_6 face the lateral wall of the left ventricle. A V_4 on the right side of the chest (V_4R) is occasionally useful (e.g. for the diagnosis of right ventricular infarction).

Most ECG machines are simultaneous three-channel recorders with output either as a continuous strip or with automatic channel switching. Many ECG machines also analyse the recordings and print the analysis on the record. Usually the machine interpretation is correct, but many arrhythmias still defy automatic analysis.

The ECG waveform

The shape of the normal ECG waveform (Fig. 13.18) has similarities, whatever the orientation. The first deflection is caused by atrial depolarization, and it is a low-amplitude slow deflection called a *P wave*. The *QRS complex* reflects ventricular activation or depolarization and is sharper and larger in amplitude than the P wave. An initial downward deflection is called the *Q wave*. An initial upward deflection is called an *R wave*. The *S wave* is the last part of ventricular activation. The *T wave* is another slow and low-amplitude deflection that results from ventricular repolarization.

The *PR interval* is the length of time from the start of the P wave to the start of the QRS complex. It is the time taken for activation to pass from the sinus node, through the atrium, AV node and His–Purkinje system to the ventricle.

The *QT interval* extends from the start of the QRS complex to the end of the T wave. This interval represents the time

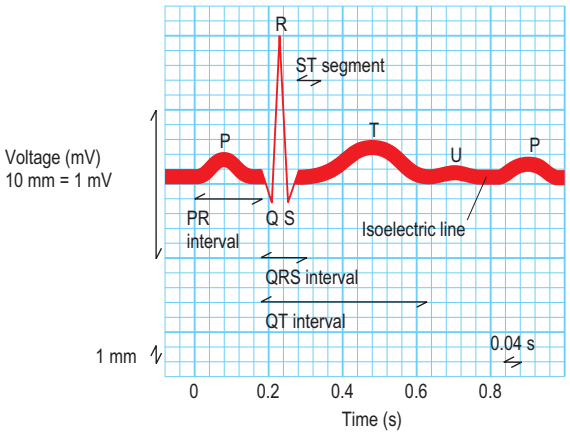


Fig. 13.18 **The waves and elaboration of the normal electrocardiogram.** From Goldman MJ. *Principles of Clinical Electrocardiography*, 9th edn. Los Altos: Lange, 1976.

Table 13.5 Normal ECG intervals

P wave duration	≤0.12 s
PR interval	0.12–0.22 s
QRS complex duration	≤0.10 s
Corrected QT (QT_c)	≤0.44 s in males
	≤0.46 s in female
$QT_{cB} = QT/\sqrt{(R-R)}$	Bazett's square root formula
$QT_{cF} = QT/\sqrt[3]{(R-R)}$	Fridericia's cube root formula

taken to depolarize and repolarize the ventricular myocardium. The QT interval varies greatly with heart rate and is often represented as a corrected QT interval (or QTc) for a given heart rate. There are a number of formulae for derivation of QTc, but the most widely accepted are Bazett's formula and Fridericia's correction (Table 13.5).

An abnormally prolonged QTc can predispose to a risk of dangerous ventricular arrhythmias. Prolongation of QT

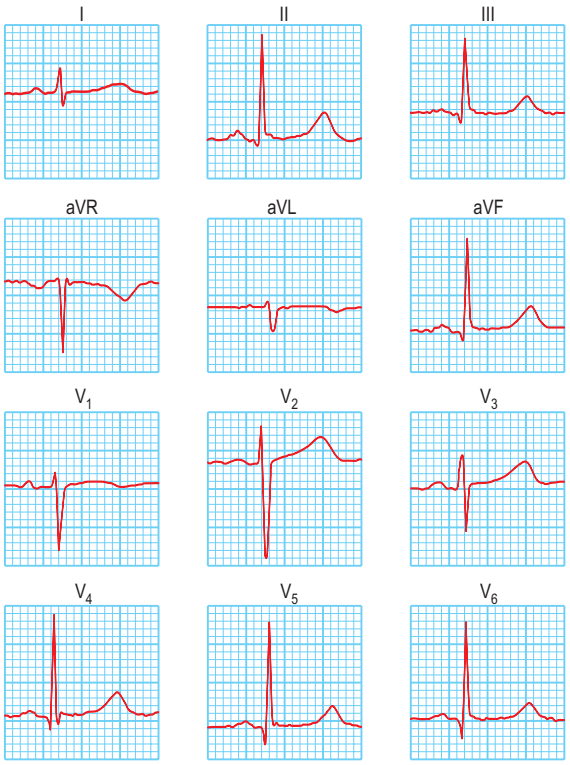


Fig. 13.19 **A normal 12-lead electrocardiogram.**

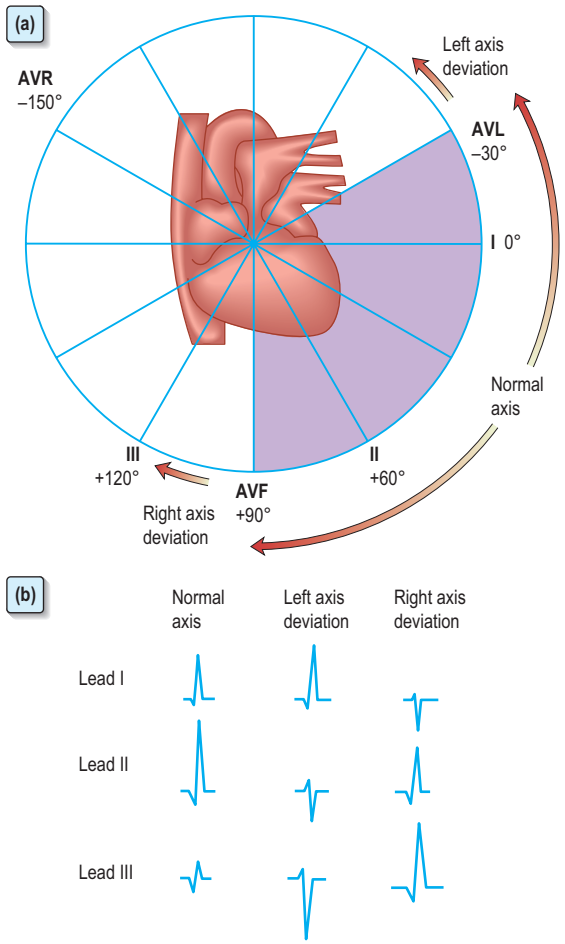


Fig. 13.20 **Cardiac vectors. (a) The hexaxial reference system,** illustrating the six leads in the frontal plane, e.g. lead I is 0°, lead II is +60°, lead III is +120°. **(b) Calculating the direction of the cardiac vector.** In the first column the QRS complex with zero net amplitude (i.e. when the positive and negative deflections are equal) is seen in lead III. The mean QRS vector is therefore perpendicular to lead III and is either –150° or +30°. Lead I is positive, so the axis must be +30°, which is normal. In left axis deviation (second column) the main deflection is positive (R wave) in lead I and negative (S wave) in lead III. In right axis deviation (third column) the main deflection is negative (S wave) in lead I and positive (R wave) in lead III. The frontal plane QRS axis is normal only if the QRS complexes in leads I and II are predominantly positive.

interval may be congenital or can occur in many acquired conditions (see Table 13.14, p. 727).

The *ST segment* is the period between the end of the QRS complex and the start of the T wave. In the normal heart, all cells are depolarized by this phase of the ECG, i.e. the ST segment represents ventricular repolarization.

A normal ECG is shown in Figure 13.19, and the normal values for the electrocardiographic intervals are indicated in Table 13.5. Leads that face the lateral wall of the left ventricle have predominantly positive deflections, and leads looking into the ventricular cavity are usually negative. Detailed patterns depend on the size, shape and rhythm of the heart and the characteristics of the torso.

Cardiac vectors

At any point in time during depolarization and repolarization, electrical potentials are being propagated in different directions. Most of these cancel each other out and only the net force is recorded. This net force in the frontal plane is known as the cardiac vector.

The mean QRS vector can be calculated from the six standard leads (Fig. 13.20):

- normal between –30° and +90°
- left axis deviation between –30° and –90°
- right axis deviation between +90° and +150°.

Calculation of this vector is useful in the diagnosis of some cardiac disorders.

Exercise electrocardiography

The ECG is recorded whilst the patient walks or runs on a motorized treadmill. The test is based upon the principle that exercise increases myocardial demand on coronary blood supply, which may be inadequate during exercise, and at peak stress can result in relative myocardial ischaemia. Most exercise tests are performed according to a standardized method, e.g. the Bruce protocol. Recording an ECG after exercise is not an adequate form of stress test. Normally there is little change in the T wave or ST segment during exercise.

The patient's exercise capacity (the total time achieved) will depend on many factors; however, patients who can only exercise for less than 6 minutes generally have a poorer prognosis.

Myocardial ischaemia provoked by exertion results in ST segment depression (>1 mm) in leads facing the affected area. The form of ST segment depression provoked by ischaemia is characteristic: it is either planar or shows down-sloping depression (Fig. 13.21). Up-sloping depression is a non-specific finding. The degree of ST segment

Fig. 13.21 **Electrocardiographic changes during exercise test.** Upper trace – significant horizontal ST segment depression during exercise. Lower trace – ST elevation during exercise.

depression is positively correlated to the degree of myocardial ischaemia.

ST segment elevation during an exercise test is induced much less frequently than ST depression. When it occurs, it reflects transmural ischaemia caused by coronary spasm or critical stenosis.

Although most abnormalities are detected in leads V_5 (anterior and lateral ischaemia) or AVF (inferior ischaemia), it is best to record a full 12-lead ECG. During an exercise test the blood pressure and rhythm responses to exercise are also assessed. Exercise normally causes an increase in heart rate and blood pressure. A sustained fall in blood pressure usually indicates severe coronary artery disease. A slow recovery of the heart rate to basal levels has also been reported to be a predictor of mortality.

Frequent premature ventricular depolarizations during the test are associated with a long-term increase in the risk of death from cardiovascular causes and further testing is required in these patients. Use of the exercise test in angina is described on page 749.

24-Hour ambulatory taped electrocardiography

This records transient changes such as a brief paroxysm of tachycardia, an occasional pause in the rhythm, or intermittent ST segment shifts (Fig. 13.22). A conventional 12-lead ECG is recorded in less than a minute and usually samples less than 20 complexes. In a 24-hour period over 100 000 complexes are recorded. Such a large amount of data must be analysed by automatic or semi-automatic methods. This technique is called 'Holter' electrocardiography after its inventor.

Event recording is another technique that is used to record less frequent arrhythmias. The patient is provided with a pocket-sized device that can record and store a short segment of the ECG. The device may be kept for several days or weeks until the arrhythmia is recorded. Most units of this kind will also allow transtelephonic ECG transmission so that the physician can determine the need for treatment or the continued need for monitoring.

A very small event recorder, known as an implantable loop recorder (ILR), can also be implanted subcutaneously, triggered by events or a magnet, and interrogated by the physician.

Other tests

Non-invasive methods that make use of digitalized Holter recordings to identify increased risk of ventricular arrhythmias include assessment of heart rate variability, signal averaged ECG (SAECG) and T wave alternans.

Heart rate variability (HRV) can be assessed from a 24-hour ECG. HRV is decreased in some patients following myocardial infarction and represents an abnormality of autonomic tone or cardiac responsiveness. Low HRV is a major risk factor for sudden death and ventricular arrhythmias in patients discharged from hospital following myocardial infarction.

Signal-averaged ECG (SAECG) is a technique that requires amplification and averaging of abnormal low-amplitude signals which occur beyond the end of the QRS complex and extend well into the ST segment. These signals are therefore also known as late potentials and are too small to be detected on a surface ECG. They arise in areas of slow conduction in the myocardium, such as the border zone of an infarct, where re-entrant ventricular arrhythmias can originate.

T wave alternans (TWA) is a valuable technique used as a non-invasive marker of susceptibility to ventricular arrhythmias and sudden cardiac death. TWA represents microvolt level changes in the morphology of the T waves in every other beat and can be detected during acute myocardial ischaemia using amplification techniques. Visible TWA on an ordinary surface ECG is quite a rare phenomenon, except in patients with long QT syndromes, particularly during emotion or exercise.

Tilt testing

Patients with suspected neurocardiogenic (vasovagal) syncope should be investigated by upright tilt testing. The patient is secured to a table which is tilted to + 60° to the vertical for 45 minutes or more. The ECG and blood pressure are monitored throughout. If neither symptoms nor signs develop, isoprenaline may be slowly infused or glyceryl trinitrate inhaled and the tilt repeated. A positive test results in hypotension, sometimes bradycardia (Fig. 13.23) and pre-syncope/syncope, and supports the diagnosis of neurocardiogenic syncope. If symptoms and signs appear, placing the patient flat can quickly reverse them. The effect of treatment can be evaluated by repeating the tilt test, but it is not always reproducible. The overall sensitivity is low.

Carotid sinus massage

Carotid sinus massage (see p. 698, Practical Box 13.3) may lead to asystole (>3 s) and/or a fall of systolic blood pressure (>50 mmHg). This hypersensitive response occurs in many of the normal (especially elderly) population, but may also be responsible for loss of consciousness in some patients with carotid sinus syndrome (p. 689). In one-third of cases carotid sinus massage is only positive when the patient is standing. Atherosclerosis can cause narrowing and stenosis of carotid arteries. Carotid sinus massage should thus be avoided in patients with carotid bruits.

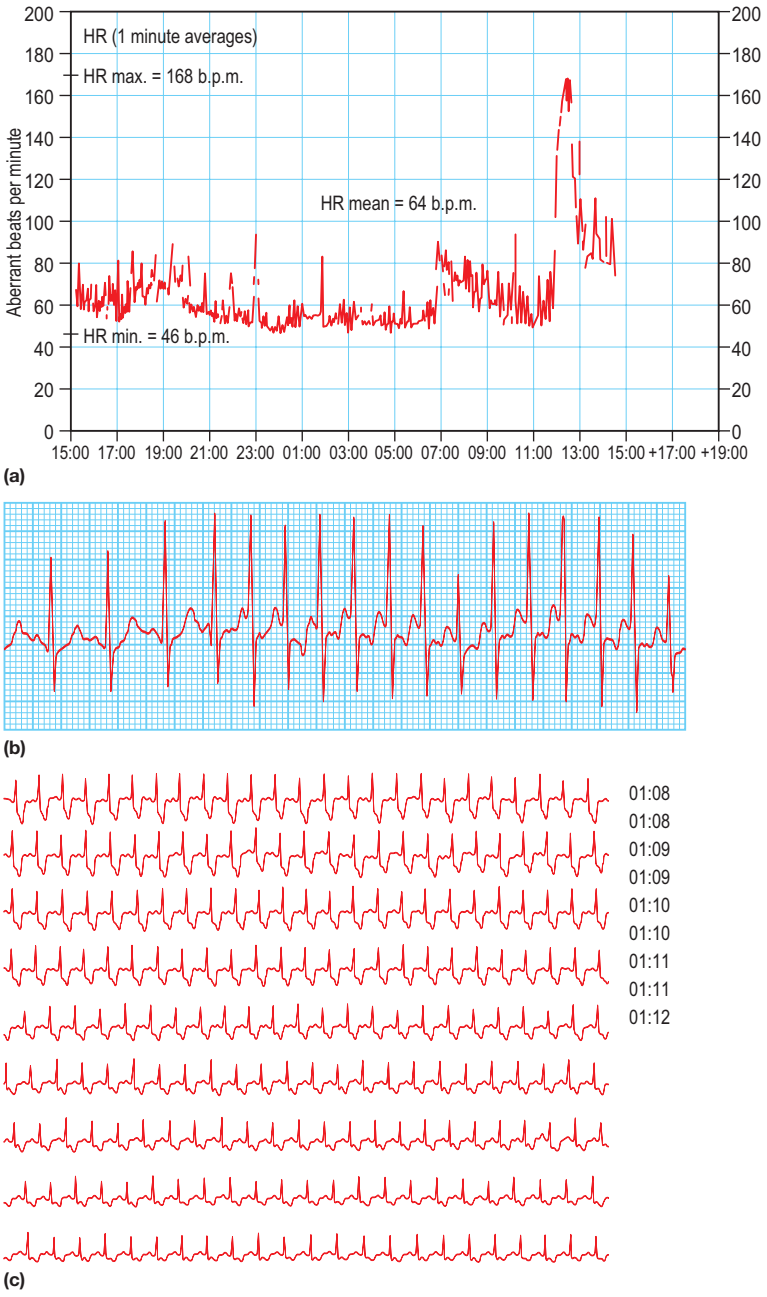


Fig. 13.22 **Examples of Holter recordings.** (a) Tachograph showing sudden increase in heart rate between 1100 and 1300 hours. (b) Ambulatory ECG record of the same patient, revealing supraventricular tachycardia. (c) Ambulatory ECG showing ST depression in a patient with silent myocardial ischaemia.

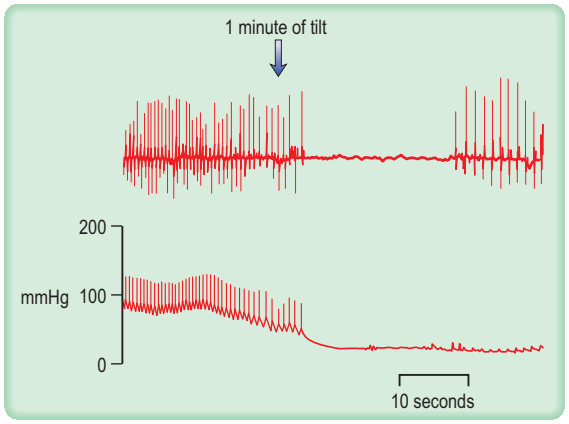


Fig. 13.23 **Tilt test.** The ECG (top trace) and arterial blood pressure (bottom trace) recorded during a tilt test. After 1 minute of tilt, hypotension, bradycardia and syncope occur.

Practical Box 13.3 Carotid sinus massage

- Ensure there is no significant carotid artery disease (carotid bruits).
- Provide continuous electrocardiographic monitoring.
- Patient is in supine position with the head slightly extended.
- Start with right carotid sinus massage.
- Apply firm rotary pressure to the carotid artery at the level of the third cervical vertebra for 5 seconds.
- Alternatively, steady pressure can be applied.
- If no response, massage left carotid sinus.
- Generally, right carotid sinus massage decreases the sinus node discharge, and left carotid sinus massage slows atrioventricular conduction.
- Do not massage both carotid sinuses at the same time.
- Single application of carotid sinus pressure may be effective in about 20–30% of patients with paroxysmal supraventricular tachycardias; multiple applications can terminate tachycardia in about 50% of patients.
- Asystole is a potential but rare complication.

Echocardiography

Echocardiography is a non-invasive diagnostic technique that is widely used in clinical cardiology. It involves the use of ultrasound (either alone or with contrast agent) to assess cardiac structure and function.

Physics

Echocardiography uses transmitted ultrasound wavelengths of 1 mm or less, which correspond to frequencies of approximately 2 MHz (2 million cycles per second) or more. At such high frequencies, the ultrasound waves can be focused into a 'beam' and aimed at a particular region of the heart. The waves are generated in very short bursts or pulses a few microseconds long by a crystal transducer, which also detects returning echoes and converts them into electrical signals.

When the handheld crystal transducer is placed on the body surface the emitted ultrasound pulses encounter interfaces between various body tissues as they pass through the body. In crossing each interface, some of the wave energy is reflected, and if the beam path is approximately at right angles to the plane of the interface, the reflected waves return to the transducer as an echo. Since the velocity of sound in body tissues is almost constant (1550 m/s), the time delay for the echo to return measures the distance of the reflecting interface. Thus, if a single ultrasound pulse is transmitted, a series of echoes return, the first of which is from the closest interface.

Echocardiographic modalities

M mode

M-mode echocardiography is a technique that details the changing motion of structures along the ultrasound beam with time. Thus the motion of the interventricular septum during the cardiac cycle (either towards or away from the transducer placed on the chest wall) can be assessed and quantified. Stationary structures thus generate horizontal straight lines, the distances of which from the top of the screen indicate their depths, and movements, such as those of heart valves, are indicated by zigzag lines (Fig. 13.24c).

Two-dimensional echocardiography

Alternatively, a series of views from different positions can be obtained in the form of a two-dimensional image (cross-sectional 2-D echocardiography) (Fig. 13.24a, b, d). This method is useful for delineating anatomical structures and for quantifying volumes of cardiac chambers. It is conventional for different imaging planes to be used to create standard 2-D views.

Three-dimensional echocardiography

Three-dimensional echocardiography is a novel development in cardiac imaging in which a volumetric dataset is acquired using a multiplane probe rotating around a fixed axis. Current clinical uses include accurate volumetric assessment of ventricular function and mass, assessment of mitral and aortic valve disease and assessment of adult congenital heart disease (Fig. 13.25).

Doppler echocardiography

Echocardiography imaging utilizes echoes from tissue interfaces. Using high amplification, it is also possible to detect weak echoes scattered by small targets, including those from red blood cells. Blood velocity in the heart chambers is typically much more rapid (>1 m/s) than the movement of myocardial tissue. If the blood is moving in the same direction as the direction of the ultrasound beam, the frequency of the returning echoes will be changed according to the Doppler phenomenon. The Doppler shift frequency is directly proportional to the blood velocity. Blood velocity data can be acquired and displayed in several ways.

Pulsed-wave (PW) Doppler extracts velocity data from the pulse echoes used to form a 2-D image and gives useful qualitative information. PW echoes can be specified from locations within an image identified by a sample volume cursor placed on the screen. Such information from the left ventricular outflow tract (LVOT) and right ventricular outflow tract provides the stroke distance, and is used to estimate cardiac output (CO) and also to quantify intracardiac shunts.

Cardiac output can then be derived using the formula: CO = stroke volume × heart rate. Stroke volume is the stroke distance multiplied by the area of the LVOT, which can also be measured echocardiographically. PW Doppler of the flow across the mitral valve and into the left atrium through the pulmonary veins can be used as part of the estimation of left ventricular filling pressure.

Colour flow Doppler. Doppler colour flow imaging uses one colour for blood flowing towards the transducer and another colour for blood flowing away. This technique allows the direction, velocity and timing of the flow to be measured with a simultaneous view of cardiac structure and function. Colour flow Doppler is used to help assess valvular regurgitation (Fig. 13.26) and may be useful in the assessment of coronary blood flow.

Continuous-wave (CW) Doppler collects all the velocity data from the path of the beam and analyses it to generate a spectral display. This is unlike PW Doppler, which provides information from a particular sample volume at one location along a line. Thus, CW Doppler does not provide any depth information.

The outline of the envelope of the spectral display is used to estimate the value of peak velocity throughout the cardiac cycle. CW Doppler is used typically to assess valvular obstruction, which then causes increased velocities. For example, normal flow velocities are of the order of 1 m/s across the normal aortic valve, but if there is a severe obstructive lesion, such as a severely stenotic aortic valve, velocities of 4 m/s or more can occur (Fig. 13.27a, b). These velocities are generated by the pressure gradient that exists across the lesion.

According to the Bernoulli equation, the pressure difference between two chambers is calculated as: 4 multiplied by the square of the CW Doppler velocity between chambers. Thus a velocity of 5 m/s across the aortic valve suggests a peak gradient of $4 \times 5 \times 5 = 100$ mmHg between the ascending aorta and the left ventricle. This equation has been validated in a wide variety of clinical situations, including valve stenoses, ventricular septal defects and intraventricular obstruction (as in hypertrophic cardiomyopathy). It is often

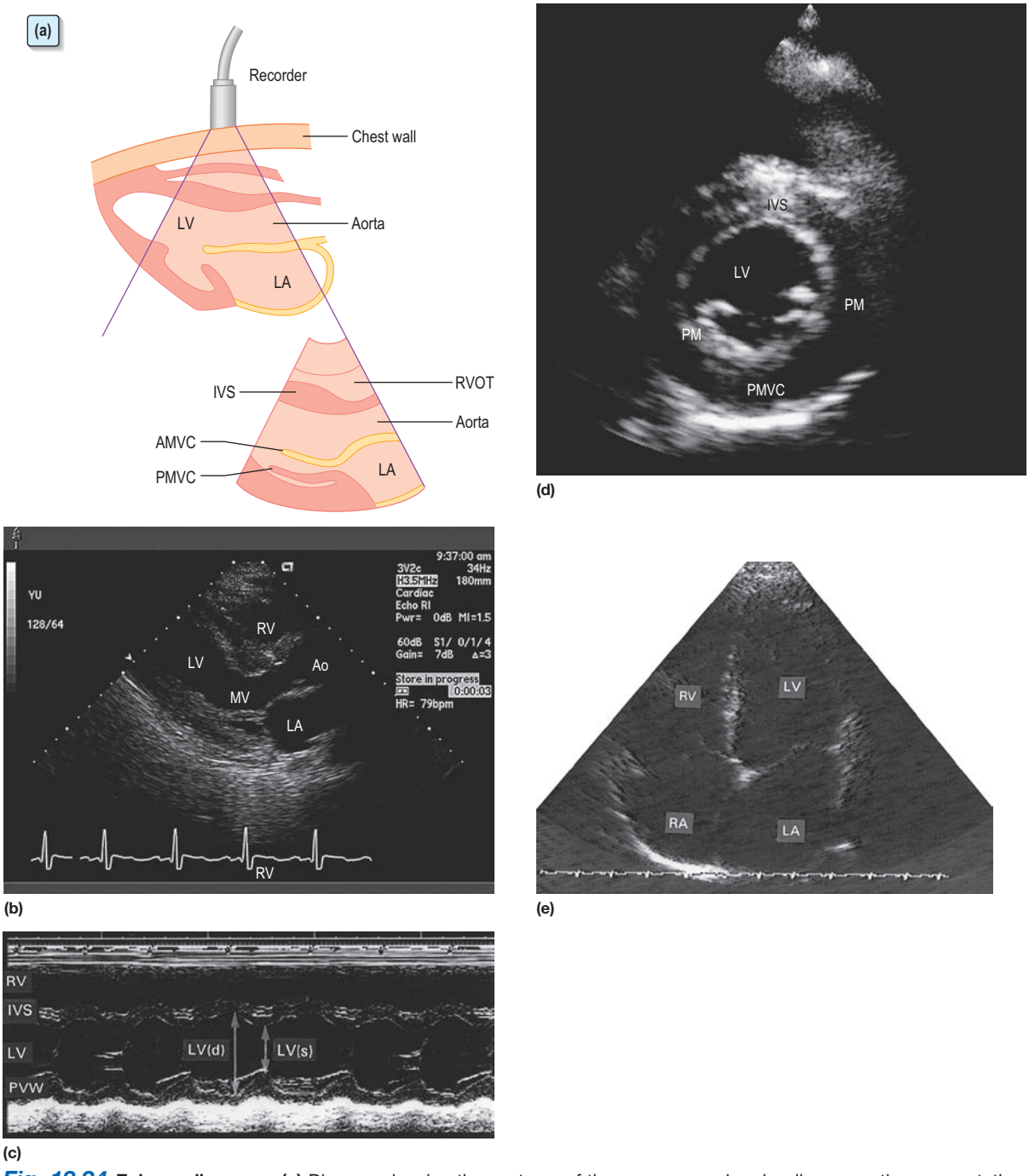


Fig. 13.24 **Echocardiograms. (a)** Diagram showing the anatomy of the area scanned and a diagrammatic representation of the echocardiogram. **(b)–(e) Echocardiograms from a normal subject: (b)** *Two-dimensional long-axis view*. **(c)** *M-mode recording* with the ultrasound beam directed across the left ventricle, just below the mitral valve. **(d)** *Two-dimensional short-axis view* at the level of the tips of the papillary muscles. **(e)** *Apical four-chamber view.* Note that the convention that shows the position of the transducer (the apex of the sector image) at the top of the paper causes the heart to appear 'upside down' in these views. AMVC, anterior mitral valve cusp; Ao, aorta; IVS, interventricular septum; LA, left atrium; LV, left ventricle; LV(d), LV(s), left ventricular end-diastolic and end-systolic dimensions; MV, mitral valve; PM, papillary muscle; PMVC, posterior mitral valve cusp; PVW, posterior ventricular wall; RA, right atrium; RV, right ventricle; RVOT, right ventricular outflow tract.

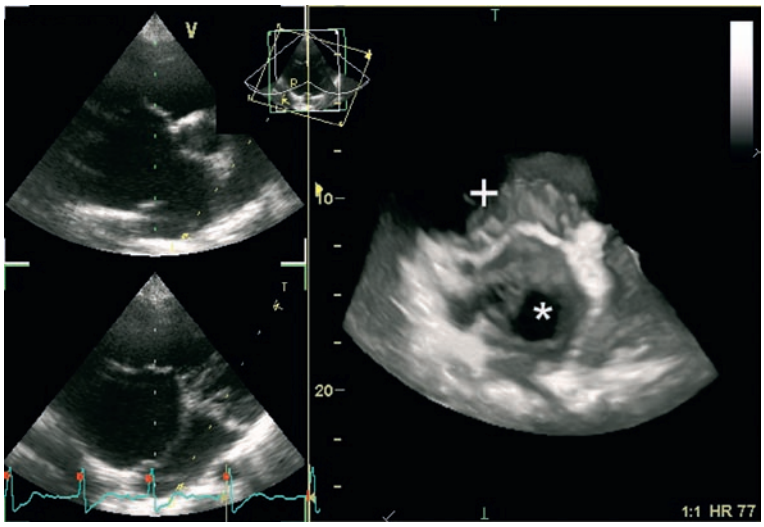


Fig. 13.25 **3-D Echocardiograph in a patient with an atrial septal defect (*).** The mitral valve (+) and the left atrial anatomy are well delineated. (AVI available online: http://www.asecho.org/freepdf/3D.pdf.)

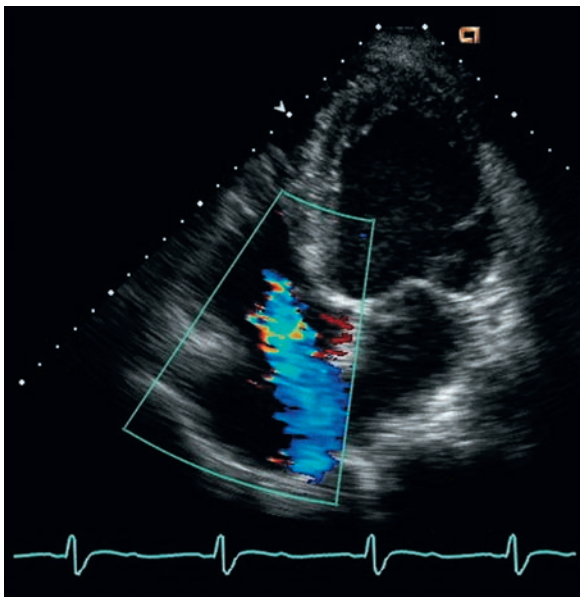


Fig. 13.26 **Colour Doppler** shows blood flowing away from the echocardiography probe as a blue signal and towards the probe as a red signal. In this patient with tricuspid regurgitation, blood leaks from the right ventrical to the right atrium during cardiac systole.

clinically unnecessary to resort to invasive methods such as cardiac catheterization to measure intracardiac pressure gradients.

Similarly, pulmonary artery (PA) systolic pressure and right ventricular diastolic pressure can be calculated using the Bernoulli equation. In this case, CW Doppler tracing of the tricuspid regurgitant jet is used to estimate the pressure gradient between the right ventricle and the right atrium. The PA systolic pressure is then calculated by adding the estimated right atrial pressure to the pressure gradient between the right ventricle and the right atrium.

Tissue Doppler is similar to PW Doppler. It measures myocardial tissue velocities within a particular sample volume placed on the image. Such velocities are of the order of 1 cm/s. Currently, tissue Doppler of the mitral annulus is used as part of the estimation of left ventricular filling pressure.

Other ultrasound modalities. Harmonic power Doppler, pulse inversion Doppler and ultraharmonics are used to detect and amplify microsphere-specific signals as part of the echocardiographic assessment of myocardial perfusion.

The echocardiographic examination

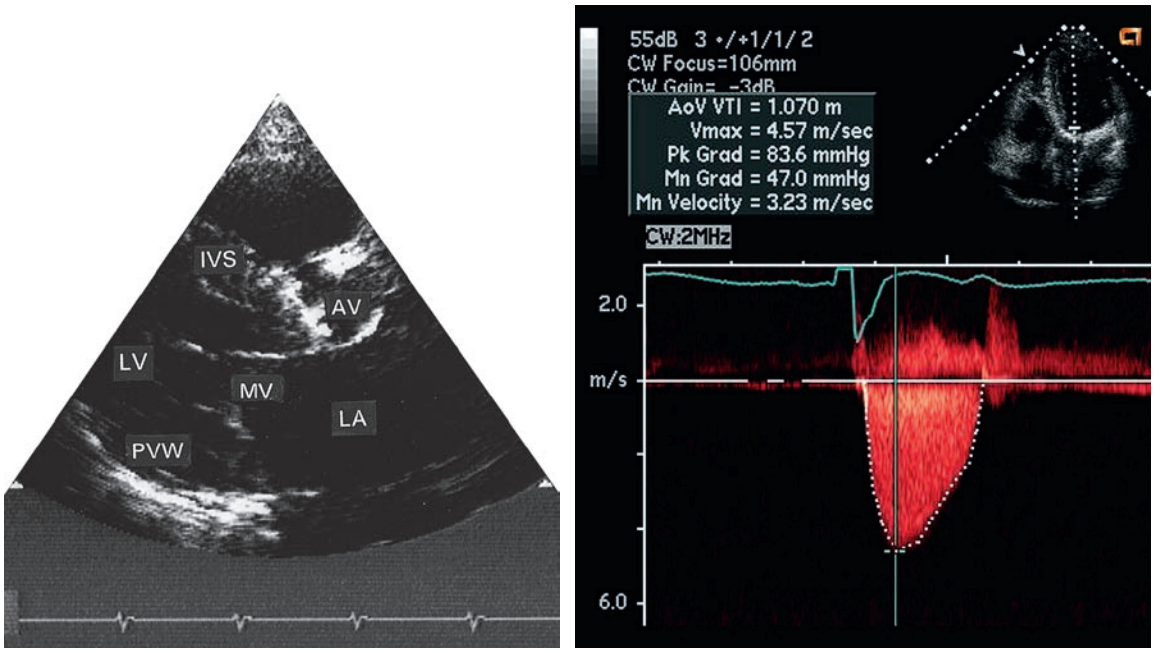
Transthoracic echocardiography is a 'non-invasive' procedure that causes the patient no discomfort and is harmless. A physician or technician performs the studies and a comprehensive examination takes 15–45 minutes. The ultrasound machines are either mobile on wheels, or are handheld.

The technical issue associated with echocardiography is that the lungs and rib cage obstruct the passage of ultrasound to the heart in the adult subject. Small 'windows' can usually be found in the third and fourth left intercostal spaces (termed left parasternal); just below the xiphoid process of the sternum (subcostal); and, with the subject turned to the left and exhaling, from the point where the apical beat is palpated (apical). By positioning the transducer successively over these sites and angling and rotating it to line up with the scan plane, a series of standard sectional views is obtained. In children, and some adults, the aortic arch can be visualized from a suprasternal position.

The standard nomenclature for 2-D echocardiographic images is shown in Figure 13.24a. The parasternal position of the transducer gives access to the long-axis and short-axis planes. The apical approach gives a second view of the long-axis plane, but with the apex in the foreground, and also shows the four-chamber plane (Fig. 13.24e). A three-chamber apical view is sometimes also obtained. The convention adopted in most adult cardiology units of showing the transducer position at the top of the image results in the apical views being 'upside-down' compared to the anatomical position of the heart. M-mode recordings are obtained from the parasternal position to document motion patterns of the aorta, aortic valve and left atrium, the mitral valve, and the left and right ventricles (Fig. 13.24c).

Quantification in echocardiography

Quantification is helpful because of the relative inaccuracy of *subjective* assessment of chamber size and cardiac function. M mode is used to assess lengths in a single plane. The 1 cm calibration markers on M-mode recordings permit



(a) (b)

Fig. 13.27 **Cardiac echograms. (a) 2-D Echocardiogram (long-axis view) in a patient with calcific aortic stenosis.** The calcium in the valve generates abnormally intense echoes. There is some evidence of the associated left ventricular hypertrophy. **(b) Continuous-wave (CW) Doppler signals** obtained from the right upper parasternal edge, where the high-velocity jet from the stenotic valve is coming towards the transducer. AV, aortic valve; LA, left atrium; MV, mitral valve; LV, left ventricle; IVS, interventricular septum; PVW, posterior ventricular wall.

measurement of cardiac dimensions at any point in the cardiac cycle with a typical accuracy of ± 2–3 mm. M mode can be used to estimate LV systolic function by comparing end-diastolic and end-systolic dimensions. For example, the percentage reduction in the left ventricular cavity size ('shortening fraction' – SF) is given by:

$$SF = \frac{LVDD - LVSD}{LVDD} \times 100\%$$

where LVDD is left ventricular diastolic diameter and LVSD is left ventricular systolic diameter, at the base of the heart. The normal range is 30–45%.

This method is easy to perform, but is an inaccurate measure of ejection fraction (EF) because it does not take account of reduced regional function of the mid or apical myocardium, due to infarction for example. For this reason, estimation of EF based upon the difference in LV volumes from systole to diastole, derived from planimetered measurements of LV area in at least two planes, is more accurate. A normal EF is >55%. This method is helpful in assessing the response of the patient with heart failure to therapy. It also permits estimation of LV mass.

Transoesophageal echocardiography (TOE)

This involves placing the transducer mounted on a flexible tube into the oesophagus. This involves the use of local anaesthesia and sometimes intravenous sedation. High-resolution images can be obtained because of the close proximity of the heart to the transducer in the oesophagus, and also because of the higher frequencies that are used relative to transthoracic imaging. TOE is most commonly used in the assessment of valve structure and function (to assess for reparability of mitral valve prolapse), for features and complications of infective endocarditis, to assess the aorta for aortic dissection and to assess for a cardiac source of embolus.

Wall motion stress echocardiography

Echocardiography can be used clinically to evaluate the patient for the presence of myocardial scars and of reversible ischaemia. Since ultrasound cannot directly detect red blood cells in capillaries, myocardial wall motion is used as a surrogate for perfusion. Myocardial segments that demonstrate a change in function (defined as a change or reduction in thickening) from rest to stress can be assumed to be supplied by a flow-limiting stenosis in the epicardial artery or graft.

Stress for this indication needs to be inotropic to induce true ischaemia. Physiological stress includes treadmill exercise, which is complicated by the difficulty in obtaining reliable images rapidly as the patient comes off the treadmill, before heart rate reduces back to sub-maximal levels. Alternatively, pharmacological stress can be induced with dobutamine at graded doses. This is relatively safe but complications such as ventricular arrhythmia have been reported.

This technique can also be used to assess for viability of the myocardium and for hibernating or stunned myocardium (p. 740).

Myocardial perfusion echocardiography

In order to assess myocardial perfusion by echocardiography (MPE), microspheres of similar size to red blood cells are used as an intravenous contrast agent. Microsphere-specific ultrasound modes such as harmonic power Doppler can be used for detection. MPE involves the use of intravenous infusion of contrast to fill the myocardium. A pulse of ultrasound destroys microspheres within the capillaries (and not the LV

cavity), and the time taken to replenish the capillaries is a measure of myocardial blood flow. The time taken to fill should be significantly shorter at stress than at rest.

Clinical use of echocardiography

The echocardiographic findings in particular conditions are discussed in relevant sections, but a brief overview is given below.

Valve stenosis

Congenitally abnormal aortic or pulmonary valves show a characteristic 'dome' shape in systole because the cusps cannot separate fully and a bicuspid configuration may be demonstrated. The presence of calcium in a valve gives rise to intense echoes that generate multiple, parallel lines on M-mode recordings. CW Doppler (p. 699) directed from the apex measures velocity of the jet crossing the diseased valve, from which the pressure gradient can be calculated (Fig. 13.27b).

Mitral stenosis

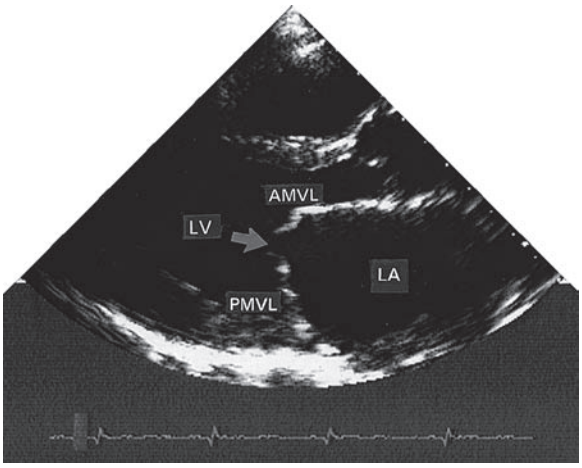
The M-mode shows restriction and reversal of direction of the posterior leaflet motion (Fig. 13.28c). A short-axis view shows the shape of the mitral orifice in diastole, and its area can be measured directly from the image. Peak, mean and end-diastolic pressure gradients can be obtained from CW Doppler. Additional imaging views indicate the size of the left atrium, and may show the presence of left atrial thrombus.

Valve regurgitation

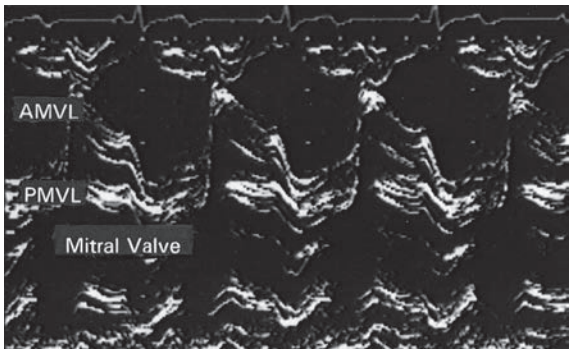
Doppler is extremely sensitive for detecting valve regurgitation and care must be taken not to overestimate severity of regurgitation. Transthoracic echocardiography is as accurate as angiography in the evaluation of valvular regurgitation, provided appropriate algorithms for interpretation are used. Echocardiography can determine the underlying cause such as rheumatic disease or mitral valve prolapse.

Aortic aneurysms and dissections

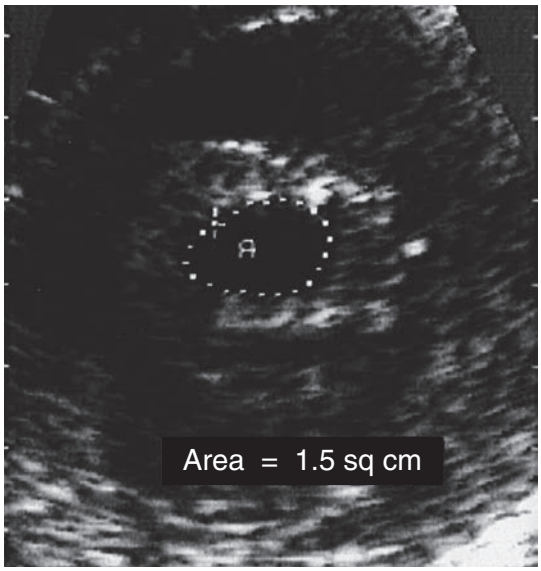
Dilatation of the aortic root can be measured accurately and the presence of a reflecting structure within the lumen of the aorta is strongly suggestive of an intimal flap associated with dissection. Transoesophageal views are particularly suitable for detecting pathology in the ascending and descending aorta. In some hospitals this is the investigation of first choice



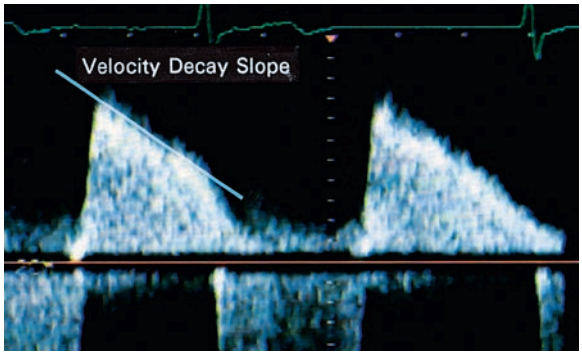
(a)



(c)



(b)



(d)

Fig. 13.28 **Echocardiograms in rheumatic mitral valve disease. (a) 2-D Long-axis view** showing enlarged left atrium and 'hooked' appearance of the mitral valve leaflets resulting from commissural fusion. **(b) Magnified short-axis view** showing the mitral valve orifice as seen from the direction of the arrow in (a). The orifice area can be planimetered to assess the severity; in this case it is 1.5 cm^2, indicating moderately severe disease. **(c) M-mode recording** of the mitral valve showing restricted motion of the thickened leaflets. **(d) Continuous-wave (CW) Doppler** recording showing slow rate of decay of flow velocity from the left atrium to the left ventricle during diastole. It is also possible to derive the valve orifice area from the velocity decay rate. LA, left atrium; LV, left ventricle; AMVL, PMVL, anterior and posterior mitral valve leaflets.

when aortic dissection is suspected although MRI is increasingly used.

Prosthetic heart valves

Each type of mechanical heart valve prosthesis has characteristic echocardiographic features. Irregularity or restriction of movement can be shown. Bioprostheses can also be assessed echocardiographically. The presence of stenosis or regurgitation may be documented by colour flow and CW Doppler. Prosthetic valves imaged transthoracically cast acoustic shadows that obscure the area behind the valve. For this reason a transoesophageal approach is often used to inspect a prosthetic mitral valve for evidence of endocarditis or thrombosis.

Infective endocarditis

Vegetations >2 mm can be detected (Fig. 13.87a, p. 772). Complications of endocarditis such as valvar regurgitation and abscess can be identified.

Cardiac failure

Left ventricular function and response to treatment is readily assessed and should be performed in all patients with known or suspected heart failure. Left- and right-sided heart cardiac output and filling pressures can also be estimated.

Cardiomyopathies

Dilated cardiomyopathy is characterized by an enlarged, globular shaped, thin-walled left ventricle with poor function and low stroke output shown by reduced movements of the valves (Fig. 13.114, p. 793).

Hypertrophic cardiomyopathy (HCM) (Fig. 13.107, p. 790) shows that the left ventricle is typically small and hypertrophied, with a thickened, interventricular septum (asymmetric septal hypertrophy – ASH). There is a characteristic, displacement of the mitral valve apparatus towards the septum in systole (systolic anterior motion – SAM). Some patients have outflow tract or other LV obstruction and others have regional hypertrophy of other segments.

Restrictive cardiomyopathy shows the typical appearances of conditions such as amyloid, and LV diastolic function can be evaluated.

Pericardial effusion

Fluid in the pericardial cavity shows as an echo-free region between the myocardium and the intense echo of the parietal pericardium (Fig. 13.29). The most severe manifestation of pericardial effusion is tamponade, which is a clinical and not an echocardiographic diagnosis. The size of a pericardial effusion alone does not indicate tamponade. Echocardiographic features of haemodynamically significant effusions include marked (>25%) respiratory variation of tricuspid, aortic or mitral Doppler tracings. If the inferior vena cava is small and collapses with respiration, it is unlikely that the patient is in tamponade.

Masses within the heart

Echocardiography is a sensitive method for detecting masses within the heart (Fig. 13.104, p. 787), and their haemodynamic significance.

Ischaemic disease

Coronary arteries cannot be imaged adequately using echo techniques, but images may be useful for the diagnosis of complications related to myocardial infarction, such as mitral papillary muscle rupture, tamponade or ventricular septal rupture. The use of stress echocardiography to identify chronic coronary artery disease is discussed above (p. 702).

In the post-infarction period, echocardiography and Doppler are used to diagnose left ventricular aneurysm, left ventricular thrombus, mitral regurgitation and pericardial effusion as well as to assess left ventricular function (ejection fraction).

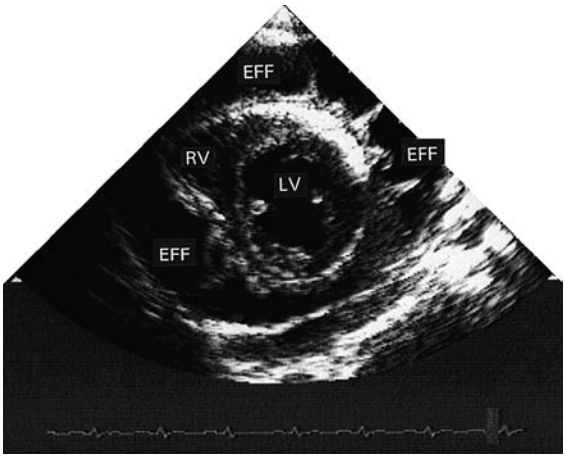


Fig. 13.29 **2-D Echocardiogram (short-axis view) from a patient with a large pericardial effusion** associated with pulmonary tuberculosis. The exudate is seen between the visceral and parietal layers of the pericardium and would give a false impression of cardiomegaly on a chest X-ray. Note the multiple fibrous strands within the effusion, showing that it is consolidating and will probably lead to constriction of cardiac function. LV, left ventricle; RV, right ventricle; EFF, effusion.

Congenital heart disease

Echocardiography has largely replaced cardiac catheterization and angiography. The aim of the examination is first to establish the sequence of blood flow through the heart, and to define anatomical abnormalities as well to evaluate the postoperative status of the operated congenital heart patients.

Contrast echo for LV opacification

Intravenous contrast agents opacify the left ventricle and can define the endocardial border. Their clinical utility has reduced with the advent of harmonic imaging, which has improved image quality in previously 'difficult to image' patients.

Intravascular (coronary) ultrasound

Intravascular (coronary) ultrasound probes can be used to image proximal coronary arteries as part of a percutaneous transluminal coronary angioplasty (PTCA) procedure, e.g. to assess for adequacy of deployment of intracoronary stents.

Nuclear imaging

Nuclear imaging is used to detect myocardial infarction or to measure myocardial function, perfusion or viability, depending on the radiopharmaceutical used and the technique of imaging. These data are particularly valuable when used in combination.

Image type

Gamma cameras produce a planar image in which structures are superimposed as in a standard radiograph. Single-photon-emission computed tomography (SPECT) imaging uses similar raw data to construct tomographic images, just as a CT image is reconstructed from X-rays. This gives finer anatomical resolution, but is technically demanding. These methods may be used with any of the radiopharmaceuticals.

Myocardial perfusion and viability

Thallium-201 is rapidly taken up by the myocardium, so an image taken immediately after injection reflects the distribution of blood flow to the myocardium. Areas of ischaemia or infarction receive less ^{201}Tl and appear dark. Between 2 and 24 hours after injection, ^{201}Tl is redistributed so that all cardiac myocytes contain a comparable concentration. Images at this time show dark areas where the myocardium has infarcted, but normal density in ischaemic areas. Comparison of the early and late images is one method of predicting whether an ischaemic area of myocardium contains enough viable tissue to justify coronary bypass or angioplasty.

Technetium-99-labelled tetrofosmin (Fig. 13.30) is also taken up rapidly by cardiac myocytes, but does not undergo redistribution. When this substance is injected during exercise, its distribution in the myocardium reflects the distribution of blood at the time of the exercise, even if the image is taken several hours later. This is a sensitive method of detecting myocardial viability. Images produced following injection of [^{99m}Tc]tetrofosmin during exercise can be compared to images produced following injection at rest to decide which areas of ischaemia are reversible (p. 740). In patients unable to exercise, the heart can be stressed with drugs, e.g. dipyridamole or dobutamine.

Infarct imaging

Perfusion images produced using compounds labelled with ^{201}Tl or [^{99m}Tc]sestamibi show a myocardial infarction as a perfusion defect or 'cold spot'. These methods are sensitive for detecting and localizing the infarct, but give no information about its age. [^{99m}Tc]Pyrophosphate is preferentially taken up by myocardium which has undergone infarction within the previous few days. Images are difficult to interpret because the isotope is also concentrated by bone and cartilage.

Cardiac computed tomography

Computed tomography is useful for the assessment of the thoracic aorta and mediastinum. The development of 64-slice multidetector CT (MDCT) scanners has enabled accurate non-invasive imaging of the coronary arteries.

Coronary artery calcification

Calcium is absent in normal coronary arteries but is present in atherosclerosis and increases with age. Studies have

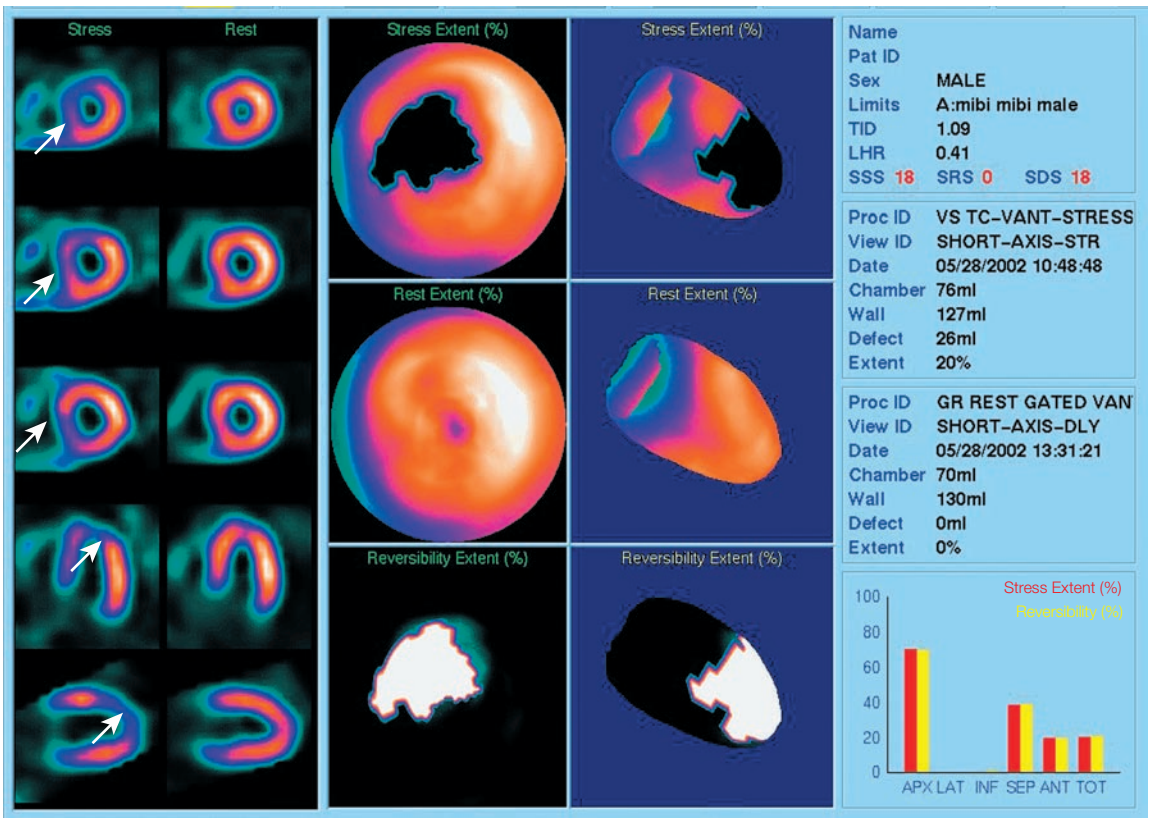


Fig. 13.30 **Myocardial SPECT study acquired with [^{99}Tc]tetrofosmin tracer.** Left panel – three short axis slices, a horizontal and a vertical axis plane of the left ventricle after stress and following a resting re-injection of tracer. The **rest** images demonstrate normal tracer uptake (orange signal) in the whole of the left ventricle. The **stress** images demonstrate reduced tracer uptake (arrows; purple–blue signal) in the anterior and septal walls consistent with a significant stenosis in the left anterior descending artery. Middle panel – polar maps of the whole myocardium can localize the ischaemic territory. Right panel – quantitative analysis can help define the extent and reversibility of the ischaemia.

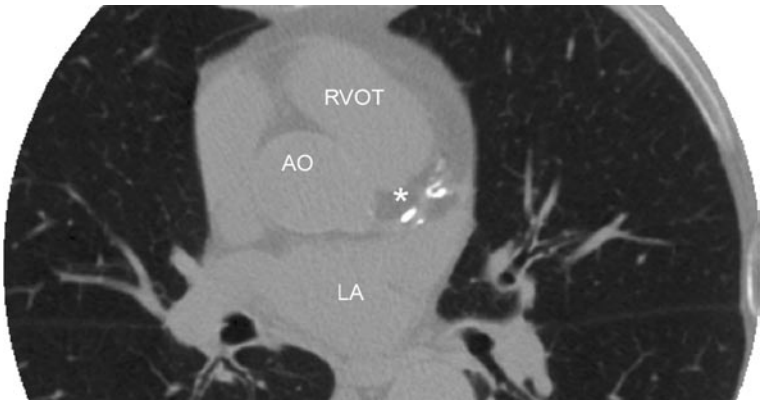


Fig. 13.31 **Multi-slice CT showing calcium deposits (*)** in the left main stem and left anterior descending coronary arteries. AO, aorta; LA, left atrium; RVOT, right ventricular outflow tract.

demonstrated a positive correlation between calcification and the presence of coronary artery stenoses, although the relationship is non-linear. Electron beam CT (EBT) and MDCT scanners are used to obtain multiple thin axial slices through the heart (Fig. 13.31) and then the calcium score is calculated. The calcium score is based on the X-ray attenuation coefficient or CT number measured in Hounsfield units. Meta-analyses have demonstrated that a higher calcium score is associated with higher event rate and higher relative risk ratios, although currently no study has shown a net effect on health outcomes of calcium scoring.

CT coronary angiography

CT coronary angiography (CTCA) is performed with a supine patient connected to a 3-lead ECG for cardiac synchronization. Sixty-four-slice MDCT scanners have a temporal resolution of 165–210 ms and image quality is optimal with a slow and steady heart rate (<65–70 b.p.m.), which can be obtained with the use of oral or intravenous beta-blockers. Sublingual nitroglycerin (0.4–0.8 mg dose) may improve visualization of the coronary artery lumen. A volume dataset containing the whole heart is acquired during a single breath-hold with the injection of 60–80 ml of iodinated contrast agent at 4–6 ml/s. The radiation dose during the scan is 11–22 mSv but this can be reduced to 7–11 mSv with ECG-controlled dose modulation; this compares to 2.5–5.0 mSv for diagnostic coronary angiography and 15–20 mSv for SPECT. The volume dataset is then analysed with multiplanar reformatting for the presence of coronary artery stenoses (Fig. 13.32). Recent studies have reported high sensitivity (>85%) and specificity (>90%) for the detection of coronary artery disease with a very high negative predictive value (>95%). CTCA may become part of an acute chest pain service in the emergency medicine department to exclude aortic dissection, pulmonary embolism and coronary artery disease.

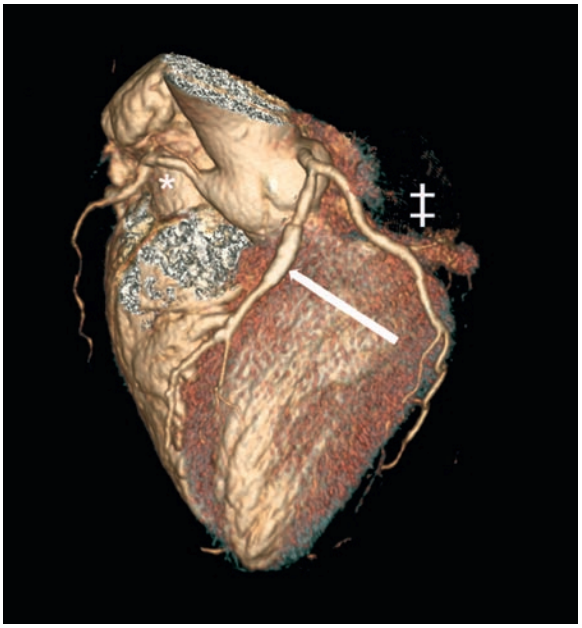


Fig. 13.32 **Multi-slice CT – volume rendered image in a patient with coronary artery stent insertion** (arrow) demonstrating a patent left anterior descending coronary artery with mild stenosis before the stent. The right (*) and the circumflex (++) coronary arteries are free from significant stenoses.

Cardiovascular magnetic resonance

Cardiovascular magnetic resonance (CMR), a non-invasive imaging technique that does not involve harmful radiation, is increasingly used in the investigation of patients with cardiovascular disease.

CMR is usually performed with multiple breath-holds to minimize respiratory motion artefacts and cardiac gating to reduce blurring during the cardiac cycle. Several different sequences are used to provide anatomical and functional information. Most sequences do not require a contrast agent but intravenous gadolinium may be required for magnetic resonance angiography, myocardial perfusion, infarct and fibrosis imaging. The major contraindications are permanent pacemaker or defibrillator, intracerebral clips and significant claustrophobia. Patients with coronary stents and prosthetic valves can be safely scanned.

Clinical use of CMR

The current indications for CMR are summarized in Table 13.6.

Congenital heart disease

CMR provides additional and complementary information to echocardiography in patients with congenital heart disease.

www.studentconsult.com
Ask the authors

Ciné imaging can accurately assess systemic and non-systemic ventricular function and mass. Extra-cardiac conduits, anomalous pulmonary venous return and aortic coarctations pre- and post-repair can be studied by CMR and the studies repeated for long-term follow-up without the risk of ionizing radiation.

Cardiomyopathies, pericardial disease and cardiac masses

In hypertrophic cardiomyopathy CMR accurately defines the extent and distribution of myocardial hypertrophy and can be used in patients with sub-optimal echocardiograms. Intravenous gadolinium can be used to demonstrate regional myocardial fibrosis which is associated with an adverse prognosis. In patients with suspected arrhythmogenic right ventricular cardiomyopathy, CMR is the imaging investigation of choice to detect global and regional wall motion abnormalities of the right ventricle and right ventricular outflow tract, and to detect fatty or fibro-fatty infiltration of the right and left ventricles. In constrictive pericarditis and restrictive cardiomyopathy CMR can demonstrate the effects of the impaired ventricular filling common to both conditions (dilated right atrium and inferior vena cava), but can also determine the thickness of the pericardium (usually 4 mm in normal individuals) (Fig. 13.33). In patients with dilated cardiomyopathy CMR can accurately quantify bi-ventricular function and with gadolinium can demonstrate myocardial fibrosis. In inflammatory and infiltrative conditions of the myocardium, such as myocarditis, sarcoidosis and amyloidosis, CMR is increasingly used as a diagnostic investigation due to different patterns of signal enhancement seen with gadolinium. In patients with thalassaemia, CMR can detect iron deposition within the myocardium and guide chelation therapy. CMR can be useful in patients with cardiac masses to differentiate benign from malignant tumours and to identify thrombus not visualized on echocardiography.

Table 13.6 Cardiovascular magnetic resonance (CMR)

Indications

1. Congenital heart disease (CHD):
   - anatomical assessment following echocardiography, particularly in patients with complex CHD or following surgical intervention
   - follow-up/surveillance studies
   - anomalous pulmonary or systemic venous return
   - assessment of right or left ventricular function/dilatation
2. Cardiomyopathies/cardiac infiltration/pericardial disease
3. Disease of the aorta, including aortic dissection, aneurysm, coarctation
4. Valvular heart disease:
   - quantification of stenosis and regurgitation
   - accurate ventricular dimensions/function
   - planimetry of valvular stenosis
5. Coronary artery disease:
   - left and right ventricular function
   - wall motion assessment during dobutamine stress
   - myocardial perfusion during adenosine stress
   - coronary artery and coronary artery bypass graft imaging
   - myocardial infarct imaging and viability assessment with dobutamine CMR or delayed enhancement
6. Pulmonary vessels

Diseases of the aorta

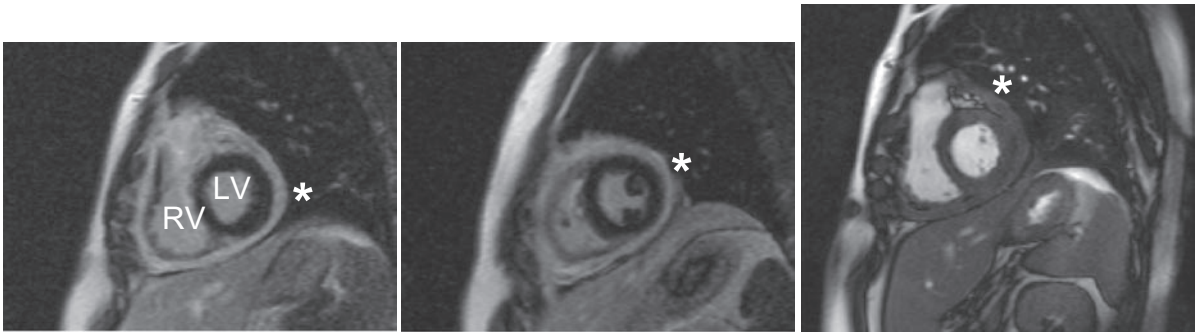
CMR is an excellent technique for assessing patients with aortic dissection and can detect the clinical features of an aortic dissection; the intimal flap, thrombosis in a false lumen, aortic regurgitation, pericardial effusion and aortic dilatation. As it does not involve radiation or need contrast, CMR is an ideal method of surveillance of patients with dilated thoracic aortas or repaired coarctation.

Valvular heart disease

Valvular stenosis produces signal void on gradient-echo CMR. CMR can quantify the velocity across a stenosed valve using phase-contrast velocity mapping. Valvular regurgitation can be accurately quantified using phase-contrast velocity mapping across the valve, or by calculating the stroke volumes of the left and right ventricle which are equal in the absence of significant regurgitation. However, in most patients, transthoracic and transoesophageal echocardiography should provide sufficient information.

Coronary artery disease

CMR can be used to assess coronary artery anatomy, left ventricular function, myocardial perfusion and viability in a 'one-stop' approach to the assessment of patients with coronary artery disease. Coronary artery anatomy and



(a) (b) (c)

Fig. 13.33 **Short-axis cardiac magnetic resonance (CMR) in a patient with constrictive pericarditis.** The pericardium is thickened (*) and on **ciné imaging (c)** there is septal bounce. LV, left ventricle; RV, right ventricle.

stenoses can be identified with ultra-fast breath-hold or respiratory-gated sequences with high accuracies. Global left ventricular function and wall motion abnormalities can be detected with ciné imaging performed at rest and during dobutamine stress. Myocardial perfusion can be assessed with gadolinium and first-pass imaging; ischaemia can be demonstrated with adenosine for coronary vasodilatation. Myocardial viability can be determined using gadolinium and 'delayed enhancement' images. With these techniques CMR is increasingly used both for the assessment of ischaemia in patients with suspected coronary disease and to assess myocardial viability prior to revascularization in patients with impaired cardiac function (Figs 13.34 and 13.35).

FURTHER READING

Budoff MJ, Ashenbach S, Blumenthal RS, et al. Assessment of coronary artery disease by cardiac computed tomography: a scientific statement from the American Heart Association Committee on Cardiovascular Imaging and Intervention, Council on Cardiovascular Radiology and Intervention, and Committee on Cardiac Imaging, Council on Clinical Cardiology. *Circulation* 2006; **114**: 1761–1791.

Hoffmann U, Ferencik M, Cury RC, Pena AJ. Coronary CT angiography. *Journal of Nuclear Medicine* 2006; **47**: 797–806.

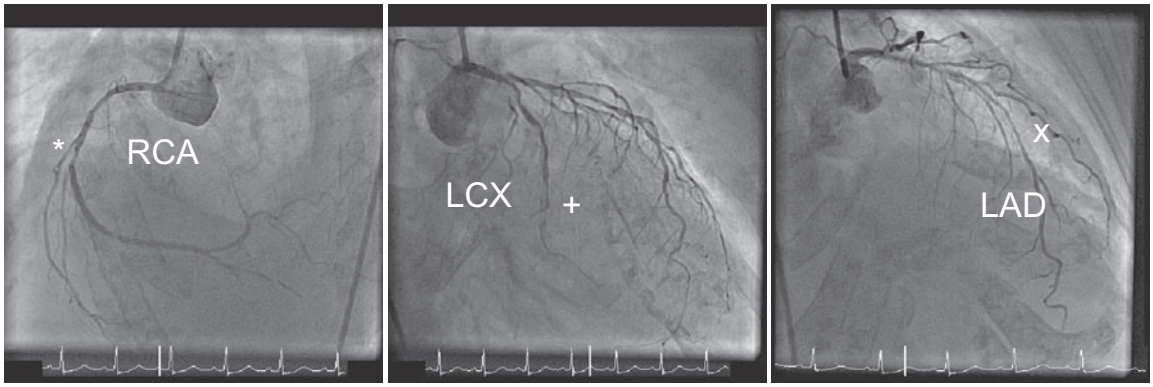


Fig. 13.34 **Cardiac magnetic resonance (CMR). A diagnostic coronary angiography in a patient with Type 2 diabetes mellitus and exertional breathlessness** demonstrates a severe stenosis (*) in the right coronary artery (RCA), a sub-totally occluded (+) circumflex coronary artery (LCX), and a long segment of disease (x) in the left anterior descending coronary artery (LAD).

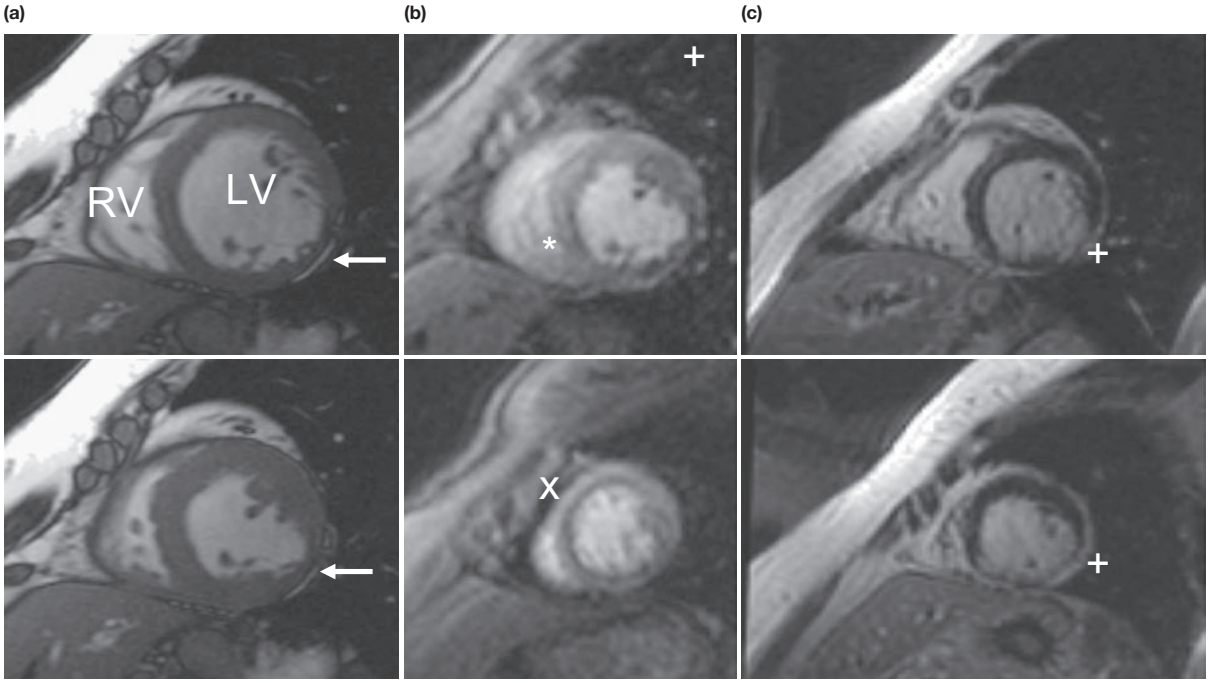


Fig. 13.35 **Cardiac magnetic resonance (CMR). (a) Short axis CMR FIESTA ciné imaging** (diastole top, systole bottom) demonstrates thinning and hypokinesia of the inferolateral wall (arrows) but with preserved function of the rest of the myocardium. LV, left ventricle; RV, right ventricle. **(b) First-pass perfusion CMR (mid top, apex bottom)** demonstrates sub-endocardial perfusion defects in the inferolateral wall (+), infero-septum (*) and apical segments (x). **(c) Delayed enhancement CMR confirms a transmural myocardial infarction** of the mid-apical infero-lateral wall (+) secondary to a previous infarction from the circumflex coronary artery and confirms inducible ischaemia in the right and left coronary territories.

Pulmonary vessels

Magnetic resonance angiography with gadolinium can provide high quality images of the pulmonary veins which can be fused with electrical data during pulmonary vein isolation for the treatment of atrial fibrillation.

Positron emission tomography (PET)

This is a technique based on detection of high-energy emissions caused by annihilation of positrons released from unstable isotopes. There are several advantages of PET over other techniques, e.g. improved spatial resolution, accurate quantification, the use of biological isotopes of carbon, nitrogen and oxygen. However, PET is expensive and requires a cyclotron to produce the short-lived tracers. PET has become a useful investigation in the detection of viable myocardium in patients who are suitable for revascularization.

Myocardial perfusion and ischaemia can be determined using PET with ^{13}N-ammonia or oxygen-15 with greater sensitivity than SPECT. Myocardial metabolism and viability can be detected with the use of ^{18}F-fluorodeoxyglucose (FDG), which the cardiac myocyte utilizes for energy production in the presence of reduced oxygen supply and blood flow. There may be reduced perfusion to infarcted or fibrotic myocardium, but also reduced FDG uptake. In hibernating myocardium, with viable but dysfunctional myocardium, PET can demonstrate reduced myocardial perfusion but with preserved or increased FDG uptake.

Cardiac catheterization

Cardiac catheterization is the introduction of a thin radio-opaque tube (catheter) into the circulation. The right heart is catheterized by introducing the catheter into a peripheral vein (usually the right femoral or internal jugular vein) and advancing it through the right atrium and ventricle into the pulmonary artery. The pressures in the right heart chambers, and pulmonary artery can be measured directly. An indirect measure of left atrial pressure can be obtained by 'wedging' a catheter into the distal pulmonary artery (p. 902). In this position the pressure from the right ventricle is obstructed by the catheter and only the pulmonary venous and left atrial pressures are recorded.

Left heart catheterization is usually performed via the right femoral artery, although the brachial and radial arteries are sometimes used in patients with significant peripheral vascular disease. A pigtail catheter is advanced up the aorta and manipulated through the aortic valve into the left ventricle. Pressure tracings are taken from the left ventricular cavity. The end-diastolic pressure is invariably elevated in patients with left ventricular dysfunction. A power injection of radio-opaque contrast material is used to opacify the left ventricular cavity (left ventriculography) and thereby assesses left ventricular systolic function. The catheter is then withdrawn across the aortic valve into the aorta and the 'pullback' gradient across the valve is measured.

Aortography (a power injection into the aortic root) can be performed to assess the aortic root and the presence and severity of aortic regurgitation.

Specially designed catheters are then used to selectively engage the left and right coronary arteries, and contrast ciné-angiograms are taken in order to define the coronary circulation and identify the presence and severity of any coronary artery disease. Coronary angiography is described further on page 749.

During cardiac catheterization, blood samples may be withdrawn to measure the concentration of ischaemic metabolites (e.g. lactate) and the oxygen content. These estimations are used to gauge ischaemia, quantify intracardiac shunts and measure cardiac output.

Digital subtraction angiography

This technique permits the injection of small volumes of radio-contrast agents during cardiac catheterization with the production of computer-analysed high-quality angiograms.

Unfortunately, peripheral injection of contrast does not give adequate visualization of the coronary arteries, but aortic lesions can be visualized.

THERAPEUTIC PROCEDURES

Cardiac resuscitation

When cardiac arrest occurs, basic life support must be started immediately. The longer the period of respiratory and circulatory arrest, the lower is the chance of restoring healthy life. After 3 minutes, irreversible anoxic cerebral damage occurs.

All healthcare professionals should be familiar with basic procedures. Organization and teamwork are essential to a successful outcome. Cardiac massage alone by bystanders is the preferable approach to resuscitation for adult patients with witnessed out-of-hospital cardiac arrest.

Basic life support (BLS)

The first step is to ensure the safety of the victim and rescuer. The next is to ascertain that the victim is unresponsive by shaking him/her and shouting into one ear. If no response is obtained, help should be sought immediately prior to commencement of basic life support. If the likely cause of the unconsciousness is due to a breathing problem (e.g. trauma, drowning, choking, drug/alcohol intoxication or if the victim is an infant or child) then the rescuer should perform resuscitation for about 1 minute before going for help. Basic life support is easily remembered as A (airway), B (breathing) and C (circulation) (Emergency box 13.1).

Airway

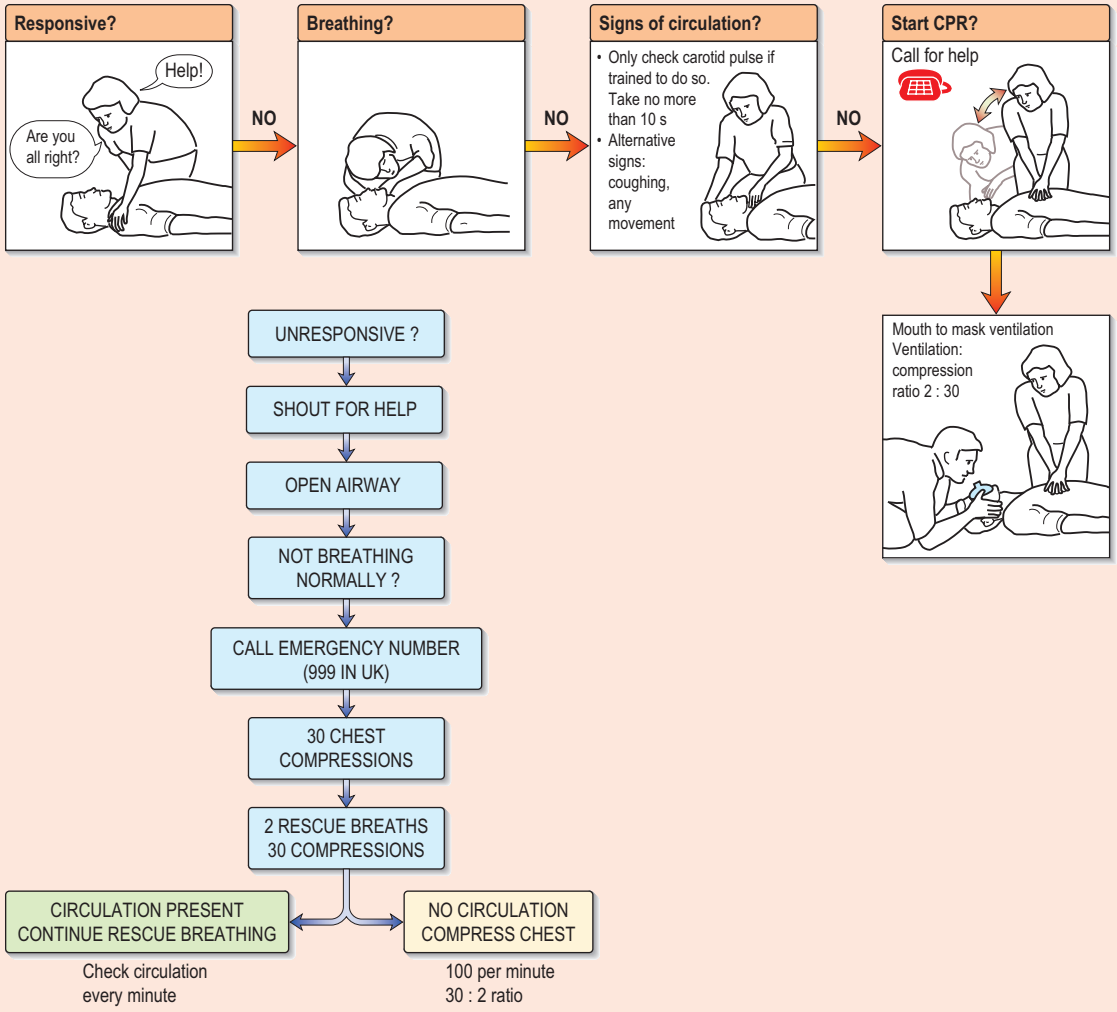
Debris (e.g. blood and mucus) in the mouth and pharynx should be removed. Loose or ill-fitting dentures should be removed. The airway should be opened gently by flexing the neck and extending the head ('sniffing the morning air' position). This manoeuvre is not recommended if a cervical spine injury is suspected. Any obstruction deep in the oral cavity or upper respiratory tract may require abdominal and/or chest thrusts (Heimlich manoeuvre, p. 834).

Breathing

Once a clear airway has been established the victim's breathing should be assessed by the 'look, listen and feel' method. By placing a cheek close to the victim's mouth, breath can be felt and breath sounds can be heard, and

Emergency Box 13.1

Basic life support



Send or go for help as soon as possible according to guidelines

the rise and fall of the chest and abdomen observed. If there is no evidence of breathing, expired air ventilation should be commenced.

With the head of the victim tilted backwards (*head tilt*), the chin pulled forward (*chin lift*) or (*jaw thrust*) and the nostrils pinched firmly, the rescuer takes a deep breath and seals his/her lips around the mouth of the victim. Two effective breaths are given over 1 second each. Expired air respiration is the only method of artificial respiration that successfully ventilates the patient. If resistance to these puffs is experienced, the airway needs to be reassessed and/or the head tilt and jaw lift corrected.

If the rescuer feels reluctant to give mouth-to-mouth resuscitation because of safety concerns (e.g. potentially infective material such as blood or vomitus in the oropharynx) it is acceptable to perform chest compressions alone, or to use a mouth piece if available.

Circulation

Cardiac arrest is accompanied by circulatory collapse. Absence of the carotid pulse confirms this and should be assessed for at least 10 seconds by someone trained to do so. The carotid pulse lies lateral and posterior to the thyroid cartilage and medial to the medial border of sternocleidomastoid: it can sometimes be difficult to palpate.

If absent, the circulation is best re-established by external chest compression. The heel of one hand is placed over the centre of the chest and the heel of the second hand is placed over the first with the fingers interlocked. The arms are kept straight and the sternum is rhythmically depressed by 2–5 cm at a rate of approximately 100 per minute. Chest compressions do not massage the heart. The thorax acts as a pump and the heart provides a system of one-way valves to ensure forward circulation. Respiratory and circulatory support is continued by providing two effective breaths for

every 30 cardiac compressions (30:2 for one or two persons).

It is better to give compressions without interruption. This maintains adequate cerebral and coronary perfusion pressures. There is evidence to suggest that if the person performing compressions tires, the quality of resuscitation deteriorates.

Advanced cardiac life support

By the time effective life support has been established, more help should have arrived and advanced cardiac life support can begin. This consists of ECG monitoring, endotracheal intubation and setting up an intravenous infusion in a large peripheral or central vein. Immediate therapy includes defibrillation, oxygen and cardioactive drugs. It is not possible to recommend an exact sequence of management because it will depend on the arrival of skilled personnel and equipment, and the nature of the cardiac arrest.

As soon as possible the cardiac rhythm should be established. At first this is easily achieved by monitoring the ECG through the paddles of a defibrillator. Later, ECG monitoring can be set up. If the ECG shows ventricular fibrillation or if there is any doubt as to the nature of the rhythm (e.g. 'fine' ventricular fibrillation may be confused with asystole), no time should be lost before defibrillating the patient (<2 minutes in hospital). If initial defibrillation attempts are unsuccessful, time can then be spent intubating the patient and setting up an intravenous infusion whilst the circulation is supported by external chest compression.

If there is any difficulty in intubating the patient, ventilation should be continued by means of an airway, a ventilating bag and oxygen. Intravenous epinephrine (adrenaline) results in vasoconstriction and increases the proportion of the cardiac output delivered to the brain.

Causes of unexpected cardiac arrest

Each year in the UK there are approximately 100 000 unexpected deaths occurring within 24 hours of the development of cardiac symptoms. About half of these deaths are almost instantaneous. There are several causes (Table 13.7).

Most deaths are due to ventricular fibrillation or rapid ventricular tachycardia, and a small proportion are due to severe bradyarrhythmias. Coronary artery disease accounts for approximately 80% of sudden cardiac deaths in western society. Transient ischaemia is suspected as the major trigger factor; however, only a small proportion of survivors have clinical evidence of acute myocardial infarction.

There are two mechanisms of sudden unexpected cardiac arrest:

- ventricular fibrillation or pulseless ventricular tachycardia (VF/VT)
- non-VF/VT (asystole and pulseless electrical activity also known as electromechanical dissociation).

Table 13.7 Causes of unexpected cardiac arrest

Cardiac arrhythmias (e.g. ventricular fibrillation)
Sudden pump failure (e.g. acute myocardial infarction)
Acute circulatory obstruction (e.g. pulmonary embolism)
Cardiovascular rupture (e.g. aortic dissection, myocardial rupture)
Vasomotor collapse (e.g. in pulmonary hypertension)

The principal difference in the management of these two groups of arrhythmias is the need for attempted defibrillation in those patients with VF/VT (Fig. 13.36).

Three-quarters of arrests are due to ventricular fibrillation or rapid ventricular tachycardia. Only a very small proportion are due to pulseless electrical activity. The remainder are due to asystole. An agonal rhythm is characterized by an inexorable slowing and widening of the QRS complexes associated with falling blood pressure and cardiac output. This type of arrhythmia is very difficult to reverse and usually no attempt should be made because it is the result rather than the cause of death.

Arrests are treated in the following ways:

- Ventricular fibrillation or pulseless ventricular tachycardia is readily treated with IMMEDIATE defibrillation, cardiopulmonary resuscitation (CPR) and drugs. Intravenous amiodarone is the first-line drug in refractory VF/pulseless VT.
- Asystole is more difficult to treat but the heart may respond to atropine or epinephrine (adrenaline). More recently vasopressin has been shown to be successful. If there is any sign of slow electromechanical activity (e.g. bradycardia with a weak pulse), emergency pacing should be used.
- Pulseless electrical activity: several potentially reversible causes are listed in the universal algorithm (Fig. 13.36). It carries a very poor prognosis. Effective treatment involves addressing the underlying cause.

Figure 13.36 shows the treatments recommended by the European Resuscitation Council and the Resuscitation Council UK.

Defibrillation

This technique is used to convert ventricular fibrillation to sinus rhythm. When the defibrillator is discharged, a high-voltage field envelopes the heart which depolarizes the myocardium and allows an organized heart rhythm to emerge. Electrical energy is discharged through two paddles placed on the chest wall.

The paddles are placed in one of two positions:

- One paddle is placed to the right of the upper sternum and the other over the cardiac apex.
- One paddle is placed under the tip of the left scapula and the other is placed over the anterior wall of the left chest.

Electrode jelly or electrolyte gel pads should be used to ensure good contact between the electrode paddles and the skin. Jelly smeared carelessly across the chest may cause short-circuits and arcing of the charge. All personnel should stand clear of the patient. The person performing defibrillation has the responsibility for ensuring the safety of the patient and other people present.

Conventional defibrillators employ a damped monophasic waveform. Biphasic defibrillators which require less energy are increasingly common. Automated external defibrillators (AEDs) which recognize ventricular fibrillation automatically deliver a shock if indicated. These are available in some public places. It is the responsibility of all healthcare practitioners to be familiar with the range of defibrillators they may be called on to use in their workplace.

Post resuscitation – therapeutic hypothermia

Recent studies have supported the use of therapeutic hypothermia in unconscious adult patients with spontaneous circulation after an out-of-hospital cardiac arrest due to ventricular fibrillation. These patients should be cooled with careful temperature monitoring to 32–34°C for 12–24 hours. External cooling methods include the use of cooling blankets; ice packs to the groin, axillae and neck; wet towels; and a cooling helmet. Invasive methods include intravenous infusions and peritoneal and pleural lavage. Shivering should be prevented by the use of a neuromuscular blocker and sedation to reduce oxygen consumption.

DC-cardioversion (DCC)

Tachyarrhythmias that do not respond to medical treatment or that are associated with haemodynamic compromise (e.g. hypotension, worsening heart failure) may be converted to sinus rhythm by the use of a transthoracic electric shock. A short-acting general anaesthetic is used. Muscle relaxants are not usually given.

When the arrhythmia has definite QRS complexes, the delivery of the shock should be timed to occur with the downstroke of the QRS complex (synchronization) (Fig. 13.37). The machine being used to perform the cardioversion

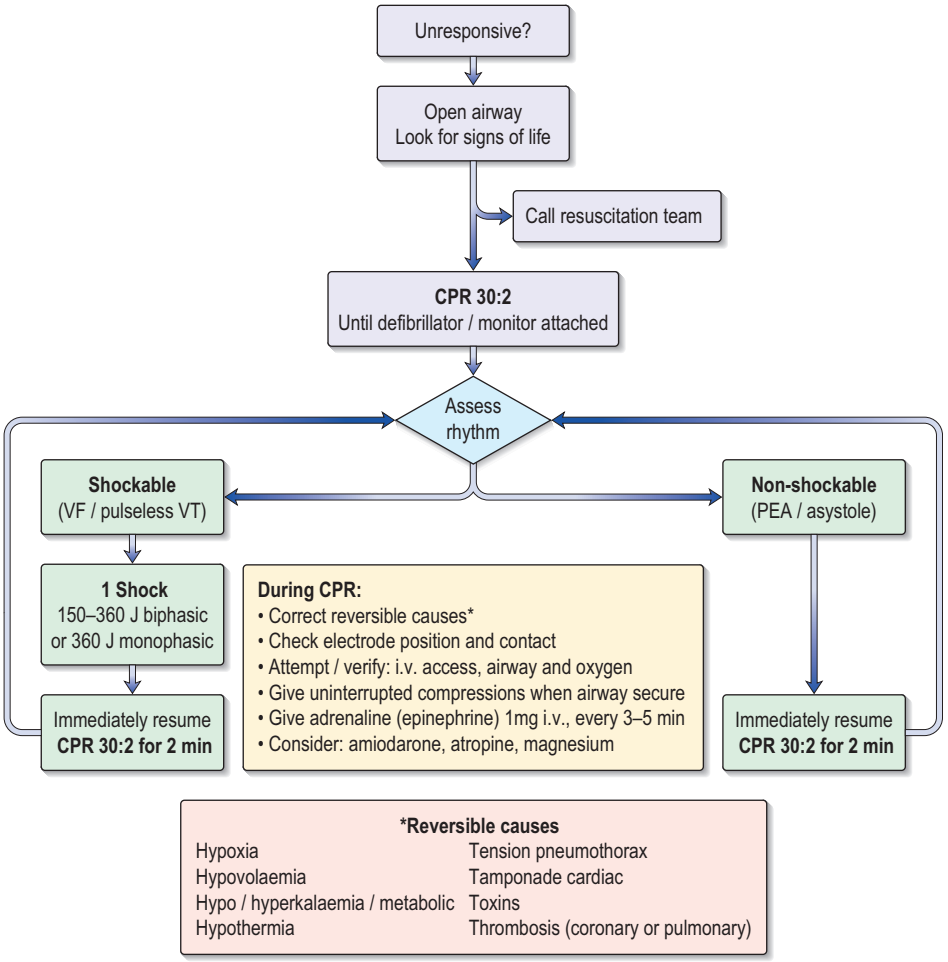


Fig. 13.36 **Adult advanced life support algorithm**. CPR, cardiopulmonary resuscitation; PEA, pulseless electrical activity; VF, ventricular fibrillation. Reproduced with permission from the Resuscitation Council http://www.resus.org.uk/siteindex.htm

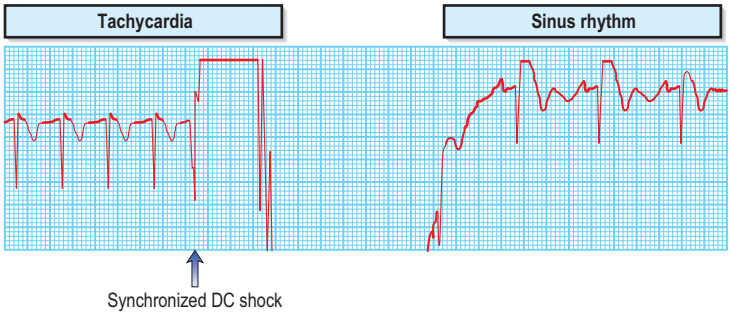


Fig. 13.37 **DC-cardioversion of a supraventricular tachycardia to sinus rhythm**. The direct current shock is delivered synchronously with the QRS complex.

will do this automatically if the appropriate button is pressed. There is a crucial difference between defibrillation and cardioversion: a non-synchronized shock is used to defibrillate. Accidental defibrillation of a patient who does not require it may itself precipitate ventricular fibrillation.

Typical indications for DCC include:

- atrial fibrillation
- atrial flutter
- sustained ventricular tachycardia
- junctional tachyarrhythmias.

If atrial fibrillation or flutter has been present for more than a few days, it is necessary to anticoagulate the patient adequately for 3 weeks before elective cardioversion to reduce the risk of embolization. The duration of anticoagulation after successful cardioversion for atrial fibrillation is a complex issue and depends on a number of factors: it should be for at least 4 weeks after the procedure and may well be for much longer.

Digoxin toxicity may lead to ventricular arrhythmias or asystole following cardioversion. Therapeutic digitalization does not increase the risks of cardioversion, but it is conventional to omit digoxin several days prior to elective cardioversion in order to be sure that toxicity is not present.

Cardiac enzyme levels may rise after a cardioversion.

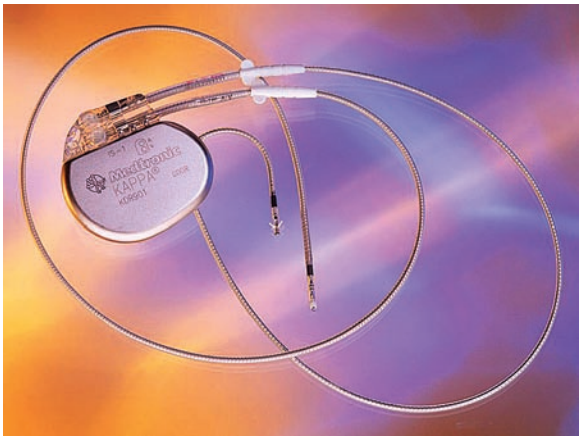
Temporary pacing

Therapeutic cardiac pacing is employed in any patient with sustained symptomatic or haemodynamically compromising bradycardia. Bradycardias may be due to either a slow intrinsic heart rate (e.g. sinus node dysfunction) or atrioventricular block. Prophylactic cardiac pacing is employed in asymptomatic patients with either bradycardia or conduction abnormalities in whom the risk of progression to symptomatic bradycardia justifies such a strategy.

Transvenous pacing is the preferred method in patients with symptomatic bradycardias. In summary, a thin (French gauge 5 or 6), bipolar pacing electrode wire is inserted via an internal jugular vein, a femoral vein or a subclavian vein and is positioned at the right ventricular apex using cardiac fluoroscopy. The energy needed for successful pacing (the pacing threshold) is assessed by reducing the energy until the pacemaker fails to stimulate the tissue (loss of capture). The output energy is then set at three times the threshold value to prevent inadvertent loss of capture. If the threshold increases above 5 V, the pacemaker wire should be resited. A temporary pacemaker unit (Fig. 13.38a) is almost always set to work 'on demand' – to fire only when a spontaneous beat has not occurred. The rate of temporary pacing is usually 60–80 per minute.

Transcutaneous pacing is the preferred method in selected patients with asymptomatic bradycardia or conduction abnormalities and may be life-saving for patients in whom a cardiac arrest is precipitated by bradycardia. In this method the myocardium is depolarized by current flow between two large adhesive electrodes positioned anteriorly and posteriorly on the chest wall. Transcutaneous pacing is uncomfortable for the conscious patient. However, it can usually be tolerated until a temporary transvenous pacemaker is inserted.

(a)



(b)

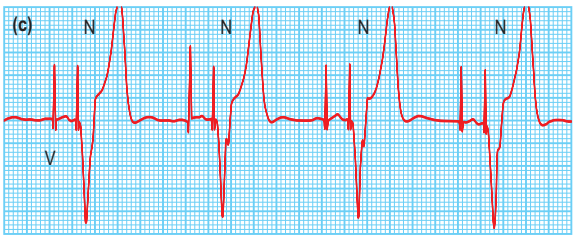


Fig. 13.38 **Pacemakers. (a)** A temporary pacemaker unit. **(b)** A permanent pacemaker with atrial (placed in the right atrium, usually in its high lateral wall) and ventricular (placed in the right ventricular apex) leads. **(c)** An electrocardiogram showing dual (atrial and ventricular) chamber pacing. A wide QRS complex results from abnormal activation of the ventricles from the right ventricular apex.

Permanent pacing

Permanent pacemakers are fully implanted in the body and connected to the heart by one or two electrode leads (Fig. 13.38b). The pacemaker is powered by solid-state lithium batteries, which usually last 5–10 years. Pacemakers are 'programmable' in that their operating characteristics (e.g. the pacing rate) can be changed by a programmer that transmits specific electromagnetic signals through the skin. The pacemaker leads are passed transvenously to the right heart chambers.

Pacemakers are designed to both pace and sense either the ventricles or the atria or more commonly both chambers. A single chamber ventricular pacemaker is described as a 'VVI' unit because it paces the ventricle (V), senses the ventricle (V) and is inhibited (I) by a spontaneous ventricular signal. Occasionally (e.g. in symptomatic sinus bradycardia), an atrial pacemaker (AAI) may be implanted. Pacemakers that are connected to both the right atrium and ventricle ('dual chamber' pacemakers) are used to simulate the natural pacemaker and activation sequence of the heart. This form of pacemaker is called DDD because it paces the two (dual) chambers, senses both (D) and reacts in two (D) ways – pacing in the same chamber is inhibited by spontaneous atrial and ventricular signals, and ventricular pacing is triggered by spontaneous atrial events (Fig. 13.38c).

In addition, pacemakers may be 'rate responsive' (R). A rate-responsive pacemaker detects motion (level of vibration or acceleration), respiration, or changes in QT interval, and by employing one or more biosensors, changes its rate of pacing so that it is appropriate to the level of exertion.

The choice of pacemaker mostly depends on the underlying rhythm abnormality and the general condition of the patient. For example, complete heart block in patients with sinus rhythm should be treated with a dual-chamber device in order to maintain AV synchrony, whereas inactive or infirm patients may not benefit from the most sophisticated units. Specialized biventricular pacemakers are used for the treatment of severe heart failure.

Permanent pacemakers are inserted under local anaesthetic using fluoroscopy to guide the insertion of the electrode leads via the cephalic or subclavian veins. Perioperative prophylactic antibiotics are routinely prescribed. The pacemaker is usually positioned subcutaneously in front of the pectoral muscle. Following surgery, which usually takes 60–90 minutes, the patient rests in bed for 6–12 hours before being discharged. Patients may not drive for at least 1 week after implantation, and must inform the licensing authorities and their motor insurers.

Complications are few but can prove to be very difficult to manage, and patients should be referred to the pacemaker clinic. They include the following:

- infection
- erosion
- pocket haematoma
- lead displacement
- electromagnetic interference.

Pericardiocentesis (Fig. 13.39)

A pericardial effusion is an accumulation of fluid between the parietal and visceral layers of pericardium. Fluid is removed to relieve symptoms that are due to haemodynamic embarrassment or for diagnostic purposes. This can be a technically difficult procedure, particularly in the acute setting. In an emergency it can be performed at the bedside.

Pericardial aspiration or pericardiocentesis is performed by inserting a needle into the pericardial space, usually via a subxiphisternal route under ultrasound guidance. Certain effusions, particularly posterior ones, require surgical drainage under a general anaesthetic. If a large volume of fluid is to be removed, a wide-bore needle and cannula are inserted. The needle may be removed and the cannula left in situ to drain the fluid. Fluid that is removed is sent for chemical analysis, microscopy, including cytology, Gram-stain and culture. If a reaccumulation of pericardial fluid is anticipated, the cannula may be left in place for several days, or an operation can be performed to cut a window in the parietal pericardium (fenestration) or to remove a large section of the pericardium.

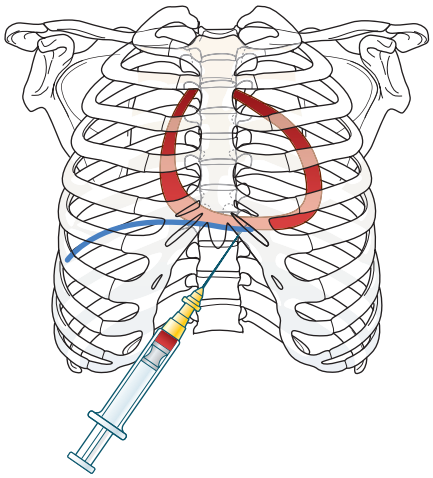


Fig. 13.39 **Schematic diagram of pericardiocentesis showing the site for insertion of the needle.**

Right-heart bedside catheterization (Fig. 15.18)

Bedside catheterization of the pulmonary artery with a pulmonary artery balloon flotation catheter (Swann–Ganz catheter) is performed in patients with:

- cardiac failure
- cardiogenic shock
- doubtful fluid status.

Intra-aortic balloon pumping

This is a technique used to assist temporarily the failing left ventricle. A catheter with a long sausage-shaped balloon at its tip is introduced percutaneously into the femoral artery and manipulated under X-ray control so that the balloon lies in the descending aorta just below the aortic arch (Fig. 13.40). The balloon is rhythmically deflated and inflated with carbon dioxide gas. Using the ECG or intra-aortic pressure changes, the inflation is timed to occur during ventricular diastole to increase diastolic aortic pressure and consequently to improve coronary and cerebral blood flow. During systole the balloon is deflated, resulting in a reduction in the resistance to left ventricular emptying. Intra-aortic balloon pumping is used for circulatory support in the following acute situations:

- *Acute heart failure.* Balloon pumping is used to improve cardiac output when there is a transient or reversible depression of left ventricular function, such as in a patient with severe mitral valve regurgitation who is awaiting surgical replacement of the mitral valve, or in a patient with a ventricular septal defect that is due to septal infarction. It may also be used to support patients awaiting heart transplantation.
- *Unstable angina pectoris.* Balloon pumping is used to treat unstable angina pectoris by improving coronary

FURTHER READING

Resuscitation Council (UK) *Resuscitation Guidelines 2005.* London: Resuscitation Council (UK), 2005.

Email: enquiries@resus.org.uk (website: www.resus.org.uk)

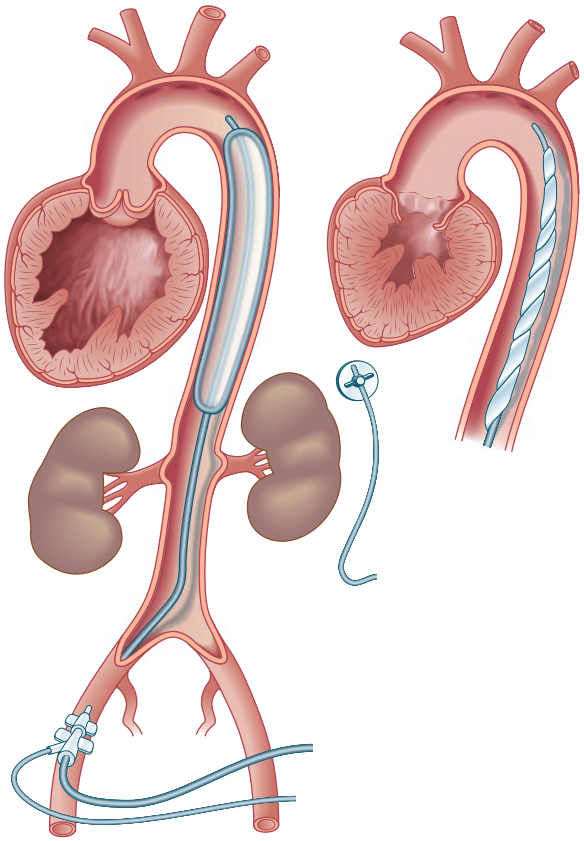


Fig. 13.40 **Intra-aortic balloon pump** – inflated in aorta on the left and deflated on the right.

flow and decreasing myocardial oxygen consumption by reducing the 'afterload'. This technique may be successful, even when medical therapy has failed. It is followed by early angiography and appropriate definitive therapy such as surgery or coronary angioplasty.

Balloon pumping should not be used when there is no remediable cause of cardiac dysfunction. It is also unsuitable in patients with severe aortic regurgitation, aortic dissection and severe peripheral vascular disease.

Complications of balloon pumping occur in about 20% of patients and include aortic dissection, leg ischaemia, emboli from the balloon and balloon rupture. Embolic complications are reduced by anticoagulation with heparin.

CARDIAC ARRHYTHMIAS

An abnormality of the cardiac rhythm is called a cardiac arrhythmia. Arrhythmias may cause sudden death, syncope, heart failure, chest pain, dizziness, palpitations or no symptoms at all. There are two main types of arrhythmia:

- bradycardia: the heart rate is slow (<60 b.p.m.)
- tachycardia: the heart rate is fast (>100 b.p.m.).

Tachycardias are more symptomatic when the arrhythmia is fast and sustained. Tachycardias are subdivided into supraventricular tachycardias, which arise from the atrium or the atrioventricular junction, and ventricular tachycardias, which arise from the ventricles.

Some arrhythmias occur in patients with apparently normal hearts, and in others arrhythmias originate from diseased tissue, such as scar, as a result of underlying structural heart disease. When myocardial function is poor, arrhythmias are more symptomatic and are potentially life-threatening.

SINUS NODE FUNCTION

The normal cardiac pacemaker is the sinus node (p. 681) and, like most cardiac tissue, it depolarizes spontaneously.

The rate of sinus node discharge is modulated by the autonomic nervous system. Normally the parasympathetic system predominates, resulting in slowing of the spontaneous discharge rate from approximately 100 to 70 b.p.m. A reduction of parasympathetic tone or an increase in sympathetic stimulation leads to tachycardia; conversely, increased parasympathetic tone and decreased sympathetic stimulation produces bradycardia. The sinus rate in women is slightly faster than in men. Normal sinus rhythm is characterized by P waves that are upright in leads I, and II of the ECG (Fig. 13.19, p. 696), but inverted in the cavity leads AVR and V1 (Fig. 13.41a).

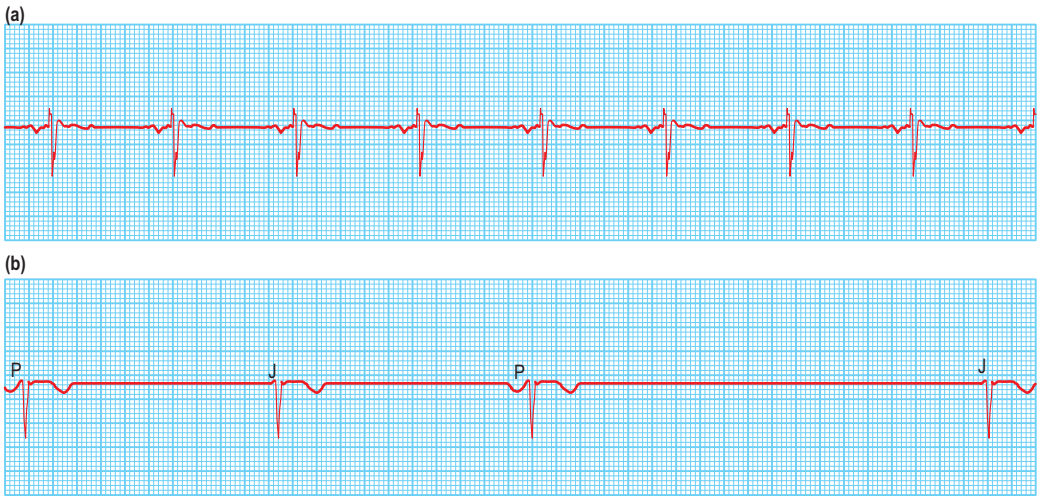


Fig. 13.41 **(a) An ECG showing normal sinus rhythm** (PR interval <0.2 s). P wave preceding each QRS complex. **(b) A patient with sick sinus syndrome**. This shows sinus arrest (only occasional sinus P waves) and junctional escape beats (J). P waves are inverted in the cavity leads that are shown here.

Sinus arrhythmia

Fluctuations of autonomic tone result in phasic changes of the sinus discharge rate. During inspiration, parasympathetic tone falls and the heart rate quickens, and on expiration the heart rate falls. This variation is normal, particularly in children and young adults. Typically sinus arrhythmia results in predictable irregularities of the pulse.

Sinus bradycardia

A sinus rate of less than 60 b.p.m. during the day or less than 50 b.p.m. at night is known as sinus bradycardia.

It is usually asymptomatic unless the rate is very slow. It is normal in athletes owing to increased vagal tone. Other causes may be divided into systemic or cardiac and are discussed below in the section entitled 'Bradycardias and heart block' (see p. 717).

Sinus tachycardia

Sinus rate acceleration to more than 100 b.p.m. is known as sinus tachycardia. Again, causes may be divided into systemic or cardiac and are discussed below in the section entitled 'Supraventricular tachycardias' (p. 720).

Mechanisms of arrhythmia production

Abnormalities of automaticity, which could arise from a single cell, and abnormalities of conduction, which require abnormal interaction between cells, account for both bradycardia and tachycardia. Sinus bradycardia is a result of abnormally slow automaticity while bradycardia due to AV block is caused by abnormal conduction within the AV node or the intraventricular conduction system. The mechanisms generating tachycardia are shown in Figure 13.42.

Accelerated automaticity (Fig. 13.42a)

The normal mechanism of spontaneous cardiac rhythmicity is slow depolarization of the transmembrane voltage during diastole until the threshold potential is reached and the action potential of the pacemaker cells takes off. This mechanism may be accelerated by increasing the rate of diastolic depolarization or changing the threshold potential. For example, sympathetic stimulation releases epinephrine (adrenaline), which enhances automaticity. Abnormal automaticity can occur in virtually all cardiac tissues and may initiate arrhythmias. Such changes are thought to produce sinus tachycardia, escape rhythms and accelerated AV nodal (junctional) rhythms.

Triggered activity (Fig. 13.42b)

Myocardial damage can result in oscillations of the transmembrane potential at the end of the action potential. These oscillations, which are called 'after depolarizations', may reach threshold potential and produce an arrhythmia. If they occur before the transmembrane potential reaches its threshold (at the end of phase 3 of the action potential), they are called 'early after depolarizations' (E in Fig. 13.42b). When they develop after the transmembrane potential is completed, they are called 'delayed after depolarizations' (D in the figure).

The abnormal oscillations can be exaggerated by pacing, catecholamines, electrolyte disturbances, hypoxia, acidosis and some medications, which may then trigger arrhythmia. The atrial tachycardias produced by digoxin toxicity are due to triggered activity. The initiation of ventricular arrhythmia in the long QT syndrome (p. 727) may be caused by this mechanism.

Re-entry (or circus movements) (Fig. 13.42c)

The mechanism of re-entry occurs when a 'ring' of cardiac tissue surrounds an inexcitable core (e.g. in a region of scarred myocardium). Tachycardia is initiated if an ectopic beat finds one limb refractory (α), resulting in unidirectional block, and the other limb excitable. Provided conduction through the excitable limb (β) is slow enough, the other limb (α) will have recovered and will allow retrograde activation to complete the re-entry loop. If the time to conduct around the ring is longer than the recovery times (refractory periods) of the tissue within the ring, circus movement will be main-

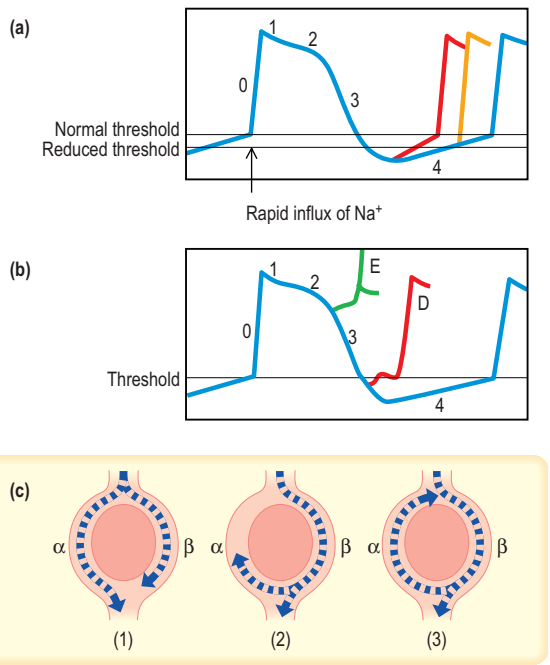


Fig. 13.42 **Mechanisms of arrhythmogenesis. (a and b) Action potentials** (i.e. the potential difference between intracellular and extracellular fluid) of ventricular myocardium after stimulation. **(a) Increased (accelerated) automaticity** due to reduced threshold potential or an increased slope of phase 4 depolarization (see p. 694). **(b) Triggered activity** due to early (E) or delayed (D) 'after depolarizations' reaching threshold potential. **(c) Mechanism of circus movement or re-entry.** In panel (1) the impulse passes down both limbs of the potential tachycardia circuit. In panel (2) the impulse is blocked in one pathway (α) but proceeds slowly down pathway β, returning along pathway α until it collides with refractory tissue. In panel (3) the impulse travels so slowly along pathway β that it can return along pathway α and complete the re-entry circuit, producing a circus movement tachycardia.

tained, producing a run of tachycardia. The majority of regular paroxysmal tachycardias are produced by this mechanism.

BRADYCARDIAS AND HEART BLOCK

Bradycardias may be due to failure of impulse formation (sinus bradycardia) or failure of impulse conduction from the atria to the ventricles (atrioventricular block).

Bradycardia

Sinus bradycardia

Sinus bradycardia is due to extrinsic factors influencing a relatively normal sinus node or due to intrinsic sinus node disease. The mechanism can be acute and reversible or chronic and degenerative. Common causes of sinus bradycardia include:

Extrinsic causes

- hypothermia, hypothyroidism, cholestatic jaundice and raised intracranial pressure
- drug therapy with beta-blockers, digitalis and other antiarrhythmic drugs
- neurally mediated syndromes (see below).

Intrinsic causes

- acute ischaemia and infarction of the sinus node (as a complication of acute myocardial infarction)
- chronic degenerative changes such as fibrosis of the atrium and sinus node (sick sinus syndrome).

Sick sinus syndrome or *sinoatrial disease* is usually caused by idiopathic fibrosis of the sinus node. Other causes of fibrosis such as ischaemic heart disease, cardiomyopathy or myocarditis can also cause the syndrome. Patients develop episodes of sinus bradycardia or sinus arrest (Fig. 13.41b) and commonly, owing to diffuse atrial disease, experience paroxysmal atrial tachyarrhythmias (tachy-brady syndrome).

Neurally mediated syndromes

Neurally mediated syndromes are due to a reflex (called Bezold–Jarisch) that may result in both bradycardia (sinus bradycardia, sinus arrest and AV block) and reflex peripheral vasodilatation. These syndromes usually present as syncope or pre-syncope (dizzy spells).

Carotid sinus syndrome occurs in the elderly and mainly results in bradycardia. Syncope occurs (see p. 689).

Neurocardiogenic (*vasovagal*) *syncope* (*syndrome*) usually presents in young adults but may present for the first time in elderly patients. It results from a variety of situations (physical and emotional) that affect the autonomic nervous system. The efferent output may be predominantly bradycardic, predominantly vasodilatory or mixed.

Postural orthostatic tachycardia syndrome (*POTS*) is a sudden and significant increase in heart rate associated with normal or mildly reduced blood pressure produced by standing. The underlying mechanism is a failure of the peripheral vasculature to appropriately constrict in response to orthostatic stress, which is compensated by an excessive increase in heart rate.

Many medications, such as antihypertensives, tricyclic antidepressants and neuroleptics can be the cause of syncope, particularly in the elderly. Careful dose titration and avoidance of combining two agents with potential to cause syncope help to prevent iatrogenic syncope.

Treatment

The management of sinus bradycardia is first to identify and if possible remove any extrinsic causes. Temporary pacing may be employed in patients with reversible causes until a normal sinus rate is restored and in patients with chronic degenerative conditions until a permanent pacemaker is implanted.

Chronic symptomatic sick sinus syndrome requires permanent pacing (DDD), with additional antiarrhythmic drugs (or ablation therapy) to manage any tachycardia element. Thromboembolism is common in tachy-brady syndrome and patients should be anticoagulated unless there is a contraindication.

Patients with carotid sinus hypersensitivity (asystole >3 s), especially if symptoms are reproduced by carotid sinus massage, and in whom life-threatening causes of syncope have been excluded, benefit from pacemaker implantation.

Treatment options in vasovagal attacks include avoidance, if possible, of situations known to cause syncope in a particular patient. Increased salt intake, compression of the lower legs with hose and drugs such as beta-blockers, alpha-agonists or myocardial negative inotropes (such as disopyramide) may be helpful. In selected patients with 'malignant' neurocardiogenic syncope (syncope associated with injuries) permanent pacemaker therapy is helpful. These patients benefit from dual chamber pacemakers with a feature called 'rate drop response' which, once activated, paces the heart at a fast rate for a set period of time in order to prevent syncope.

Heart block

Heart block or conduction block may occur at any level in the conducting system. Block in either the AV node or the His bundle results in AV block, whereas block lower in the conduction system produces bundle branch block.

Atrioventricular block

There are three forms:

First-degree AV block

This is simple prolongation of the PR interval to more than 0.22 s. Every atrial depolarization is followed by conduction to the ventricles but with delay (Fig. 13.43).

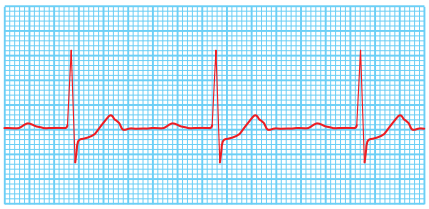


Fig. 13.43 **An ECG showing first-degree atrioventricular block** with a prolonged PR interval. In this trace coincidental ST depression is also present.

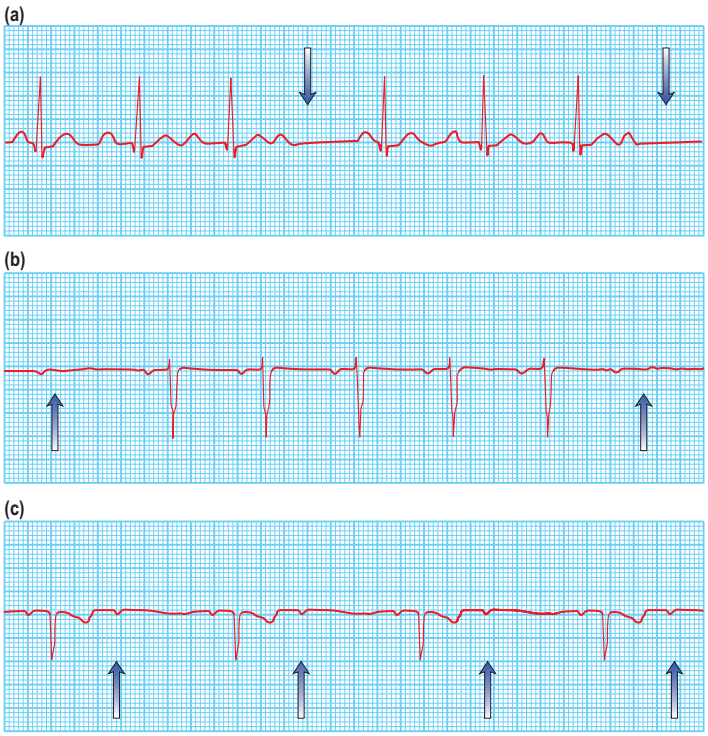


Fig. 13.44 **Three varieties of second-degree atrioventricular (AV) block. (a) Wenckebach (Mobitz type I) AV block.** The PR interval gradually prolongs until the P wave does not conduct to the ventricles (arrows). **(b) Mobitz type II AV block.** The P waves that do not conduct to the ventricles (arrows) are not preceded by gradual PR interval prolongation. **(c) Two P waves to each QRS complex.** The PR interval prior to the dropped P wave is always the same. It is not possible to define this type of AV block as type I or type II Mobitz block and it is, therefore, a third variety of second-degree AV block (arrows show P waves).

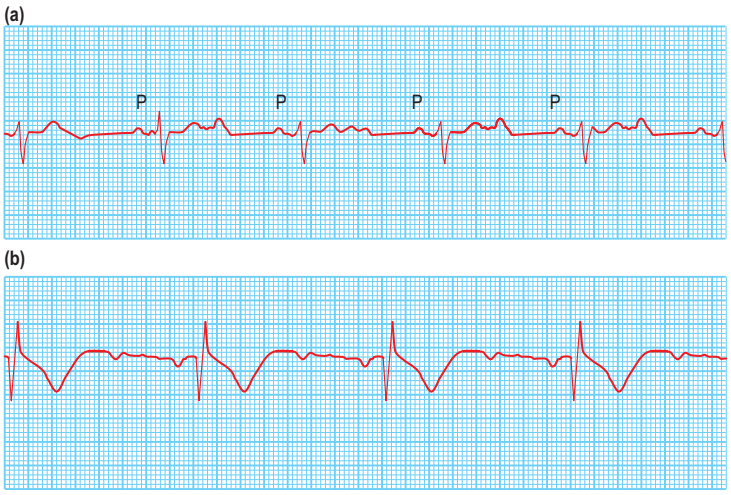


Fig. 13.45 **Two examples of complete heart block. (a) Congenital complete heart block.** The QRS complex is narrow (0.08 s) and the QRS rate is relatively rapid (52 b.p.m.). **(b) Acquired complete heart block.** The QRS complex is broad (0.13 s) and the QRS rate is relatively slow (38 b.p.m.).

Second-degree AV block

This occurs when some P waves conduct and others do not. There are several forms (Fig. 13.44):

- *Mobitz I block* (Wenckebach block phenomenon) is progressive PR interval prolongation until a P wave fails to conduct. The PR interval before the blocked P wave is much longer than the PR interval after the blocked P wave.
- *Mobitz II block* occurs when a dropped QRS complex is not preceded by progressive PR interval prolongation. Usually the QRS complex is wide (>0.12 s). Usually the QRS complex is wide (>0.12 s).
- *2:1 or 3:1* (*advanced*) *block* occurs when every second or third P wave conducts to the ventricles. This form of second-degree block is neither Mobitz I nor II.

Wenckebach AV block in general is due to block in the AV node, whereas Mobitz II block signifies block at an infra-nodal level such as the His bundle. The risk of progression to complete heart block is greater and the reliability of the resultant escape rhythm is less with Mobitz II block. Therefore pacing is usually indicated in Mobitz II block, whereas patients with Wenckebach AV block are usually monitored.

Acute myocardial infarction may produce *second-degree heart block*. In inferior myocardial infarction, close monitoring and transcutaneous temporary back-up pacing are all that is required. In anterior myocardial infarction, second-degree heart block is associated with a high risk of progression to complete heart block, and temporary pacing followed by permanent pacemaker implantation is usually indicated. 2:1 Heart block may either be due to block in the AV node or at an infra-nodal level. Management depends on the clinical setting in which it occurs.

Third-degree (complete) AV block

Complete heart block occurs when all atrial activity fails to conduct to the ventricles (Fig. 13.45). In patients with complete heart block the aetiology needs to be established

Table 13.8 Aetiology of complete heart block

Congenital
Autoimmune (e.g. maternal SLE)
Structural heart disease (e.g. transposition of the great vessels)

Idiopathic fibrosis
Lev's disease (progressive fibrosis of distal His–Purkinje system in elderly patients)
Lenegre's disease (proximal His–Purkinje fibrosis in younger patients)

Ischaemic heart disease
Acute myocardial infarct
Ischaemic cardiomyopathy

Non-ischaemic heart disease
Calcific aortic stenosis
Idiopathic dilated cardiomyopathy
Infiltrations (e.g. amyloidosis, sarcoidosis, neoplasia)

Cardiac surgery
e.g. following aortic valve replacement, CABG, VSD repair

Iatrogenic
Radiofrequency AV node ablation and pacemaker implantation

Drug-induced
e.g. digoxin, beta-blockers, non-dihydropyridine calcium-channel blockers, amiodarone

Infections
Endocarditis
Lyme disease
Chagas' disease

Rheumatic autoimmune disease
e.g. SLE, rheumatoid arthritis

Neuromuscular diseases
e.g. Duchenne muscular dystrophy

SLE, systemic lupus erythematosus; CABG, coronary artery bypass graft surgery; VSD, ventricular septal defect; AV, atrioventricular.

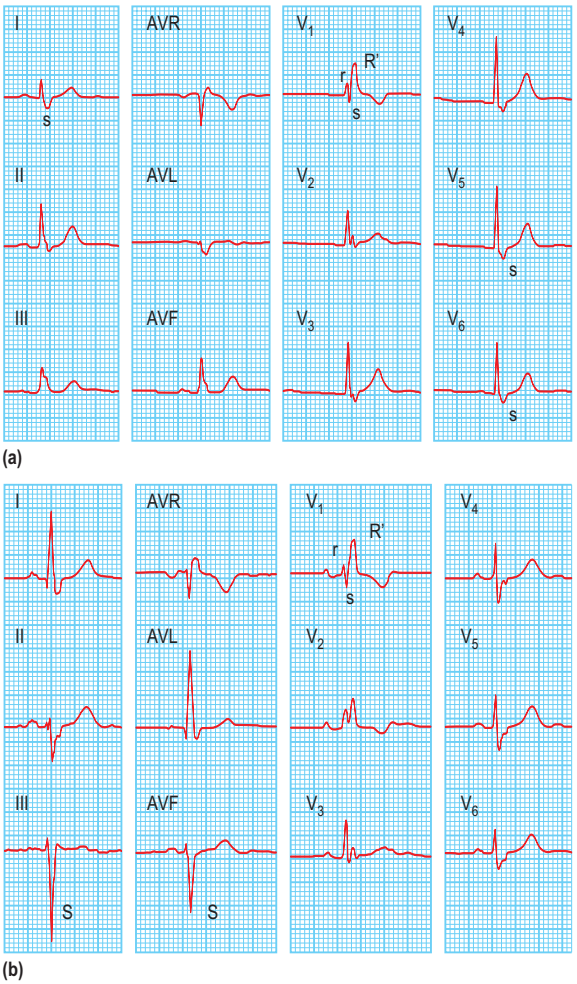


Fig. 13.46 **Right bundle branch block versus bifascicular block. (a) A 12-lead ECG showing right bundle branch block.** Note an rsR pattern with the tall R in lead V_1–V_2 and the broad S waves in leads I and V_5–V_6. **(b) Compare with an ECG showing bifascicular block.** In addition to right bundle branch block, note left axis deviation and deep S waves in leads III and AVF typical for left anterior hemiblock.

(Table 13.8). In this situation life is maintained by a spontaneous escape rhythm.

Narrow complex escape rhythm (<0.12 s QRS complex) implies that it originates in the His bundle and therefore that the region of block lies more proximally in the AV node. The escape rhythm occurs with an adequate rate (50–60 b.p.m.) and is relatively reliable.

Treatment depends on the aetiology. Recent-onset narrow-complex AV block due to transient causes may respond to intravenous atropine, but temporary pacing facilities should be available for the management of these patients. Chronic narrow-complex AV block requires permanent pacing (dual chamber, p. 714) if it is symptomatic or associated with heart disease. Pacing is also advocated for isolated, congenital AV block, even if asymptomatic.

Broad complex escape rhythm (>0.12 s) implies that the escape rhythm originates below the His bundle and therefore that the region of block lies more distally in the His–Purkinje system.

The resulting rhythm is slow (15–40 b.p.m.) and relatively unreliable. Dizziness and blackouts (*Stokes–Adams attacks*) often occur. In the elderly, it is usually caused by degenerative fibrosis and calcification of the distal conduction system (*Lev's disease*). In younger individuals, a proximal progressive cardiac conduction disease due to the inflammatory process is known as *Lenegre's syndrome*. Sodium channel abnormalities have been identified in both syndromes. Broad-complex AV block may also be caused by ischaemic heart disease, myocarditis or cardiomyopathy. Permanent pacemaker implantation (p. 713) is indicated, as pacing considerably reduces the mortality. Because ventricular arrhythmias are not uncommon, an implantable cardioverter–defibrillator (ICD) may be indicated in those with severe left ventricular dysfunction (>0.30 s).

Bundle branch block

The His bundle gives rise to the right and left bundle branches. The left bundle subdivides into the anterior and posterior divisions of the left bundle. Various conduction disturbances can occur.

Bundle branch conduction delay

This produces slight widening of the QRS complex (up to 0.11 s). It is known as incomplete bundle branch block.

Complete block of a bundle branch

This is associated with a wider QRS complex (0.12 s or more). The shape of the QRS depends on whether the right or the left bundle is blocked.

Right bundle branch block (Fig. 13.46a) produces late activation of the right ventricle. This is seen as deep S waves in

leads I and V_6 and as a tall late R wave in lead V_1 (late activation moving towards right- and away from left-sided leads).

Left bundle branch block (Fig. 13.47) produces the opposite – a deep S wave in lead V_1 and a tall late R wave in leads I and V_6. Because left bundle branch conduction is normally responsible for the initial ventricular activation, left bundle branch block also produces abnormal Q waves.

Hemiblock

Delay or block in the divisions of the left bundle branch produces a swing in the direction of depolarization (electrical axis) of the heart. When the anterior division is blocked (left anterior hemiblock), the left ventricle is activated from inferior to superior. This produces a superior and leftwards movement of the axis (left axis deviation). Delay or block in the postero-inferior division swings the QRS axis inferiorly to the right (right axis deviation).

Bifascicular block (Fig. 13.46b)

This is a combination of a block of any two of the following: the right bundle branch, the left antero-superior division and the left postero-inferior division. Block of the remaining fascicle will result in complete AV block.

Clinical features of heart blocks

Bundle branch blocks are usually asymptomatic. Right bundle branch block causes wide but physiological splitting of the second heart sound. Left bundle branch block may cause reverse splitting of the second sound. Patients with intraventricular conduction disturbances may complain of syncope. This is due to intermittent complete heart block or to ventricular tachyarrhythmias. ECG monitoring and electrophysiological studies are needed to determine the cause of syncope in these patients.

Causes

Right bundle branch block occurs as an isolated congenital anomaly or is associated with cardiac or pulmonary conditions (Table 13.9). Right bundle branch block alone does not alter the electrical axis of the heart. Axis deviations signify right ventricular hypertrophy (RV overload) or coexistent fascicular block. The combination of right bundle branch block with left axis deviation is associated with ostium primum atrial septal defects. Complete left bundle branch block is often associated with extensive left ventricular disease. The most common causes are listed in Table 13.10 and are similar to those of complete heart block.

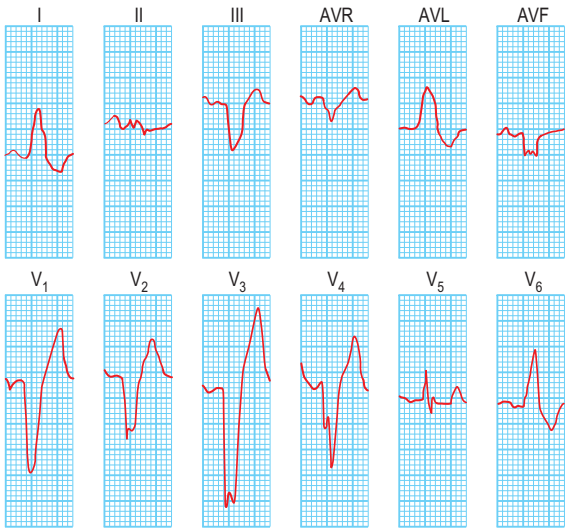


Fig. 13.47 **A 12-lead ECG showing left bundle branch block**. The QRS duration is greater than 0.12 s. Note the broad notched R waves with ST depression in leads I, AVL and V_6, and the broad QS waves in V_1–V_3.

Table 13.9 Causes of right bundle branch block

It is also a normal finding in 1% of young adults and 5% of elderly adults

Congenital heart disease
Atrial septal defect
Fallot's tetralogy
Pulmonary stenosis
Ventricular septal defect

Pulmonary disease
Cor pulmonale
Recurrent pulmonary embolism
Acute pulmonary embolism (transient)

Myocardial disease
Acute myocardial infarction
Cardiomyopathy
Conduction system fibrosis
Chagas' disease

Table 13.10 Causes of left bundle branch block

Left ventricular outflow obstruction	**Coronary artery disease**
Aortic stenosis	Acute myocardial infarction
Hypertension	Severe coronary disease (two- to three-vessel disease)

SUPRAVENTRICULAR TACHYCARDIAS

Supraventricular tachycardias (SVTs) arise from the atrium or the atrioventricular junction. Conduction is via the His–Purkinje system; therefore the QRS shape during tachycardia is usually similar to that seen in the same patient during baseline rhythm. A classification of supraventricular tachycardia is listed in Table 13.11. Some of these are discussed in more detail below.

Inappropriate sinus tachycardia

Inappropriate sinus tachycardia is a persistent increase in resting heart rate unrelated to or out of proportion with the level of physical or emotional stress. It is found predominantly in young women and is not uncommon in health professionals. Sinus tachycardia due to intrinsic sinus node abnormalities such as enhanced automaticity, or abnormal autonomic regulation of the heart with excess sympathetic and reduced parasympathetic input, is extremely rare.

In general, sinus tachycardia is a secondary phenomenon and the underlying causes need to be actively investigated. Depending on the clinical setting, acute causes include exercise, emotion, pain, fever, infection, acute heart failure, acute

Table 13.11 Causes of supraventricular tachycardia (SVT)

Tachycardia	ECG features	Comment
Sinus tachycardia	P wave morphology similar to sinus rhythm preceding QRS	Need to determine underlying cause
AV nodal re-entry tachycardia (AVNRT)	No visible P wave, or inverted P wave immediately before or after QRS complex	Commonest cause of palpitations in patients with normal hearts
AV reciprocating tachycardia (AVRT) complexes	P wave visible between QRS and T wave	Due to an accessory pathway. If pathway conducts in both directions, ECG during sinus rhythm may be pre-excited
Atrial fibrillation	Irregularly irregular RR intervals and absence of organized atrial activity	Commonest tachycardia in patients over 65 years
Atrial flutter	Visible flutter waves at 300/min (saw-tooth appearance) usually with a 2:1 AV conduction	Suspect in any patient with regular SVT at 150/min
Atrial tachycardia	Organized atrial activity with P wave morphology different from sinus rhythm preceding QRS	Usually occurs in patients with structural heart disease
Multifocal atrial tachycardia	Multiple P wave morphologies (≥3) and irregular RR intervals	Rare arrhythmia; most commonly associated with significant chronic lung disease
Accelerated junctional tachycardia	ECG similar to AVNRT	Rare in adults

pulmonary embolism and hypovolaemia. Chronic causes include pregnancy, anaemia, hyperthyroidism and catecholamine excess. The underlying cause should be found and treated, rather than treating the compensatory physiological response. If necessary, beta-blockers may be used to slow the sinus rate, e.g. in hyperthyroidism (p. 987).

Atrioventricular junctional tachycardias

AV nodal re-entry and AV re-entry tachycardias are usually referred to as paroxysmal SVTs and are often seen in young patients with no or little structural heart disease, although congenital heart abnormalities (e.g. Ebstein's anomaly, atrial septal defect, Fallot's tetralogy) can coexist in a small proportion of patients with these arrhythmias. The first presentation is commonly between ages 12 and 30, and the prevalence is approximately 2.5 per 1000.

In these tachycardias the AV node is an essential component of the re-entry circuit.

Atrioventricular nodal re-entry tachycardia (AVNRT)

This tachycardia is twice as common in women. Clinically, the tachycardia often strikes suddenly without obvious provocation, but exertion, coffee, tea and alcohol may aggravate or induce the arrhythmia. An attack may stop spontaneously or may continue indefinitely until medical intervention.

In AVNRT, there are two functionally and anatomically different pathways predominantly within the AV node: one is characterized by a short effective refractory period and slow conduction, and the other has a longer effective refractory period and conducts faster. In sinus rhythm, the atrial impulse that depolarizes the ventricles usually conducts through the fast pathway. If the atrial impulse (e.g. an atrial premature beat) occurs early when the fast pathway is still refractory, the slow pathway takes over in propagating the atrial impulse to the ventricles. It then travels back through the fast pathway which has already recovered its excitability, thus initiating the most common 'slow–fast', or typical, AVNRT.

The rhythm is recognized on ECG by normal regular QRS complexes, usually at a rate of 140–240 per minute (Fig. 13.48a). Sometimes the QRS complexes will show typical bundle branch block. P waves are either not visible or are seen immediately before or after the QRS complex because of simultaneous atrial and ventricular activation. Less commonly observed (5–10%) is tachycardia when the atrial impulse conducts anterogradely through the fast pathway and returns through the slow pathway, producing a long RP′ interval ('fast–slow', or long RP′ tachycardia).

Atrioventricular reciprocating tachycardia (AVRT)

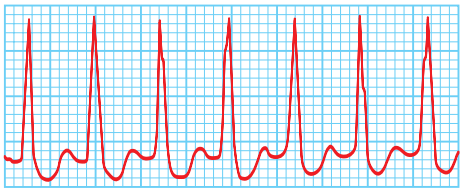
In AVRT there is a large circuit comprising the AV node, the His bundle, the ventricle and an abnormal connection from the ventricle back to the atrium. This abnormal connection consists of myocardial fibres that span the AV groove; it is called an accessory pathway or bypass tract. Bypass tracts result from incomplete separation of the atria and the ventricles during fetal development.

In contrast to AVNRT, this tachycardia is due to a macro-reentry circuit and each part of the circuit is activated sequentially. As a result atrial activation occurs after ventricular activation and the P wave is usually clearly seen between the QRS and T wave (Fig. 13.48b).

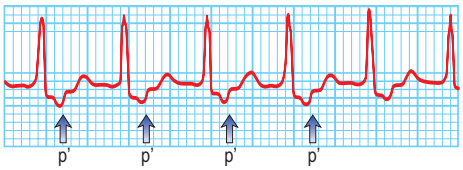
Accessory pathways are most commonly situated on the left but may occur anywhere around the AV groove. The most common accessory pathways, known as *Kent bundles*, are in the free wall or septum. In about 10% of cases multiple pathways occur. *Mahaim fibres* are atrio-fascicular or nodo-fascicular fibres entering the ventricular myocardium in the region of the right bundle branch. Accessory pathways that conduct from the ventricles to the atria only are not visible on the surface ECG during sinus rhythm and are therefore 'concealed'. Accessory pathways that conduct bidirectionally usually are manifest on the surface ECG. If the accessory pathway conducts from the atrium to the ventricle during sinus rhythm, the electrical impulse can conduct quickly over this abnormal connection to depolarize part of the ventricles abnormally (pre-excitation). A pre-excited ECG is

Ask the authors
www.studentconsult.com

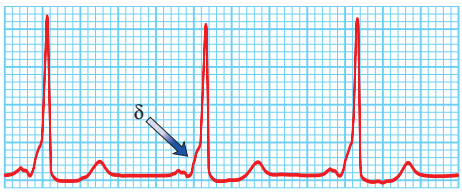
Fig. 13.48 **Atrioventricular junctional tachycardia.**



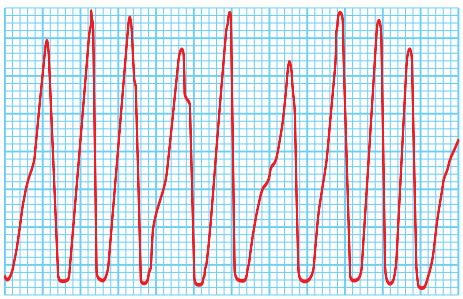
(a) *Atrioventricular nodal re-entry* tachycardia. The QRS complexes are narrow and the P waves cannot be seen.



(b) *Atrioventricular re-entry* tachycardia (Wolff–Parkinson–White syndrome). The tachycardia P waves (arrows) are clearly seen after narrow QRS complexes.



(c) An electrocardiogram taken in a patient with *Wolff–Parkinson–White (WPW) syndrome* during sinus rhythm. Note the short PR interval and the δ wave (arrow).



(d) *Atrial fibrillation in the WPW syndrome*. Note tachycardia with broad QRS complexes with fast and irregular ventricular rate.

characterized by a short PR interval and a wide QRS complex that begins as a slurred part known as the δ wave (Fig. 13.48c). Patients with a history of palpitations and a pre-excited ECG have a syndrome known as Wolff–Parkinson–White (WPW) syndrome.

During AVRT the AV node and ventricles are activated normally (orthodromically), resulting usually in a narrow QRS complex. Less commonly, the tachycardia circuit can be reversed, with activation of the ventricles via the accessory pathway and atrial activation via retrograde conduction through the AV node (antidromic AVRT). This results in a broad complex tachycardia. These patients are also prone to atrial fibrillation.

During atrial fibrillation, the ventricles may be depolarized by impulses travelling over both the abnormal and the normal pathways. This results in pre-excited atrial fibrillation, a characteristic tachycardia that is characterized by irregularly irregular broad QRS complexes (Fig. 13.48d). If an accessory pathway has a short antegrade effective refractory period (<250 ms), it may conduct to the ventricles at an extremely high rate and may cause ventricular fibrillation. The incidence of sudden death is 0.15–0.39% per patient-year and it may be a first manifestation of the disease in younger individuals. Verapamil and digoxin may allow a higher rate of conduction over the abnormal pathway and precipitate ventricular fibrillation. Therefore, neither verapamil nor digoxin should be used to treat atrial fibrillation associated with the WPW syndrome.

Symptoms

The leading symptom of most SVTs, in particular AV node re-entry and AV re-entry tachycardias, is rapid regular palpitations, usually with abrupt onset and sudden termination, which can occur spontaneously or be precipitated by simple movements. A common feature is termination by Valsalva manoeuvres. In younger individuals with no structural heart disease, the rapid heart rate can be the main pathological finding.

Irregular palpitations may be due to atrial premature beats, atrial flutter with varying AV conduction block, atrial fibrillation or multifocal atrial tachycardia. In patients with depressed ventricular function, uncontrolled atrial fibrillation can reduce cardiac output and cause hypotension and congestive heart failure.

Other symptoms may include anxiety, dizziness, dyspnoea, neck pulsation, central chest pain and weakness. Polyuria may occur because of release of atrial natriuretic peptide in response to increased atrial pressures during the tachycardia, especially during AVNRT and atrial fibrillation. Prominent jugular venous pulsations due to atrial contractions against closed atrioventricular valves may be observed during AVNRT.

Syncope has been reported in 10–15% of patients, usually just after initiation of the arrhythmia or in association with a prolonged pause following its termination. It is more common if the patient is standing. However, in older patients with concomitant heart disease, such as aortic stenosis, hypertrophic cardiomyopathy and cerebrovascular disease, significant hypotension and syncope may result from moderately fast ventricular rates.

Acute management

In an emergency, distinguishing between AVNRT and AVRT may be difficult, but it is usually not critical as both tachycardias respond to the same treatment. Patients presenting with SVTs and haemodynamic instability (e.g. hypotension, pulmonary oedema) require emergency cardioversion. If the patient is haemodynamically stable, vagal manoeuvres, including right carotid massage (Practical box 13.3), Valsalva manoeuvre (Practical box 13.4) and facial immersion in cold water can be successfully employed.

Of these techniques, the Valsalva manoeuvre is the best and is often easier for the patient to perform successfully. It should be undertaken when the patient is resting in the supine position (thus avoiding elevated background sympathetic tone). Several seconds after the release of strain, the resulting intense vagal effect may terminate AVNRT or AVRT or may produce sufficient AV block to reveal an underlying atrial tachyarrhythmia.

If physical manoeuvres have not been successful, intravenous adenosine (initially 6 mg by i.v. push, followed by 12 mg if needed) should be tried. This is a very short-acting (half-life <10 s) naturally occurring purine nucleoside that causes complete heart block for a fraction of a second fol-

Practical Box 13.4 Valsalva manoeuvre

- Valsalva manoeuvre is an abrupt voluntary increase in intrathoracic and intra-abdominal pressures by straining.
- Provide continuous electrocardiographic monitoring.
- Patient is in supine position.
- Patient should not take deep inspiration before straining.
- Ideally, the patient blows into the mouthpiece of a manometer against a pressure of 30–40 mmHg for 15 seconds.
- Alternatively, the patient strains for 15 seconds while breath-holding.
- Transient acceleration of tachycardia usually occurs during the strain phase as a result of sympathetic excess.
- On release of strain, the rate of tachycardia slows because of the compensatory increase in vagal tone (baroreceptor reflex), and it may be terminated in about 50% of patients.
- Termination of tachycardia may be followed by pauses and transient ventricular ectopics.

lowing i.v. administration. It is highly effective at terminating AVNRT and AVRT or unmasking underlying atrial activity. It rarely affects ventricular tachycardia. The side-effects of adenosine are very brief but include:

- bronchospasm
- flushing
- chest pain
- heaviness of the limbs
- sense of impending doom.

It is contraindicated in patients with a history of asthma. In some patients, adenosine can induce atrial fibrillation.

An alternative treatment is verapamil 5–10 mg i.v. over 5–10 minutes, i.v. diltiazem, or beta-blockers (esmolol, propranolol, metoprolol). Verapamil (or diltiazem) must not be given after beta-blockers *or* if the tachycardia presents with broad (>0.12 s) QRS complexes.

Long-term management

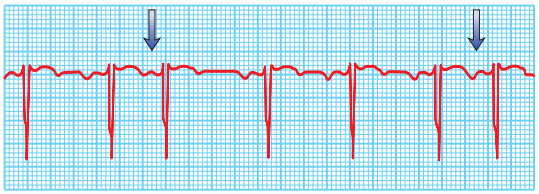
Patients with suspected cardiac arrhythmias should always be referred to the cardiologist for electrophysiological evaluation and long-term management, as both pharmacological and non-pharmacological alternatives, including ablation of an accessory pathway, are readily available. Verapamil, diltiazem and beta-blockers have proven efficacy in 60–80% of patients. Sodium-channel blockers (flecainide and propafenone), potassium repolarization current blockers (sotalol, dofetilide, azimilide), and the multichannel blocker amiodarone may also prevent the occurrence of tachycardia.

Refinement of catheter ablation techniques has rendered many AV junctional tachycardias entirely curable. Modification of the slow pathway is successful in 96% of patients with AVNRT, although a 1% risk of AV block is present. In AVRT, the target for catheter ablation is the accessory pathway(s). The success rate for ablation of a single accessory pathway is approximately 95%, with a recurrence rate of 5%, requiring a repeat procedure.

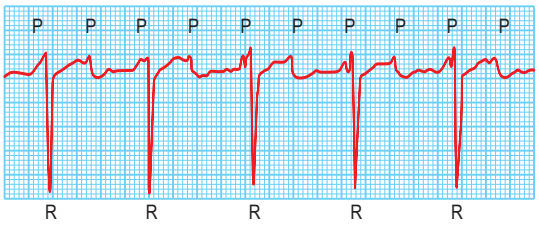
Atrial tachyarrhythmias

Atrial tachyarrhythmias including atrial fibrillation, atrial flutter, atrial tachycardia and atrial ectopic beats all arise from the atrial myocardium (Fig. 13.49a–d). They share common aetiologies, which are listed in Table 13.12.

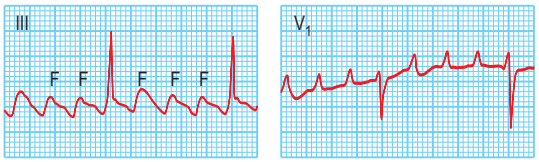
Fig. 13.49 **ECGs of a variety of atrial arrhythmias.**



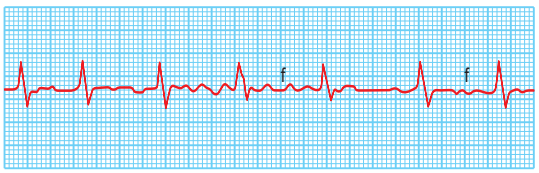
(a) Atrial premature beats (arrows). The premature P wave is different from the sinus P wave and conducts to the ventricle with a slightly prolonged PR interval.



(b) Atrial tachycardia with second-degree atrioventricular block. Note the fast atrial (P wave) rate of 150/min and the slower ventricular (R wave) rate of 75/min.



(c) Atrial flutter. Some flutter waves are marked with an F. In this case the flutter frequency is 240/min. Every fourth flutter wave is transmitted to the ventricles and the ventricular rate is therefore 60/min.



(d) Atrial fibrillation. Note absolute rhythm irregularity and baseline undulations (f waves).

Table 13.12 Causes of atrial tachyarrhythmias

Cardiac	Non-cardiac
Hypertension	Thyrotoxicosis
Congestive heart failure	Phaeochromocytoma
Coronary artery disease and myocardial infarction	Acute and chronic pulmonary disease (pneumonia, chronic obstructive pulmonary disease)
Valvular heart disease	Pulmonary vascular disease (pulmonary embolism)
Cardiomyopathy: dilated, hypertrophic	Electrolyte disturbances (hypokalaemia)
Myocarditis and pericarditis	Increased sympathetic tone (exercise, adrenergically mediated arrhythmia)
Wolff–Parkinson–White syndrome	Increased parasympathetic tone (vagally induced and postprandial arrhythmia)
Sick sinus syndrome	Alcohol abuse ('holiday heart' and long-term use)
Cardiac tumours	Caffeine, smoking, recreational drug use, e.g. cannabis
Cardiac surgery	Myotonic dystrophy type 1
Familial tachyarrhythmia (e.g. lone atrial fibrillation)	Chagas' disease
Genetic predisposition	

Atrial fibrillation

This is a common arrhythmia, occurring in 5–10% of patients over 65 years of age. It also occurs, particularly in a paroxysmal form, in younger patients. Any condition resulting in raised atrial pressure, increased atrial muscle mass, atrial fibrosis, or inflammation and infiltration of the atrium, may cause atrial fibrillation. There are also many systemic causes of atrial fibrillation (Table 13.12).

Although rheumatic heart disease, alcohol intoxication and thyrotoxicosis are the 'classic' causes of atrial fibrillation, hypertension and heart failure are the most common causes. Hyperthyroidism may provoke atrial fibrillation, sometimes as virtually the only feature of the disease, and thyroid function tests are mandatory in any patient with unaccounted atrial fibrillation. Atrial fibrillation occurs in one-third of patients after coronary bypass surgery and in more than half of those undergoing valvular surgery. It usually manifests during the first 4 days and is associated with increased morbidity and mortality, largely due to stroke and circulatory failure, and longer hospital stay.

In some patients no cause can be found, and this group is labelled as 'lone' atrial fibrillation. The pathogenesis of 'lone', or 'idiopathic', atrial fibrillation is unknown but genetic predisposition or even specific genetically predetermined forms of the arrhythmia have been proposed. Thirty to 40% of those with AF, especially those who present at a young age, have at least one parent with AF and genes associated with the sodium channel, the potassium channel and gap junction proteins have been implicated. Gene defects linked to chromosomes 10, 6, 5 and 4 have been associated with familial AF.

Atrial fibrillation is maintained by continuous, rapid (300–600 per minute) activation of the atria by multiple meandering re-entry wavelets, often driven by rapidly depolarizing automatic foci, located predominantly within the pulmonary veins. The atria respond electrically at this rate but there is no coordinated mechanical action and only a proportion of the impulses are conducted to the ventricles. The ventricular response depends on the rate and regularity of atrial activity, particularly at the entry to the AV node, the refractory properties of the AV node itself, and the balance between sympathetic and parasympathetic tone.

Symptoms and signs

Symptoms attributable to atrial fibrillation are highly variable. In some patients (about 30%) it is an incidental finding, whilst others attend hospital as an emergency following the onset of atrial fibrillation. Most patients experience some deterioration of exercise capacity or well-being, but this may only be appreciated once sinus rhythm is restored. When caused by rheumatic mitral stenosis, the onset of atrial fibrillation results in considerable worsening of cardiac failure.

The patient has a very irregular pulse, as opposed to a basically regular pulse with an occasional irregularity (e.g. extrasystoles) or recurring irregular patterns (e.g. Wenckebach block). The irregular nature of the pulse in atrial fibrillation is maintained during exercise.

The ECG shows fine oscillations of the baseline (so-called fibrillation or f waves) and no clear P waves. The QRS rhythm is rapid and irregular. Untreated, the ventricular rate is usually 120–180 per minute, but it slows with treatment.

The clinical classification of atrial fibrillation includes first detected, paroxysmal, persistent, and permanent forms of the arrhythmia and is essential for the decision-making between rhythm restoration and rate control. For example, atrial fibrillation may be asymptomatic and the 'first detected episode' should not be regarded as necessarily the true onset.

Management

When atrial fibrillation is due to an acute precipitating event such as alcohol toxicity, chest infection or hyperthyroidism, the provoking cause should be treated. Strategies for the acute management of AF are ventricular rate control or cardioversion (± anticoagulation). Ventricular rate control is achieved by drugs which block the AV node (see below), while cardioversion may be achieved electrically by DC shock (p. 712) or medically either by intravenous infusion of an antiarrhythmic drug such as a class Ic or a class III agent or by taking an oral agent previously tested in hospital and found to be safe in a particular patient ('pill-in-pocket' approach).

The choice depends upon:

- how well the arrhythmia is tolerated (is cardioversion urgent?)
- whether anticoagulation is required before considering elective cardioversion
- whether spontaneous cardioversion is likely (previous history? reversible cause?).

Conversion to sinus rhythm can be achieved by electrical DC cardioversion (p. 712) in about 80% of patients. Biphasic waveform defibrillation is more effective than conventional (monophasic) defibrillation, and biphasic defibrillators are becoming standard. To minimize the risk of thromboembolism associated with cardioversion patients are fully anticoagulated (INR 2.0–3.0) with warfarin for 3 weeks before cardioversion (unless atrial fibrillation is of less than 1–2 days' duration) and at least 4 weeks after the procedure. The patient is then assessed for the necessity for long-term anticoagulation. If cardioversion is urgent, transoesophageal echocardiography may be used to document the presence or absence of atrial thrombus as a guide to possibly avoiding pre-cardioversion anticoagulation.

Two strategies are available for the long-term management of atrial fibrillation:

- 'rate control' (AV nodal slowing agents *plus warfarin*)
- 'rhythm control' (antiarrhythmic drugs plus DC cardioversion *plus warfarin*).

Major randomized studies in patients predominantly over the age of 65 years (AFFIRM) or in patients with heart failure (AF-CHF) have shown that there is no net mortality or symptom benefit to be gained from one strategy compared with the other. Which strategy to adopt needs to be assessed for each individual patient. Factors to consider include the likelihood of maintaining sinus rhythm and the safety/tolerability of antiarrhythmic drugs in a particular patient.

Rhythm control is advocated for younger, symptomatic and physically active patients. Recurrent paroxysms may be prevented by oral medication. In general, class Ic agents are employed in patients with no significant heart disease and class III agents are preferred in patients with significant structural heart disease. Catheter ablation techniques such as pulmonary vein isolation can be used in patients who do not respond to antiarrhythmic drugs.

Rate control as a primary strategy is appropriate in patients who:

- have the permanent form of the arrhythmia associated with symptoms which can be further improved by slowing heart rate or are older than 65 years with recurrent atrial tachyarrhythmias (primary 'accepted' atrial fibrillation)
- have persistent tachyarrhythmias and have failed cardioversion(s) and serial prophylactic antiarrhythmic drug therapy and in whom the risk/benefit ratio from using specific antiarrhythmic agents is shifted towards increased risk.

Rate control is usually achieved by a combination of *digoxin, beta-blockers* or *calcium-channel blockers* (*verapamil* or *diltiazem*). Digoxin monotherapy may be sufficient for elderly non-ambulant patients. In younger patients the effect of catecholamines easily overwhelms the vagotonic effect of digoxin and additional AV nodal slowing agents are needed. The ventricular rate response is generally considered controlled if the heart rate is between 60 and 80 b.p.m at rest and 90 and 115 b.p.m. during moderate exercise. To assess the adequacy of rate control, an ECG rhythm strip may be sufficient in an elderly patient but ambulatory 24-hour Holter monitoring and an exercise stress test (treadmill) are needed in younger individuals. Older patients with poor rate control despite optimal medical therapy should be considered for AV node ablation and pacemaker implantation (*'ablate and pace' strategy*). These patients usually experience a marked symptomatic improvement but because of the ongoing risk of thromboembolism require life-long *anticoagulation*.

Anticoagulation (target INR 2.0–3.0)

This is indicated in patients with atrial fibrillation and one of the following major or two of the following moderate risk factors:

Major risk factors

- prosthetic heart valve
- rheumatic mitral valve disease
- prior history of CVA/TIA.

Moderate risk factors

- age ≥ 75 years
- congestive heart failure
- hypertension
- diabetes mellitus.

In addition, anticoagulation with warfarin can be used for other less well validated features and include: age 65–74 years, coronary artery disease, history of thyrotoxicosis, female gender and echocardiographic evidence of left atrial enlargement, LV dysfunction of any cause, atrial thrombus or a reduced atrial appendage emptying velocity.

Aspirin is also sometimes required depending on the scoring system used.

A popular scoring system known as CHADS2 is often used to determine the need for anticoagulation. CHADS is an acronym standing for Congestive heart failure, Hypertension, Age ≥ 75, Diabetes mellitus and previous Stroke or TIA. Each factor scores 1 except previous stroke or TIA which scores 2. A total score of 0 implies no need for anticoagulation and with a score of 2 warfarin is needed. For a score of 1 warfarin or aspirin may be used depending on patient preference and the presence or absence of the additional less validated risk factors (see above).

For young patients (<65 years) with no significant heart disease (lone atrial fibrillation) aspirin 75–300 mg daily is generally prescribed. New oral anticoagulants, such as dabigatran, apixaban and rivaroxaban, which are easier to use than warfarin, are currently under trial.

Atrial flutter

Atrial flutter is often associated with atrial fibrillation and often requires a similar initial therapeutic approach. Atrial flutter is usually an organized atrial rhythm with an atrial rate typically between 250 and 350 b.p.m. Typical, or isthmus-dependent, atrial flutter involves a macro re-entrant right atrial circuit around the tricuspid annulus. The wavefront circulates down the lateral wall of the right atrium, through the Eustachian ridge between the tricuspid annulus and the inferior vena cava, and up the interatrial septum, giving rise to the most frequent pattern, referred to as counter-clockwise flutter. Re-entry can also occur in the opposite direction (clockwise or reverse flutter).

The ECG shows regular sawtooth-like atrial flutter waves (F waves) between QRS complexes (Fig. 13.49c). In typical counter-clockwise atrial flutter, the F waves are negative in the inferior leads and positive in leads V_1 and V_2. In clockwise atrial flutter, the deflection of the F waves is the opposite. If F waves are not clearly visible, it is worth trying to reveal them by slowing AV conduction by carotid sinus massage or by the administration of AV nodal blocking drugs such as adenosine or verapamil.

Symptoms are largely related to the degree of AV block. Most often, every second flutter beat conducts, giving a ventricular rate of 150 b.p.m. Occasionally, every beat conducts, producing a heart rate of 300 b.p.m. More often, especially when patients are receiving treatment, AV conduction block reduces the heart rate to approximately 75 b.p.m.

Management

Treatment of a symptomatic acute paroxysm is electrical *cardioversion*. Patients who have been in atrial flutter more than 1–2 days should be treated in a similar manner to patients with atrial fibrillation and *anticoagulated* for 3 weeks prior to cardioversion.

Recurrent paroxysms may be prevented by oral medication. In general, *class Ic* agents are employed in patients with no significant heart disease and *class III* agents are preferred in patients with significant structural heart disease. *AV nodal blocking agents* may be used to control the ventricular rate if the arrhythmia persists. However, the treatment of choice for patients with recurrent atrial flutter is radiofrequency catheter *ablation* (p. 731), which permanently interrupts re-entry by creating a line of conduction block within the isthmus between the inferior vena cava and the tricuspid valve ring. This technique offers patients whose only arrhythmia is typical atrial flutter an almost certain chance of a cure.

Atrial tachycardia

This is an uncommon arrhythmia. Its prevalence is believed to be less than 1% in patients with arrhythmias. It is usually associated with structural heart disease but in many cases it is referred to as idiopathic. Macro re-entrant tachycardia often occurs after surgery for congenital heart disease.

Atrial tachycardia with block is often a result of digitalis poisoning.

The mechanisms of atrial tachycardia are attributed to enhanced automaticity, triggered activity or intra-atrial re-entry. Atrial re-entry tachycardia is usually relatively slow (125–150 b.p.m.) and can be initiated and terminated by atrial premature beats. The P′P′ intervals are regular. The PR interval depends on the rate of tachycardia and is longer than in sinus rhythm at the same rate.

Automatic tachycardia usually presents with higher rates (125–250 b.p.m.) and is often characterized by a progressive increase in the atrial rate with onset of the tachycardia ('warm-up') and progressive decrease prior to termination ('cool-down'). Atrial tachycardia is typically caused by a focus which is frequently located along the crista terminalis in the right atrium, adjacent to a pulmonary vein in the left atrium, or around one of the atrial appendages. Automatic atrial tachycardia may also present as an incessant variety leading to tachycardia-induced cardiomyopathy. Short runs of atrial tachycardia may provoke more sustained episodes of atrial fibrillation.

Figure 13.49b demonstrates an atrial tachycardia at an atrial rate of 150 per minute. The P waves are abnormally shaped and occur in front of the QRS complexes. Carotid sinus massage may increase AV block during tachycardia, thereby facilitating the diagnosis, but does not usually terminate the arrhythmia. Treatment options include *cardioversion*, *antiarrhythmic drug therapy* to maintain sinus rhythm, *AV nodal slowing agents* to control rate and, in selected cases, radiofrequency catheter *ablation*.

Atrial ectopic beats

These often cause no symptoms, although they may be sensed as an irregularity or heaviness of the heart beat. On the ECG they appear as early and abnormal P waves, and are usually, but not always, followed by normal QRS complexes (Fig. 13.49a). Treatment is not normally required unless the ectopic beats provoke more significant arrhythmias, when *beta-blockers* may be effective.

VENTRICULAR TACHYARRHYTHMIAS

Ventricular tachyarrhythmias can be considered under the following headings:

- life-threatening ventricular tachyarrhythmias
- torsades de pointes
- normal heart ventricular tachycardia
- non-sustained ventricular tachycardia
- ventricular premature beats.

Life-threatening ventricular tachyarrhythmias

Sustained ventricular tachycardia and ventricular fibrillation with haemodynamic instability (e.g. syncope, hypotension) are life-threatening ventricular tachyarrhythmias.

Sustained ventricular tachycardia

Sustained ventricular tachycardia (>30 s) often results in *pre-syncope* (dizziness), *syncope*, *hypotension* and *cardiac arrest*, although it may be remarkably well tolerated in some patients. Examination reveals a pulse rate typically between 120 and 220 b.p.m. Usually there are clinical signs of atrioventricular dissociation (i.e. intermittent cannon 'a' waves in the neck, p. 691) and variable intensity of the first heart sound).

The ECG shows a rapid ventricular rhythm with broad (often 0.14 s or more), abnormal QRS complexes. AV dissociation may result in visible P waves which appear to march through the tachycardia, capture beats (intermittent narrow QRS complex owing to normal ventricular activation via the AV node and conducting system) and fusion beats (intermediate between ventricular tachycardia beat and capture beat).

Supraventricular tachycardia with bundle branch block may resemble ventricular tachycardia on the ECG. However, if a broad complex tachycardia is due to SVT with either right or left bundle branch block, then the QRS morphology should resemble a typical RBBB or LBBB pattern (Figs 13.46a and 13.47). Other ECG criteria to differentiate VT from SVT with aberrancy are indicated in Table 13.13. Eighty per cent of all broad complex tachycardias are due to ventricular tachycardia, and the proportion is even higher in patients with structural heart disease. Therefore, in all cases of doubt, ventricular tachycardia should be diagnosed.

Treatment may be urgent, depending on the haemodynamic situation. If the patient is haemodynamically compromised (e.g. hypotensive or pulmonary oedema), emergency DC cardioversion may be required. On the other hand, if the blood pressure and cardiac output are well maintained, intravenous therapy with class I drugs or amiodarone is usually used. First-line drug treatment consists of lidocaine (50–100 mg i.v. over 5 minutes) followed by a lidocaine infusion (2–4 mg i.v. per minute). Amiodarone is given as a dilute intravenous infusion (5 mg/kg over 1 hour (loading dose)) followed by a maintenance infusion of 1200 mg over 24 hours. DC cardioversion is necessary if medical therapy is unsuccessful.

Ventricular fibrillation

This is very rapid and irregular ventricular activation with no mechanical effect. The patient is pulseless and becomes rapidly unconscious, and respiration ceases (cardiac arrest). The ECG shows shapeless, rapid oscillations and there is no hint of organized complexes (Fig. 13.50). It is usually provoked by a ventricular ectopic beat. Ventricular fibrillation rarely reverses spontaneously. The only effective treatment is electrical defibrillation. Basic and advanced cardiac life support is needed (p. 709).

If the attack of ventricular fibrillation occurs during the first day or two of an acute myocardial infarction, it is probable that prophylactic therapy will be unnecessary. If the

Table 13.13 ECG distinction between supraventricular tachycardia (SVT) with bundle branch block and ventricular tachycardia (VT)

VT is more likely than SVT with bundle branch block where there is:

- a very broad QRS (>0.14 s)
- atrioventricular dissociation
- a bifid, upright QRS with a taller first peak in V_1
- a deep S wave in V_6
- a concordant (same polarity) QRS direction in all chest leads (V_1–V_6)

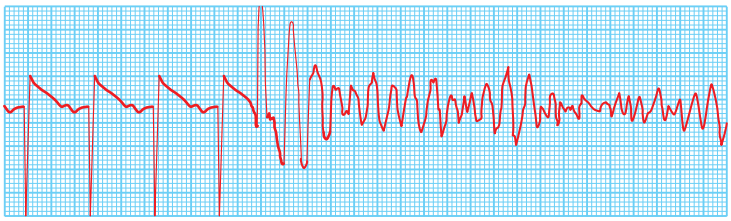


Fig. 13.50 **Ventricular fibrillation.** Four beats of sinus rhythm followed by a ventricular ectopic beat that initiates ventricular fibrillation. The ST segment during sinus rhythm is elevated owing to acute myocardial infarction in this case.

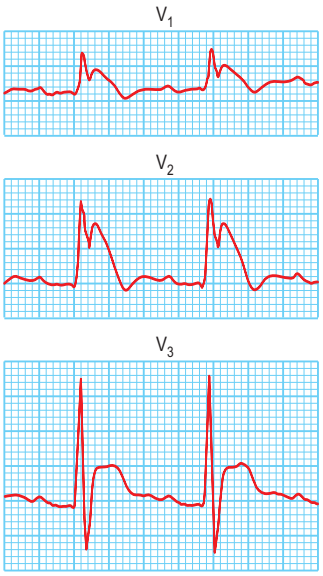


Fig. 13.51 **The Brugada ECG.** ST elevation in V_1–V_3 with right bundle branch block.

ventricular fibrillation was not related to an acute infarction, the long-term risk of recurrent cardiac arrest and sudden death is high.

Survivors of these ventricular tachyarrhythmias are, in the absence of an identifiable reversible cause (e.g. acute myocardial infarction, severe metabolic disturbance), at high risk of sudden death. Implantable cardioverter–defibrillators (ICDs) are first-line therapy in the management of these patients (p. 732).

The Brugada syndrome

This inheritable condition accounts for part of a group of patients with idiopathic ventricular fibrillation who have no evidence of causative structural cardiac disease. It is more common in young male adults and in South East Asia. The diagnosis is made by identifying the classic ECG changes that may be present spontaneously or be provoked by the administration of a class I antiarrhythmic (flecainide or ajmaline): right bundle branch block with coved ST elevation in leads V_1–V_3 (Fig. 13.51).

In 20% of cases it is a monogenic inheritable condition associated with loss of sodium channel function due to a mutation in the *SCN5A gene*. Recently other mutations in the glycerol-3-phosphate dehydrogenase-1-like gene (*GPD1LL*-type) and genes related to calcium channel subunits (*CACNA1C* and *CACNB2*) have also been implicated in the genesis of this syndrome. It can present with sudden death during sleep, resuscitated cardiac arrest and syncope, or the patient may be asymptomatic and diagnosed incidentally or during familial assessment. There is a high risk of sudden death, particularly in the symptomatic patient or those with spontaneous ECG changes. The only successful treatment is an ICD. Beta-blockade is not helpful and may be harmful in this syndrome.

Table 13.14 Causes of long QT syndrome

Congenital	Acquired
Jervell–Lange–Nielsen (autosomal recessive) Romano–Ward (autosomal dominant)	**Electrolyte abnormalities** Hypokalaemia Hypomagnesaemia Hypocalcaemia **Drugs** Quinidine, disopyramide Sotalol, amiodarone Tricyclic antidepressants, e.g. amitriptyline Phenothizine drugs, e.g. chlorpromazine Antipsychotics, e.g. haloperidol, olanzapine Macrolides, e.g. erythromycin Quinolones, e.g. ciprofloxacin Methadone **Poisons** Organophosphate insecticides **Miscellaneous** Bradycardia Mitral valve prolapse Acute myocardial infarction Diabetes Prolonged fasting and liquid protein diets (long-term) Central nervous system diseases, e.g. dystrophia myotonica

Long QT syndrome

This describes an ECG where the ventricular repolarization (QT interval) is greatly prolonged. The causes of long QT syndrome are listed in Table 13.14.

Congenital long QT syndrome

Two major syndromes have been described, which may (Jervell and Lange–Nielsen syndrome) or may not (Romano–Ward syndrome) be associated with congenital deafness.

The molecular biology of the congenital long QT syndromes has been shown to be heterogeneous. It is usually a monogenic disorder and has been associated with mutations in cardiac potassium and sodium channel genes (Table 13.15). The different genes involved appear to correlate with different phenotypes (Fig. 13.52b) that can exhibit such variable penetrance that carriers may have completely normal ECGs. There are three major sub-types: LQT1 in which the arrhythmia is usually provoked by exercise, particularly swimming; LQT2 with provocation associated with emotion and acoustic stimuli; and LQT3 where the arrhythmias occur during rest or when asleep. It is likely that identification of the mutation involved will not only improve diagnostic accuracy but also guide future therapy for the congenital long QT syndrome.

FURTHER READING

Morita H, Wu J, Zipes DP. The QT syndromes: long and short. *Lancet* 2008; **372**: 750–763.

In *acquired long QT syndrome*, QT prolongation and torsades de pointes are usually provoked by bradycardia.

Clinical features

Patients with a long QT develop syncope and palpitations as a result of polymorphic ventricular tachycardia (torsade de pointes). They usually terminate spontaneously but may degenerate to ventricular fibrillation, resulting in sudden death.

Torsades de pointes is characterized on the ECG by rapid, irregular, sharp complexes that continuously change from an upright to an inverted position (Fig. 13.52a).

Between spells of tachycardia or immediately preceding the onset of tachycardia the ECG shows a prolonged QT interval; the corrected QT (Table 13.5, p. 695) is usually greater than 0.50 s.

Table 13.15 Single gene mutations responsible for congenital long QT syndrome

Subtype	Chromosome	Gene	Channel
LQT1	11p15.5	KCNQ1	I_{ks} α subunit
LQT2	7q35–36	KCNH2	I_{kr} α subunit
LQT3	3p21–24	SCN5A	I_{Na} α subunit
LQT4	4q24–27	ANK2	Na, K, Ca
LQT5	21q22.1	KCNE1	I_{ks} α subunit
LQT6	21q22.1	KCNE2	I_{kr} α subunit
LQT7	17q23	KCNJ2	I_{ks} α subunit
LQT8	12p13.3	CACNA1C	I_{Ca} α subunit
LQT9	3p25.3	CAV3	
LQT10	11q23.3	SCN4B	

Management

Acute (acquired):

- any electrolyte disturbance is corrected
- causative drugs are stopped
- the heart rate is maintained with atrial or ventricular pacing
- magnesium sulphate 8 mmol (Mg^{2+}) over 10–15 minutes for acquired long QT
- intravenous isoprenaline may be effective when QT prolongation is acquired (isoprenaline is contraindicated for congenital long QT syndrome).

Long-term, congenital long QT syndrome is generally treated by beta-blockade, left cardiac sympathetic denervation and pacemaker therapy. LQT1 patients seem to respond well to beta-blockade and LQT3 patients are better treated with sodium channel blockers. All LQT patients should avoid drugs known to prolong the QT interval. Patients who remain symptomatic despite conventional therapy and those with a strong family history of sudden death usually need ICD therapy.

Short QT syndrome

Five types have been described caused by genetic abnormalities leading to faster repolarization. Ventricular arrhythmias and sudden death may occur and an ICD is the best treatment.

Normal heart ventricular tachycardia

Monomorphic ventricular tachycardia in patients with structurally normal hearts (idiopathic VT) is usually a benign condition with an excellent long-term prognosis. Occasionally, it

Fig. 13.52 **Prolonged QT. (a)** An ECG demonstrating a supraventricular rhythm with a long QT (LQT) interval giving way to atypical ventricular tachycardia (torsades de pointes). The tachycardia is short-lived and is followed by a brief period of idioventricular rhythm. **(b)** Further three examples of a prolonged QT interval corresponding to three different types of LQT.

is incessant (so called Galavardin's tachycardia) and if untreated may lead to cardiomyopathy.

Normal heart VT either arises from a focus in the right ventricular outflow tract or in the left ventricular septum. Treatment of symptoms is usually with beta-blockers. There is a special form of verapamil-sensitive tachycardia that responds well to non-dihydropyridine calcium antagonists. In symptomatic patients, radiofrequency catheter ablation is highly effective, resulting in a cure in >90% cases. It is sometimes difficult to distinguish ARVC (p. 791) from this seemingly benign disorder.

Non-sustained ventricular tachycardia

Non-sustained ventricular tachycardia (NSVT) is defined as ventricular tachycardia that is ≥5 consecutive beats but lasts <30 s (Fig. 13.53d). NSVT can be found in 6% of patients with normal hearts and usually does not require treatment. NSVT is documented in up to 60–80% of patients with heart disease. There is insufficient evidence on prognosis, but an ICD has been shown to improve survival of patients with particularly poor left ventricular function (ejection fraction 30% or less) by preventing arrhythmic death. Antiarrhythmic suppression of NSVT is not usually advocated but beta-blockers may improve quality of life in symptomatic individuals.

Ventricular premature beats

These may be uncomfortable, especially when frequent. The patient complains of extra beats, missed beats or heavy beats because it may be the premature beat, the post-ectopic pause or the next sinus beat that is noticed by the patient. The pulse is irregular owing to the premature beats. Some early beats may not be felt at the wrist. When a premature beat occurs regularly after every normal beat, 'pulsus bigeminus' may occur. If premature ventricular beats are highly symptomatic, treatment with beta-blockade may be helpful. If the ectopics stem from a single focus, especially when in the right ventricle, catheter ablation can be very effective.

These premature beats (Fig. 13.53a-c) have a broad (>0.12 s) and bizarre QRS complex because they arise from an abnormal (ectopic) site in the ventricular myocardium. Following a premature beat there is usually a complete compensatory pause because the AV node or ventricle is refractory to the next sinus impulse. Early 'R-on-T' ventricular premature beats (occurring simultaneously with the upstroke or peak of the T wave of the previous beat) may induce ventricular fibrillation in patients with heart disease, particularly in patients following myocardial infarction.

Ventricular premature beats are usually treated only if symptomatic. Usually simple measures such as reassurance and beta-blocker therapy are all that is required.

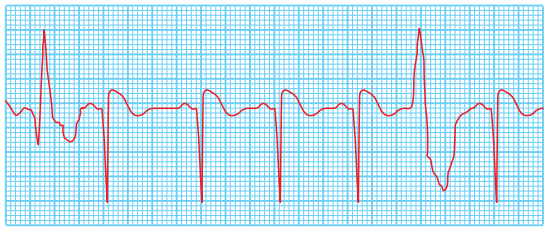
LONG-TERM MANAGEMENT OF CARDIAC TACHYARRHYTHMIAS

Options for the long-term management of cardiac tachyarrhythmias include:

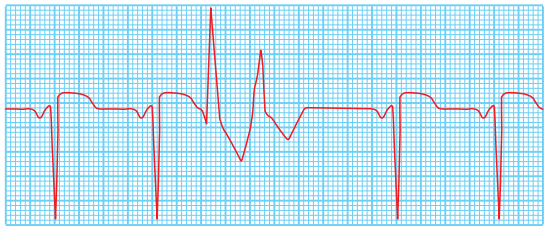
- antiarrhythmic drug therapy
- ablation therapy
- device therapy.

To determine the optimal strategy for a given patient the following questions must be addressed:

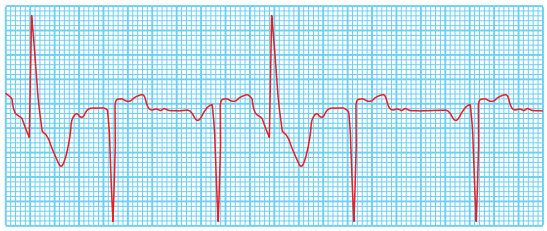
Fig. 13.53 **Varieties of ventricular ectopic activity.**



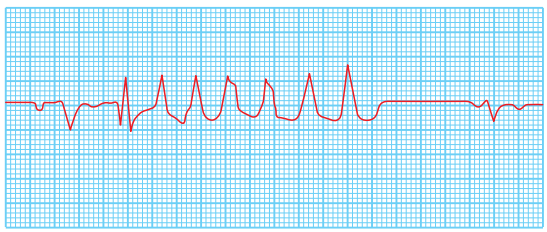
(a) Two ventricular ectopic beats of different morphology (multimorphological).



(b) Two ventricular premature beats (VPBs) occurring one after the other (a pair or couplet of VPBs).



(c) Frequently repetitive ventricular ectopic activity of a single morphology.



(d) Brief run of ventricular tachycardia (non-sustained ventricular tachycardia) that follows previous ectopic activity.

- Is the principal aim of treatment symptom relief or prevention of sudden death?
- Is maintaining sinus rhythm or controlling ventricular rates the treatment goal?

Commonly employed treatment strategies for the management of specific tachyarrhythmias are outlined in Table 13.16.

Antiarrhythmic drugs

Drugs that modify the rhythm and conduction of the heart are used to treat cardiac arrhythmias. Antiarrhythmic drugs may aggravate or produce arrhythmias (proarrhythmia) and they may also depress ventricular contractility and must therefore be used with caution. They are classified according to their effect on the action potential (Vaughan Williams' classification; Table 13.17 and Fig. 13.54).

Table 13.16 Long-term management of tachyarrythmias

Tachycardia	Management aims	Management strategies
AV node re-entry tachycardia (AVNRT)	Relieve symptoms	AV-node-blocking agents Class Ic or class III Catheter ablation
AV reciprocating tachycardia (AVRT)	Relieve symptoms	AV-node-blocking agents Class Ic or class III Catheter ablation
Wolff–Parkinson–White (WPW) syndrome	Relieve symptoms Prevent sudden death (esp. if documented pre-excited atrial fibrillation)	Class Ic or class III Catheter ablation
Atrial fibrillation	Relieve symptoms Prevent worsening heart failure due to poor rate control Prevent thromboembolic complications	Maintenance of sinus rhythm: class Ic or class III ± cardioversion catheter ablation of ectopic focus Rate control: AV-node-slowing agents AV node ablation plus plus pacemaker Anticoagulation
Atrial flutter	Relieve symptoms Prevent worsening heart failure due to poor rate control Prevent thromboembolic complications	Class Ic or class III Catheter ablation AV-node-blocking agents Anticoagulation
Atrial tachycardia	Relieve symptoms Prevent worsening heart failure due to poor rate control Prevent thromboembolic complications	Class Ic or class III Catheter ablation AV-node-blocking agents Anticoagulation
Life-threatening ventricular tachyarrhythmias	Prevent sudden death	Implantable cardioverter–defibrillator (ICD) Beta-blockers
Congenital long QT	Prevent sudden death	Beta-blockers ± pacemaker ICD Correct bradycardia Correct electrolytes
Acquired long QT	Prevent sudden death	Avoidance of all QT-prolonging drugs
Normal heart ventricular tachycardias	Relieve symptoms	Beta-blockers, calcium channel blockers Catheter ablation
Non-sustained VT (NSVT)	Relieve symptoms Prevent sudden death in certain situations	Beta-blockers Class Ic and class III ICD in clearly defined subgroups

Class I drugs

These are membrane-depressant drugs that reduce the rate of entry of sodium into the cell (sodium-channel blockers). They may slow conduction, delay recovery or reduce the spontaneous discharge rate of myocardial cells. Class Ia drugs (e.g. disopyramide) lengthen the action potential, class Ib drugs (e.g. lidocaine) shorten the action potential, and class Ic (flecainide, propafenone) do not affect the duration of the action potential. Class I agents have been found to increase mortality compared to placebo in post-myocardial infarction patients with ventricular ectopy (Cardiac Arrhythmia Suppression Trial (CAST) trials – class Ic agents) and in patients treated for atrial fibrillation (class Ia agent, quinidine). In view of this, class Ic agents such as flecainide and all other class I drugs should be reserved for patients who do not have significant coronary artery disease, left ventricular dysfunction, or other forms of significant structural heart disease.

Class II drugs

These antisympathetic drugs prevent the effects of catecholamines on the action potential. Most are β-adrenoceptor antagonists. Cardioselective beta-blockers (β_1) include metoprolol, bisoprolol, atenolol and acebutalol. Beta-blockers suppress AV node conduction, which may be effective in preventing attacks of junctional tachycardia, and may help to control the ventricular rate during paroxysms of other forms of SVT (e.g. atrial fibrillation). In general beta-blockers are anti-ischaemic and anti-adrenergic, and have proven beneficial effects in patients post-myocardial infarction (by preventing ventricular fibrillation) and in patients with congestive heart failure. It is therefore advisable to use beta-blocker therapy either alone or in combination with other antiarrhythmic drugs in patients with symptomatic tachyarrhythmias, particularly in patients with coronary artery disease.

Class III drugs

These prolong the action potential, usually by blocking the rapid component of the delayed rectifier potassium current (I_{Kr}) and do not affect sodium transport through the membrane. The drugs in this class are amiodarone and sotalol. Sotalol is also a beta-blocker.

Some drugs, such as ibutilide and dofetilide, are only available in some countries. Sotalol may result in acquired long QT syndrome and torsades de pointes. The risk of torsades is increased in the setting of hypokalaemia, and particular care should be taken in patients taking diuretic therapy. Amiodarone therapy in contrast to most other antiarrhythmic drugs carries a low risk of proarrhythmia in patients with significant structural heart disease. Because it has many toxic and potentially serious side-effects, patients need to be counselled prior to commencing amiodarone and

monitored at regular follow-up intervals. Dofetilide has been used to treat atrial fibrillation and flutter in patients with recent myocardial infarction and poor LV function.

Class IV drugs

The non-dihydropyridine calcium channel blockers that reduce the plateau phase of the action potential are particularly effective at slowing conduction in nodal tissue. Verapamil and diltiazem are two drugs in this group. These drugs can prevent attacks of junctional tachycardia (AVNRT and AVRT) and may help to control ventricular rates during paroxysms of other forms of SVT (e.g. atrial fibrillation).

Table 13.17 Vaughan Williams' classification of antiarrhythmic drugs

Class I	Membrane-depressant drugs (sodium-channel blockers)
Ia	Disopyramide
Ib	Lidocaine, mexiletine
Ic	Flecainide, propafenone
Class II	β-Adrenoceptor blocking drugs, e.g. atenolol, propranolol, esmolol
Class III	Prolong action potential, e.g. amiodarone, sotalol
Class IV	Calcium-channel blockers, e.g. verapamil, diltiazem
(Other	Adenosine, digoxin)

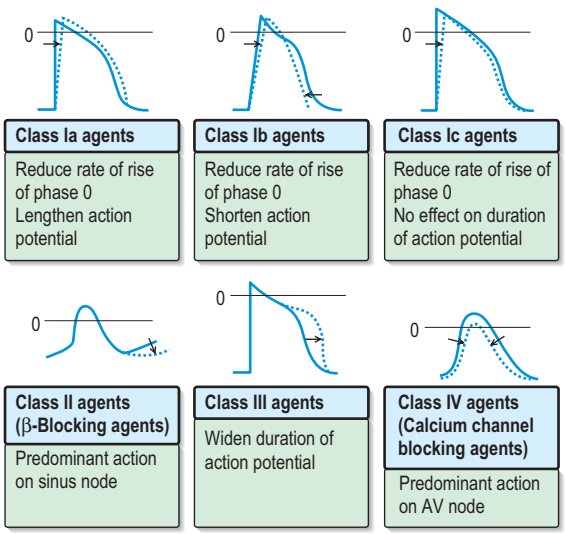


Fig. 13.54 **Vaughan Williams' classification of antiarrhythmic drugs** based on their effect on cardiac action potentials. 0, = 0 mV. The dotted curves indicate the effects of the drugs.

Drug therapy is commonly used for symptomatic relief in patients who do not have life-threatening tachyarrhythmias. Patient safety is the main factor determining the choice of antiarrhythmic therapy and proarrhythmic risks need to be carefully assessed prior to initiating therapy. As a generalization, class Ic agents are employed in patients with structurally normal hearts and class III agents are used in patients with structural heart disease.

In order to administer antiarrhythmic agents as safely as possible it is helpful to be familiar with the different proarrhythmia mechanisms and their main predisposing risk factors (Table 13.18).

Patients with structurally normal hearts and normal QT intervals, or with implantable defibrillators, are either at very low risk of proarrhythmia or are protected from any life-threatening consequences, and in these patients it is possible to persevere with drug therapy until an efficacious, well-tolerated agent is identified.

Catheter ablation

Radiofrequency catheter ablation is frequently employed in the management of symptomatic tachyarrhythmias. Ablations are performed by placing three or four electrode catheters into the heart chambers in order to record and pace from various sites. Pacing the atria or the ventricles is used to trigger the tachycardia and to study the tachycardia mechanism. Successful ablation depends on accurate identification of either the site of origin of a focal tachycardia or of a critical component of a macro-reentry tachycardia. The following tachyarrhythmias can be readily ablated:

- AV node re-entry tachycardia (AVNRT)
- AV re-entry tachycardia (AVRT) with an accessory pathway, including WPW syndrome
- normal heart VT
- atrial flutter
- atrial tachycardia
- atrial fibrillation.

Symptomatic patients with a pre-excited ECG because of accessory pathway conduction (WPW syndrome) are advised to undergo catheter ablation as first-line therapy, owing to the risk of sudden death associated with this condition. This is especially the case in patients with pre-excited atrial fibrillation. Patients with accessory pathways that only

Table 13.18 Proarrhythmia mechanisms and the predisposing risk factors

Proarrhythmia mechanism	Drug	Predisposing risk factors
Torsades de pointes	Class I (disopyramide, procainamide, quinidine) Sotalol 'Pure' class III (ibutilide)	Baseline QT prolongation Female gender Hypokalaemia/diuretic use Clinical heart failure Advanced structural heart disease Conversion of AF to sinus rhythm
Atrial flutter with 1:1 AV conduction and wide QRS complexes	Class 1_a (disopyramide, quinidine) Class I_c (flecainide, propafenone)	Risk reduced by adding an AV-nodal-slowing agent
'Late' sudden death (arrhythmic mechanism not clearly defined)	Class I_c (flecainide) Quinidine	Myocardial ischaemia

conduct retrogradely from the ventricles to the atrium are not at increased risk of sudden death but experience symptoms due to AVRT. These patients are commonly offered an ablation procedure if simple measures such as AV nodal slowing agents fail to suppress tachycardia. Asymptomatic patients with the WPW ECG pattern are now frequently offered an ablation procedure for prophylactic reasons. The main risk associated with accessory pathway ablation is thromboembolism in patients with left-sided accessory pathways. The success rate for catheter ablation of AVNRT and accessory pathways is greater than 95%.

Patients with normal hearts and documented ventricular tachycardia should be referred for specialist evaluation. Unlike VT in patients with structural heart disease, normal heart VT is not associated with increased risk of sudden death and is easily cured by catheter ablation.

Catheter ablation is recommended in patients with atrial flutter that is not easily managed medically. Ablation of typical flutter is effective in 90–95% cases. In the direct comparison of catheter ablation and antiarrhythmic therapy, the rate of recurrence was significantly lower following ablation. Atrial tachycardia, especially in patients with structurally normal hearts, may also be cured by catheter ablation. In atrial fibrillation, adequate control of ventricular rates is sometimes not possible despite optimal medical therapy. These patients experience a marked symptomatic improvement following AV node ablation and pacemaker implantation. Unlike other forms of catheter ablation this does not cure the arrhythmia; atrial fibrillation continues and anticoagulation is still required following ablation.

In *younger patients* with structurally normal hearts, atrial ectopic beats, which commonly arise from a focus situated in the pulmonary veins, may trigger atrial fibrillation. Catheter ablation of this ectopic focus includes the application of radiofrequency energy in or around the pulmonary veins in order to abolish the connection between the sleeves of arrhythmogenic atrial myocardium surrounding or extending into the veins from the remaining atrium (pulmonary vein isolation). The trigger is therefore eliminated and the arrhythmia does not recur. These techniques appear to be highly effective, especially in young patients with paroxysmal atrial fibrillation, normal atrial size and no underlying heart disease (70–80% long-term success), but are presently time-consuming procedures (4 hours or more) and are associated with some serious complications such as stroke, pericardial haemorrhage, pulmonary vein stenosis and atrio-oesophageal fistula in a small minority of patients (<2% in all).

Implantable cardioverter–defibrillator (ICD)

Life-threatening ventricular arrhythmias (ventricular fibrillation or rapid ventricular tachycardia with hypotension) result in death in up to 40% within 1 year of diagnosis. Large multicentre prospective trials such as the Antiarrhythmics (amiodarone) Versus Implantable Defibrillator (AVID) trial have proven that implantable defibrillators improve overall survival in patients who have experienced an episode of life-threatening ventricular tachyarrhythmia (Fig. 13.55).

The ICD recognizes ventricular tachycardia or fibrillation and automatically delivers pacing or a shock to the heart to cause cardioversion to sinus rhythm. Modern ICDs are only

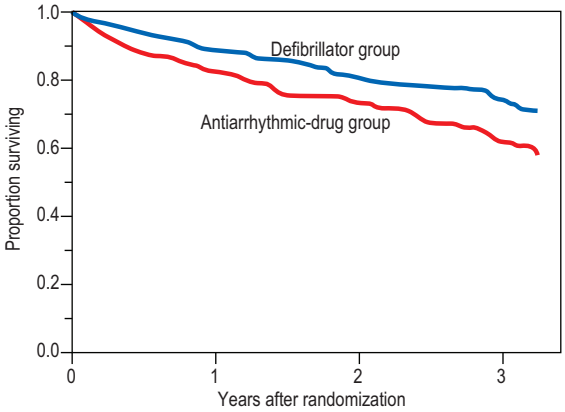


Fig. 13.55 **Survival curves of the Multicentre Automatic Defibrillator Implantation Trial (MADIT)** following myocardial infarction. Patients with left ventricular ejection fractions of ≤0.35 and documented asymptomatic non-sustained ventricular tachycardia and inducible ventricular tachycardia were randomized to receive an ICD or conventional therapy.

a little larger than a pacemaker and are implanted in a pectoral position (Fig. 13.56). The device may have leads to sense and pace both the right atrium and ventricle, and the lithium batteries employed are able to provide energy for over 100 shocks each of around 30 J. ICD discharges are painful if the patient is conscious. However, ventricular tachycardia may often be terminated by overdrive pacing the heart, which is painless. The ICD is superior to all other treatment options at preventing sudden cardiac death. The use of this device has cut the sudden death rate in patients with a history of serious ventricular arrhythmias to approximately 2% per year. However, the majority of these patients have significant structural heart disease and overall cardiac mortality due to progressive heart failure remains high. As a result the ICD is now first-line therapy in the secondary prevention of sudden death.

Implantable cardioverter–defibrillators are also employed in the primary prevention of sudden cardiac death. The chances of surviving an out-of-hospital cardiac arrest are as low as 10%. Therefore selected patients who have never experienced a spontaneous episode of life-threatening ventricular tachyarrhythmia but who are assessed to be at high risk of sudden death are advised to undergo ICD implantation. In two large primary prevention ICD trials, Multicenter Automated Defibrillator Implantation Trial (MADIT II) and Sudden Cardiac Death in Heart Failure Trial (SCD-HeFT), therapy with an ICD reduced mortality by 23–31% on top of conventional treatment, which included revascularization, beta-blockers and angiotensin-converting enzyme inhibitors.

The following groups of patients may merit prophylactic ICD placement:

- Patients with coronary artery disease; significant impairment of left ventricular function (LVEF ≤35–40%), spontaneous non-sustained ventricular tachycardia in whom sustained ventricular tachycardia was induced by pacing the heart during an electrophysiological study.

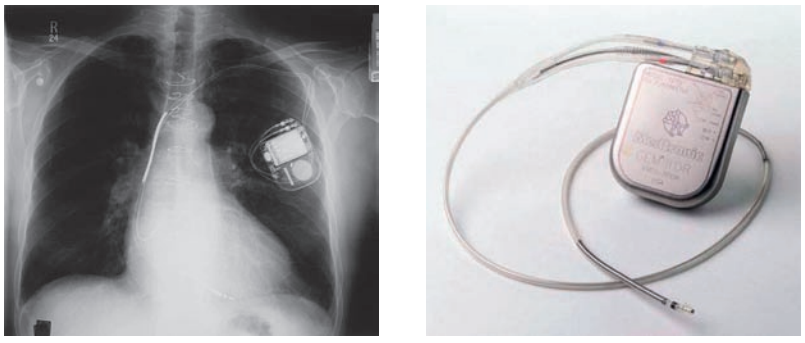


Fig. 13.56 **Implantable cardioverter–defibrillator. (a) X-ray of a 'dual-chamber' ICD** in a left pectoral position with atrial and ventricular leads. **(b) Termination of ventricular fibrillation** by the direct current shock at 20 J. Electrogram recorded internally from the ventricular lead of an ICD reveals chaotic ventricular activity consistent with ventricular fibrillation. This is confirmed by an electrocardiogram (lead V_2). The marker channel demonstrates that the device detects ventricular fibrillation correctly (FS, fibrillation sensing, lower line) and delivers an appropriate shock (CD) that terminates the arrhythmia and restores normal sinus rhythm (VS, ventricular sensing, lower line). Incidentally, atrial fibrillation is also detected before shock delivery (upper line on the marker channel).

- Patients with very poor LV function post-MI (LVEF ≤35%).
- Patients with dilated and particularly hypertrophic cardiomyopathy, long QT syndrome and Brugada syndrome or other channelopathies who have a strong family history of sudden cardiac death.

FURTHER READING

Fuster V, Ryden LE, Cannom DS, et al. ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation: *Europace* 2006; **8**: 651–745.

National Institute of Health and Clinical Excellence. The management of atrial fibrillation. Clinical Guideline 36, 2006. www.nice.org.uk.

National Institute of Health and Clinical Excellence. Implantable cardioverter defibrillators (ICDs) for the treatment of arrhythmias. Technology appraisal review TA11, 2006. www.nice.org.uk.

Rodes DM. Long-QT syndrome. *New England Journal of Medicine* 2008; **358**: 169–176.

Vardas PE, Auricchio A, Blanc JJ, et al. European Society of Cardiology; European Heart Rhythm Association. Guidelines for cardiac pacing and cardiac resynchronization therapy. The Task Force for Cardiac Pacing and Cardiac Resynchronization Therapy of the European Society of Cardiology. Developed in collaboration with the European Heart Rhythm Association. *Europace* 2007; **9**: 959–998.

Zipes DP, Camm AJ, Borggrefe M, et al. ACC/AHA/ESC 2006 guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death. *Europace* 2006; **8**: 746–837.

HEART FAILURE

Heart failure is a complex syndrome that can result from any structural or functional cardiac disorder that impairs the ability of the heart to function as a pump to support a physiological circulation.

World-wide the incidence of heart failure is variable but increases with advancing age. For example, in Scotland the prevalence of heart failure is high at 7.1 in 1000, increasing with age to 90.1 in 1000 among patients over 85 years. In the UK overall incidence is about 2 in 1000. Approximately 23 million people world-wide have heart failure.

The prognosis of heart failure has improved over the past 10 years, but the mortality rate is still high with approximately 50% of patients dead at 5 years. Heart failure accounts for 5% of admissions to hospital medical wards. The cost of managing heart failure in the UK exceeds £1 billion per year. Coronary artery disease is the commonest cause of heart failure in western countries.

The causes of heart failure are shown in Table 13.19.

PATHOPHYSIOLOGY

When the heart fails, considerable changes occur to the heart and peripheral vascular system in response to the haemodynamic changes associated with heart failure (Table 13.20). These physiological changes are compensatory and maintain cardiac output and peripheral perfusion. However, as heart failure progresses, these mechanisms are overwhelmed and become pathophysiological. The development of pathological peripheral vasoconstriction and sodium retention in heart failure by activation of the renin–angiotensin–aldosterone system are a loss of beneficial compensatory mechanisms and represent cardiac decompensation. Factors involved are venous return, outflow resistance, contractility of the myocardium, and salt and water retention.

Table 13.19 Causes of heart failure

Main causes

Ischaemic heart disease (35–40%)

Cardiomyopathy (dilated) (30–34%)

Hypertension (15–20%)

Other causes

Cardiomyopathy (undilated): hypertrophic/obstructive, restrictive (amyloidosis, sarcoidosis)

Valvular heart disease (mitral, aortic, tricuspid)

Congenital heart disease (ASD, VSD)

Alcohol and drugs (chemotherapy – trastuzamab, imatinib)

Hyperdynamic circulation (anaemia, thyrotoxicosis, haemochromatosis, Paget's disease)

Right heart failure (RV infarct, pulmonary hypertension, pulmonary embolism, cor pulmonale (COPD))

Tricuspid incompetence

Arrhythmias (atrial fibrillation, bradycardia (complete heart block, the sick sinus syndrome))

Pericardial disease (constrictive pericarditis, pericardial effusion)

Infections (Chagas' disease)

Table 13.20 Pathophysiological changes in heart failure

Ventricular dilatation

Myocyte hypertrophy

Increased collagen synthesis

Altered myosin gene expression

Altered sarcoplasmic Ca^{2+}-ATPase density

Increased ANP secretion

Salt and water retention

Sympathetic stimulation

Peripheral vasoconstriction

Venous return (preload)

In the intact heart, myocardial failure leads to a reduction of the volume of blood ejected with each heart beat and an increase in the volume of blood remaining after systole. This increased diastolic volume stretches the myocardial fibres and, as Starling's law of the heart (p. 684) would suggest, myocardial contraction is restored. However, the failing myocardium results in depression of the ventricular function curve (cardiac output plotted against the ventricular diastolic volume) (Fig. 13.5, p. 684).

Mild myocardial depression is not associated with a reduction in cardiac output because it is maintained by an increase in venous pressure (and hence diastolic volume). However, the proportion of blood ejected with each heart beat (ejection fraction) is reduced early in heart failure. Sinus tachycardia also ensures that any reduction of stroke volume is compensated for by the increase in heart rate; cardiac output (stroke volume × heart rate) is therefore maintained.

When there is more severe myocardial dysfunction, cardiac output can be maintained only by a large increase in venous pressure and/or marked sinus tachycardia. The increased venous pressure contributes to the development of dyspnoea, owing to the accumulation of interstitial and alveolar fluid, and to the occurrence of hepatic enlargement, ascites and dependent oedema, due to increased systemic venous pressure. However, the cardiac output at rest may not be much depressed, but myocardial and haemodynamic reserve is so compromised that a normal increase in cardiac output cannot be produced by exercise.

In very severe heart failure the cardiac output at rest is depressed, despite high venous pressures. The inadequate cardiac output is redistributed to maintain perfusion of vital organs, such as the heart, brain and kidneys, at the expense of the skin and muscle.

Outflow resistance (afterload) (see Fig. 15.4)

This is the load or resistance against which the ventricle contracts. It is formed by:

- pulmonary and systemic resistance
- physical characteristics of the vessel walls
- the volume of blood that is ejected.

An increase in afterload decreases the cardiac output. This results in a further increase of end-diastolic volume and dilatation of the ventricle itself, which further exacerbates the problem of afterload. This is expressed by Laplace's law: the tension of the myocardium (T) is proportional to the intraventricular pressure (P) multiplied by the radius of the ventricular chamber (R): i.e. $T \propto PR$.

Myocardial contractility (inotropic state)

The state of the myocardium also influences performance. The sympathetic nervous system is activated in heart failure via baroreceptors as an early compensatory mechanism, which provides inotropic support and maintains cardiac output. Chronic sympathetic activation, however, has deleterious effects by further increasing neurohormonal activation and myocyte apoptosis. This is compensated by a down-regulation of β-receptors. Increased contractility (positive inotropism) can result from increased sympathetic drive, and this is a normal part of the Frank–Starling relationship (Fig. 13.5, p. 684). Conversely, myocardial depressants (e.g. hypoxia) decrease myocardial contractility (negative inotropism).

Neurohormonal and sympathetic system activation: salt and water retention

The increase in venous pressure that occurs when the ventricles fail leads to retention of salt and water and their accumulation in the interstitium, producing many of the physical signs of heart failure. Reduced cardiac output also leads to diminished renal perfusion, activating the renin–angiotensin system and enhancing salt and water retention (Fig. 12.3, p. 651), which further increases venous pressure (Fig. 13.57). The retention of sodium is in part compensated by the action of circulating atrial natriuretic peptides and antidiuretic hormone (see p. 653 and below).

The interaction of haemodynamic and neurohumoral factors in the progression of heart failure remains unclear. Increased ventricular wall stress promotes ventricular dilatation and further worsens contractile efficiency. In addition, prolonged activation of the sympathetic nervous and renin-angiotensin-aldosterone systems exerts direct toxic effects on myocardial cells.

Myocardial remodelling in heart failure

Left ventricular remodelling is a process of progressive alteration of ventricular size, shape and function owing to the influence of mechanical, neurohormonal and possibly genetic factors in several clinical conditions, including myocardial

infarction, cardiomyopathy, hypertension and valvular heart disease. Its hallmarks include hypertrophy, loss of myocytes and increased interstitial fibrosis. Remodelling continues for months after the initial insult, and the eventual change in the shape of the ventricle becomes responsible for significant impairment of overall function of the heart as a pump (Fig. 13.58a). In cardiomyopathy, the process of progressive ventricular dilatation or hypertrophy occurs without ischaemic myocardial injury or infarction (Fig. 13.58b).

Changes in myocardial gene expression

Haemodynamic overload of the ventricle stimulates changes in cardiac contractile protein gene expression. The overall effect is to increase protein synthesis, but many proteins also switch to fetal and neonatal isoforms. Human myosin is composed of a pair of heavy chains and two pairs of light chains. Myosin heavy chains (MHC) exist in two isoforms, α and β, that have different contractile properties and ATPase activity. αα-MHC predominates in the atria and ββ-MHC in the ventricles. In animal models, pressure overload results in a shift from αα- to ββ-MHC in the atria, in parallel with atrial size. This results in reduction in atrial contractility but reduced energy demands. This shift is less significant in the human ventricle, as the ββ-MHC isoform already predominates. Other genes affected in heart failure include those encoding Na^+/K^+-ATPase, Ca^{2+}-ATPase and β_1-adrenoceptors.

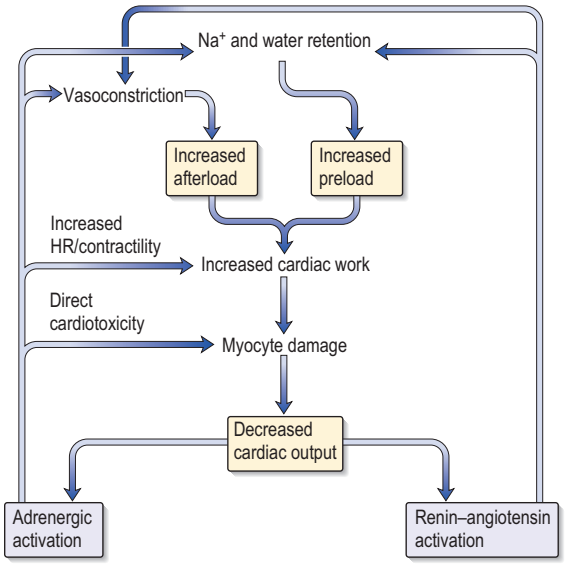


Fig. 13.57 **The compensatory physiological response to heart failure**. Chronic activation of the renin–angiotensin and adrenergic systems results in a 'vicious cycle' of cardiac deterioration that further exacerbates the physiological response.

Abnormal calcium homeostasis

Calcium ion flux within myocytes plays a pivotal role in the regulation of contractile function. Excitation of the myocyte cell membrane causes the rapid entry of calcium into myocytes from the extracellular space via calcium channels. This triggers the release of intracellular calcium from the sarcoplasmic reticulum and initiates contraction (see Fig. 13.3, p. 683). Relaxation results from the uptake and storage of calcium by the sarcoplasmic reticulum (see Fig. 13.9, p. 686) controlled by changes in nitric oxide. In heart failure, there is a prolongation of the calcium current in association with prolongation of contraction and relaxation.

Apoptosis (see also p. 27)

Apoptosis (or 'programmed cell death') of myocytes has been demonstrated in animal models of ischaemic reperfusion, rapid ventricular pacing, mechanical stretch and pressure overload. Apoptosis is associated with irreversible congestive heart failure, and the spiral of ventricular dysfunction, characteristic of heart failure, results from the initiation of apoptosis by cytokines, free radicals and other triggers.

Natriuretic peptides (ANP, BNP and C-type)

Atrial natriuretic peptide (ANP) is released from atrial myocytes in response to stretch. ANP induces diuresis, natriuresis, vasodilatation and suppression of the renin–angiotensin

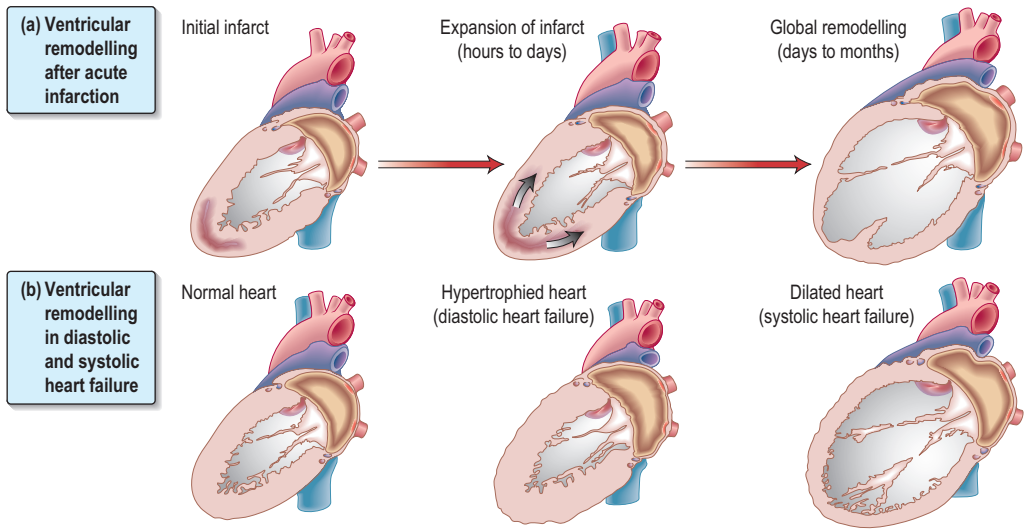


Fig. 13.58 **Ventricular remodelling in (a) ischaemic (MI) and (b) non-ischaemic heart failure**, e.g. in cardiomyopathy. From Jessop M, Brozana S. *New England Journal of Medicine* 2003; **348**: 2011, with permission. Copyright © 2003 Massachusetts Medical Society. All rights reserved.

Ask the authors
www.studentconsult.com

system. Levels of circulating ANP are increased in congestive cardiac failure and correlate with functional class, prognosis and haemodynamic state. The renal response to ANP is attenuated in heart failure, probably secondarily to reduced renal perfusion, receptor downregulation, increased peptide breakdown, renal sympathetic activation and excessive renin–angiotensin activity. Brain natriuretic peptide (BNP) (so called because it was first discovered in brain) is predominantly secreted by the ventricles, and has an action similar to that of ANP but greater diagnostic and prognostic value (p. 1023). *C-type peptide*, which is limited to vascular endothelium and the central nervous system, has similar effects to those of ANP and BNP.

The therapeutic benefits of the natriuretic peptides have been investigated in two ways. The administration of *synthetic BNP* (Nesritide) produces beneficial haemodynamic effects in acute heart failure (reducing pulmonary capillary wedge pressure and peripheral vascular resistance) but has been associated with an increased risk of death.

Neutral endopeptidase (NEP) is a metalloendopeptidase involved in the degradation of a variety of vasoactive peptides (including ANP, BNP, CNP and bradykinin). In animal studies NEP inhibitors can produce diuresis and natriuresis. Omapatrilat is a combined ACE-I and NEP-I that is an effective antihypertensive agent, but it has not received FDA approval because of an increased risk of severe angio-oedema.

Endothelial function in heart failure

The endothelium has a central role in the regulation of vasomotor tone. In patients with heart failure, endothelium-dependent vasodilatation in peripheral blood vessels is impaired and may be one mechanism of exercise limitation. The cause of abnormal endothelial responsiveness relates to abnormal release of both nitric oxide and vasoconstrictor substances, such as endothelin (ET). The activity of nitric oxide, a potent vasodilator, is blunted in heart failure. ET secretion from a variety of tissues is stimulated by many factors, including hypoxia, catecholamines and angiotensin II. The plasma concentration of ET is elevated in patients with heart failure, and levels correlate with the severity of haemodynamic disturbance. The major source of circulating ET in heart failure is the pulmonary vascular bed.

ET has many actions that potentially contribute to the pathophysiology of heart failure: vasoconstriction, sympathetic stimulation, renin–angiotensin system activation and left ventricular hypertrophy. Acute intravenous administration of endothelin antagonists improves haemodynamic abnormalities in patients with congestive cardiac failure, and oral endothelin antagonists are being developed. Plasma concentrations of some cytokines, in particular TNF, are increased in patients with heart failure.

Antidiuretic hormone (vasopressin)

Antidiuretic hormone (ADH) is raised in severe chronic heart failure, particularly in patients on diuretic treatment. High ADH concentration precipitates hyponatraemia which is an ominous prognostic indicator.

CLINICAL SYNDROMES OF HEART FAILURE

There are many causes of heart failure (see Table 13.19) that can present suddenly with acute heart failure (AHF) or more insidiously with chronic heart failure (CHF).

Left ventricular systolic dysfunction (LVSD) is commonly caused by ischaemic heart disease but can also occur with valvular heart disease and hypertension.

Right ventricular systolic dysfunction (RVSD) may be secondary to chronic LVSD but can occur with primary and secondary pulmonary hypertension, right ventricular infarction, arrhythmogenic right ventricular cardiomyopathy and adult congenital heart disease.

Diastolic heart failure is a syndrome consisting of symptoms and signs of heart failure with preserved left ventricular ejection fraction above 45–50% and abnormal left ventricular relaxation assessed by echocardiography. There is increased stiffness in the ventricular wall and decreased left ventricular compliance, leading to impairment of diastolic ventricular filling and hence decreased cardiac output. Diastolic heart failure is more common in elderly hypertensive patients but may occur with primary cardiomyopathies (hypertrophic, restrictive, infiltrative).

The symptoms and signs of heart failure are:

Symptoms:

- exertional dyspnoea
- orthopnoea
- paroxysmal nocturnal dyspnoea
- fatigue.

Signs:

- cardiomegaly
- third and fourth heart sounds
- elevated JVP
- tachycardia
- hypotension
- bi-basal crackles
- pleural effusion
- ankle oedema
- ascites
- tender hepatomegaly.

The New York Heart Association classification of heart failure (Table 13.21) can be used to describe the symptoms of heart failure and limitation of exercise capacity, and are useful to assess response to therapy. They do not include left ventricular ejection fraction to determine severity of heart failure.

Diagnosis of heart failure

The diagnosis of heart failure should not be based on history and clinical findings; it requires evidence of cardiac dysfunction with appropriate investigation using objective measures of left ventricular structure and function (usually echocardiography). Similarly, the underlying cause of heart failure

Table 13.21 New York Heart Association (NYHA) Classification of heart failure

Class I	No limitation. Normal physical exercise does not cause fatigue, dyspnoea or palpitations
Class II	Mild limitation. Comfortable at rest but normal physical activity produces fatigue, dyspnoea or palpitations
Class III	Marked limitation. Comfortable at rest but less gentle physical activity produces marked symptoms of heart failure
Class IV	Symptoms of heart failure occur at rest and are exacerbated by any physical activity

Table 13.22 Diagnosis of heart failure (European Society of Cardiology guidelines)

Essential features (criteria 1 and 2)

1. Symptoms and signs of heart failure (e.g. breathlessness, fatigue, ankle swelling)
2. Objective evidence of cardiac dysfunction (at rest)

Non-essential features

3. Response to treatment directed towards heart failure (in cases where the diagnosis is in doubt)

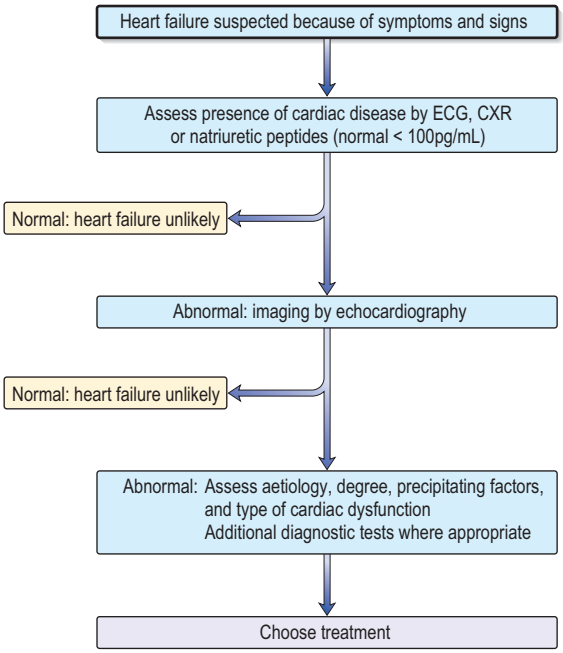


Fig. 13.59 **Algorithm for the diagnosis of heart failure.** Based on the European Society of Cardiology and NICE guidelines.

should be established in all patients (Table 13.22 and Fig. 13.59).

INVESTIGATIONS IN HEART FAILURE

- **Blood tests.** Full blood count, liver biochemistry, urea and electrolytes, cardiac enzymes in acute heart failure, BNP or N-terminal portion of proBNF (NPproBNP), thyroid function.
- **Chest X-ray.** Look for cardiomegaly, pulmonary congestion with upper lobe diversion, fluid in fissures, Kerley B lines, and pulmonary oedema.
- **Electrocardiogram** for ischaemia, hypertension or arrhythmia.
- **Echocardiography.** Cardiac chamber dimension, systolic and diastolic function, regional wall motion abnormalities, valvular heart disease, cardiomyopathies.
- **Stress echocardiography.** Assessment of viability in dysfunctional myocardium – dobutamine identifies contractile reserve in stunned or hibernating myocardium.
- **Nuclear cardiology.** Radionucleotide angiography (RNA) can quantify ventricular ejection fraction, single-photon-emission computed tomography (SPECT) or positron emission tomography (PET) can demonstrate myocardial ischaemia and viability in dysfunctional myocardium.
- **CMR (cardiac MRI).** Assessment of viability in dysfunctional myocardium with the use of dobutamine for contractile reserve or with gadolinium for delayed enhancement ('infarct imaging').
- **Cardiac catheterization.** Diagnosis of ischaemic heart failure (and suitability for revascularization), measurement of pulmonary artery pressure, left atrial (wedge) pressure, left ventricular end-diastolic pressure.
- **Cardiac biopsy.** Diagnosis of cardiomyopathies, e.g. amyloid, follow-up of transplanted patients to assess rejection.
- **Cardiopulmonary exercise testing.** Peak oxygen consumption (VO_2) is predictive of hospital admission and death in heart failure. A 6-minute exercise walk is an alternative.
- **Ambulatory 24-hour ECG monitoring (Holter).** In patients with suspected arrhythmia. May be used in patients with severe heart failure or inherited cardiomyopathy to determine if a defibrillator is appropriate (non-sustained ventricular tachycardia).

TREATMENT OF HEART FAILURE

Treatment is aimed at relieving symptoms, prevention and control of disease leading to cardiac dysfunction and heart failure, retarding disease progression and improving quality and length of life.

Measures to prevent heart failure include cessation of smoking, alcohol and illicit drugs, effective treatment of hypertension, diabetes and hypercholesterolaemia, and pharmacological therapy following myocardial infarction.

The management of heart failure requires that any factor aggravating the failure should be identified and treated. Similarly, the cause of heart failure must be elucidated and where possible corrected. Community nursing programmes to help with drug compliance and to detect early deterioration may prevent acute hospitalization.

General lifestyle advice

Education

Effective counselling of patients and family, emphasizing weight monitoring and dose adjustment of diuretics, may prevent hospitalization.

Obesity control

Maintain desired weight and body mass index.

Dietary modification

Large meals should be avoided and, if necessary, weight reduction instituted. Salt restriction is necessary and foods rich in salt or added salt in cooking and at the table should be avoided. A low-sodium diet is unpalatable and of questionable value. In severe heart failure fluid restriction is necessary. Alcohol has a negative inotropic effect and heart failure patients should moderate consumption.

A recent study has shown that omega-3 polyunsaturated fatty acids reduce mortality and admission to hospital.

Smoking

Smoking should be stopped, with help from antismoking clinics (p. 830) if necessary.

Physical activity, exercise training and rehabilitation

For patients with exacerbations of congestive cardiac failure, bed rest reduces the demands on the heart and is useful for a few days. Migration of fluid from the interstitium promotes a diuresis, reducing heart failure. Prolonged bed rest may lead to development of deep vein thrombosis; this can be avoided by daily leg exercises, low-dose subcutaneous heparin and elastic support stockings. Low-level endurance exercise (e.g. 20–30 minutes walking three or five times per week, or 20 minutes cycling at 70–80% of peak heart rate five times per week) is actively encouraged in patients with compensated heart failure in order to reverse 'deconditioning' of peripheral muscle metabolism. Strenuous isometric activity should be avoided.

Vaccination

While prospective clinical trials are lacking, it is recommended that patients with heart failure be vaccinated against pneumococcal disease and influenza.

Air travel

This is possible for most patients, subject to clinical circumstances. Check with the airline – most have guidelines on who should travel.

Sexual activity

Tell patients on nitrates not to take phosphodiesterase type 5 inhibitors (e.g. sildenafil) as it may induce profound hypotension.

Driving

Driving motor cars and motorcycles may continue provided there are no symptoms that distract the driver's attention. The DVLA in the UK does not need to be notified. Symptomatic heart failure disqualifies patients from driving large lorries and buses. Re/licensing may be permitted provided that the LV ejection fraction is good (i.e. LVEF is >0.4), the exercise test requirements can be met and there is no other disqualifying condition.

Monitoring of heart failure patients

The clinical condition of a person with heart failure fluctuates, and repeated admissions to hospital is common. Monitoring of clinical status is necessary and this responsibility should be shared between primary and secondary care health professionals.

Essential monitoring includes assessment of:

- functional capacity (e.g. VO_2 max, exercise tolerance test, echocardiography)
- fluid status (bodyweight, U&Es, clinical)
- cardiac rhythm (ECG, 24-hour tape).

Multidisciplinary team approach

Heart failure care should be delivered by a multidisciplinary team with an integrated approach across the healthcare community. The use of multiple biomarkers, e.g. troponins and CRP, is of prognostic value. The multidisciplinary team should involve specialist healthcare professionals: heart failure nurse, dietitian, pharmacist, occupational therapist, physiotherapist and palliative care adviser.

Understanding the information needs of patients and carers is vital. Good communication is essential for best clinical management, which should include advice on anxiety, depression and 'end of life' issues.

Drug management (Table 13.23)

The recommendations for pharmacological treatment of heart failure are based on consensus guidelines produced by the European Society of Cardiology and NICE (UK) (Figs 13.60 and 13.61).

Diuretics (p. 657)

These act by promoting the renal excretion of salt and water by blocking tubular reabsorption of sodium and chloride. Loop diuretics (e.g. furosemide and bumetanide) and thiazide diuretics (e.g. bendroflumethiazide, hydrochlorothiazide) should be given in patients with fluid overload. Although diuretics provide symptomatic relief of dyspnoea and improve exercise tolerance, there is limited evidence that they affect survival. In severe heart failure patients the combination of a loop and thiazide diuretic may be required. Serum electrolytes and renal function must be monitored regularly (risk of hypokalaemia and hypomagnesaemia).

Angiotensin-converting enzyme inhibitors ACEI

(Fig. 18.27, p. 1023)

Trials have shown that in addition to producing considerable symptomatic improvement in patients with symptomatic heart failure, prognosis is markedly improved and development of heart failure is slowed. ACEI also benefit patients with asymptomatic heart failure following myocardial infarction. Thus ACEI improve survival in patients in all functional classes (NYHA I–IV) and are recommended in all patients at risk of developing heart failure. The main adverse effects of ACE inhibitors are cough, hypotension, hyperkalaemia and renal dysfunction. The contraindications to their use include renal artery stenosis, pregnancy and previous angioedema. In patients with heart failure ACEI should be introduced gradually with a low initial dose and gradual titration with regular blood pressure monitoring. Serum creatinine should be measured concomitantly; potassium-sparing diuretics should be discontinued.

Angiotensin II receptor antagonists (ARA)

The angiotensin II receptor antagonists (ARA) (candesartan and valsartan) are indicated as second-line therapy in patients intolerant of ACEI. Unlike ACEI they do not affect bradykinin metabolism and do not produce a cough. The CHARM Alternative Trial showed that candesartan reduced the risk of heart failure hospitalization compared to placebo in patients intolerant to ACEI. Other trials (Val HeFT and ELITE II) have assessed other ARA.

Beta-blockers

Beta-blockers have been shown to improve functional status and reduce cardiovascular morbidity and mortality in patients with heart failure. Several trials (CIBIS, CIBIS II, MERIT-HF, COMET, SENOIRS) have assessed the effects of beta-blockers in varying degrees of heart failure. Bisoprolol and carvedilol reduce mortality in any grade of heart failure.

Nebivolol is used in the treatment of stable mild–moderate heart failure in patients over 70 years old.

In patients with heart failure, only evidence-based beta-blockers from successful trials should be used. In patients

Table 13.23 Drugs used in heart failure

Drug	Dose (initial/target)	Precautions
ACE inhibitors/angiotensin II receptor antagonists		Monitor renal function and avoid if baseline serum creatinine >250 μmol/L or baseline BP < 90 mmHg
Captopril	6.25 mg × 3 daily/25–50 mg × 3 daily	
Enalapril	2.5 mg daily/10 mg × 2 daily	
Ramipril	1.25–2.5 mg daily/2.5–5 mg × 2 daily	
Candesartan	4 mg daily/32 mg daily	
Valsartan	80 mg daily/320 mg daily	
Losartan	50 mg daily/100 mg daily	
Beta-blockers		Caution in obstructive airways disease, bradyarrhythmias
Bisoprolol	1.25 mg daily/10 mg daily	Avoid in acute heart failure until cardiovascularly stable
Carvedilol	3.125 mg daily/50 mg daily	
Nebivolol	1.25 mg daily/10 mg daily	
Diuretics		Monitor renal function and check for hypokalaemia and hypomagnesaemia
Furosemide	20–40 mg daily/max. 250–500 mg daily	
Bumetanide	0.5–1.0 mg daily/max. 5–10 mg daily	
Bendroflumethiazide	2.5 mg daily/max. 10 mg daily	
Metolazone	2.5 mg daily/max. 10 mg daily	For severe heart failure
Aldosterone antagonists		Monitor renal function, check for hyperkalaemia, gynaecomastia with spironolactone
Spironolactone	12.5–25 mg daily/max. 50–200 mg daily (lower if on ACE inhibitor)	
Eplerenone	25 mg daily/50 mg daily	
Cardiac glycosides		
Digoxin	0.125–0.25 mg daily (reduce dose in elderly or if renal impairment)	Caution in renal impairment, conduction disease and if on amiodarone
Vasodilators		
Isosorbide dinitrate	20–40 mg × 3 daily	
Hydralazine	37.5–75 mg × 3 daily	

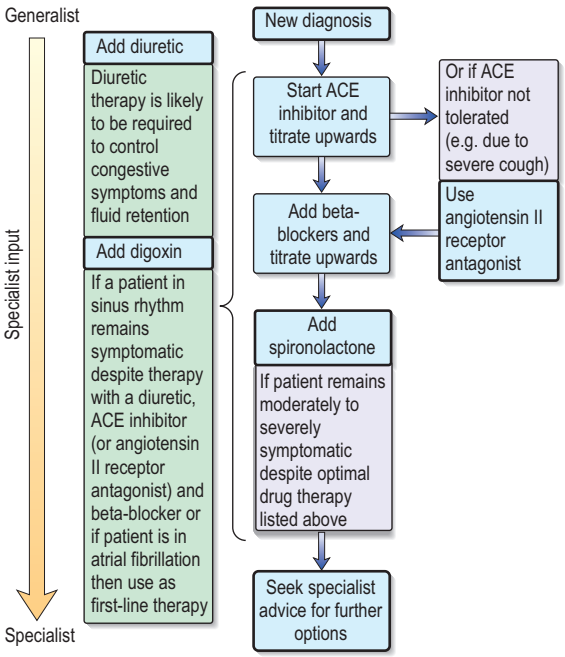


Fig. 13.60 Heart failure treatment guidelines. From NICE. 'Algorithm for the pharmacological treatment of symptomatic heart failure due to left ventricular systolic dysfunction' in CG5 chronic heart failure: management of chronic heart failure in adults in primary and secondary care. London: NICE, 2003. Available from www.nice.org.uk/CG005. Reproduced with permission.

with significant heart failure, beta-blockers are started at a low dose and gradually increased.

Aldosterone antagonists

The aldosterone antagonists spironolactone and eplerenone have been shown to improve survival in patients with heart failure. In RALES, spironolactone reduced total mortality by 30% in patients with severe heart failure. However, gynaecomastia or breast pain occurred in 1 in 10 men taking spironolactone. In EPHESUS, eplerenone given to patients with an acute myocardial infarction and heart failure reduced total mortality by 15% and sudden cardiac death by 21%, with no gynaecomastia.

Cardiac glycosides

Digoxin is a cardiac glycoside that is indicated in patients in atrial fibrillation with heart failure. It is used as add-on therapy in symptomatic heart failure patients already receiving ACEI and beta-blockers. The DIG study demonstrated that digoxin reduced hospital admissions in patients with heart failure.

Vasodilators and nitrates

The combination of hydralazine and nitrates reduces afterload and pre-load and is used in patients intolerant of ACEI or ARA. The Veterans Administration Cooperative Study demonstrated that the combination of hydralazine (with nitrates) improved survival in patients with chronic heart failure. The A-HeFT trial showed that the same combination reduced mortality and hospitalization for heart failure in black patients with heart failure.

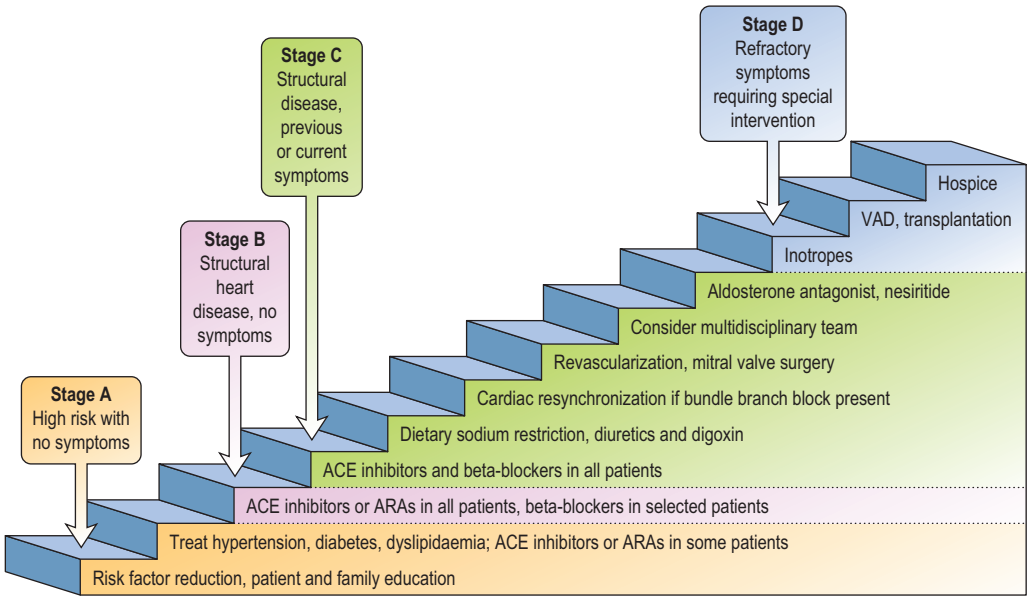


Fig. 13.61 **Stages of heart failure and treatment options for systolic heart failure**. ARA, angiotensin II receptor antagonist; ACE, angiotensin-converting enzyme; VAD, ventricular assisted device. From Jessop M, Brozana S. *New England Journal of Medicine* 2003; **348**: 2013, with permission. Copyright © 2003 Massachusetts Medical Society. All rights reserved.

Inotropic and vasopressor agents

Intravenous inotropes and vasopressor agents (see Table 13.23) are used in patients with chronic heart failure who are not responding to oral medication. Although they produce haemodynamic improvements they have not been shown to improve long-term mortality when compared with placebo.

Other medications

In hospital all patients require prophylactic anticoagulation. Heart failure is associated with a four-fold increase in the risk of a stroke. Oral anticoagulants are recommended in patients with atrial fibrillation and in patients with sinus rhythm with a history of thromboembolism, left ventricular thrombus or aneurysm. In patients with known ischaemic heart disease antiplatelet therapy (aspirin, clopidogrel) and statin therapy should be continued. Arrhythmias are frequent in heart failure and are implicated in sudden death. Although treatment of complex ventricular arrhythmias might be expected to improve survival, there is no evidence to support this and it may increase mortality. In the Sudden Cardiac Death Heart Failure Trial (SCD-HeFT), amiodarone showed no benefit compared to placebo in patients with impaired left ventricular function and mild–moderate heart failure (whereas an ICD reduced mortality by 23% compared to placebo). Patients with heart failure and symptomatic ventricular arrhythmias should be assessed for suitability for an ICD.

Ivabradine, which inhibits the I_f channels in the sincatrial node, slows the heart rate and this may be beneficial in heart failure.

Non-pharmacological treatment of heart failure

Revascularization

While coronary artery disease is the most common cause of heart failure, the role of revascularization in patients with heart failure is unclear. Patients with angina and left ventricular dysfunction have a higher mortality from surgery (10–20%), but have the most to gain in terms of improved symptoms and prognosis. Factors that must be considered before recommending surgery include age, symptoms and evidence for reversible myocardial ischaemia.

Hibernating myocardium and myocardial stunning

'Hibernating' myocardium can be defined as reversible left ventricular dysfunction due to chronic coronary artery disease that responds positively to inotropic stress and indicates the presence of viable heart muscle that may recover after revascularization. It is due to reduced myocardial perfusion, which is just sufficient to maintain viability of the heart muscle. Myocardial hibernation results from repetitive episodes of cardiac stunning that occur, for example, with repeated exercise in a patient with coronary artery disease.

Myocardial stunning is reversible ventricular dysfunction that persists following an episode of ischaemia when the blood flow has returned to normal, i.e. there is a mismatch between flow and function.

The prevalence of hibernating myocardium in patients with coronary artery disease can be estimated from the frequency of improvement in regional abnormalities in wall motion after revascularization and is estimated to be 33% of such patients. Techniques to try to identify hibernating myocardium include stress echocardiography, nuclear imaging techniques, cardiovascular magnetic resonance and positron emission tomography.

The clinical relevance of the hibernating and stunned myocardium is that ventricular dysfunction due to these mechanisms may be wrongly ascribed to myocardial necrosis and scarring which seems untreatable, whereas reversible hibernating and stunning respond to coronary revascularization.

Biventricular pacemaker or implantable cardioverter–defibrillator (p. 732) (Fig. 13.62)

Pacemakers are indicated in patients with sinoatrial disease and atrioventricular conduction block. Pacemakers are also valuable in patients without AV block but with prolonged PR intervals, left bundle branch block and severe mitral regurgitation. In patients with heart failure and left bundle branch block and NYHA 3 heart failure (MUSTIC study) biventricular pacing is more beneficial than conventional right ventricular pacing. The results showed an improvement in symptoms and exercise tolerance. The effect on prognosis is being addressed by an ongoing trial (CARE HF).

Biventricular pacing can also be used in patients not responding to therapy in the following situations:

- systolic heart failure
- non-reversible cause
- highly symptomatic (New York Heart Association grade 3/4)
- on optimal medical therapy
- ventricular dys-synchrony: left bundle branch block (QRS >120 ms)
- sinus rhythm, and possibly atrial fibrillation
- significant mitral regurgitation.

New guidelines recommend advanced pacing technologies should be used for cardiac resynchronization therapy (CRT) in selected patients with heart failure. Atrio-biventricular pacing has been shown to improve symptoms and reduce admission to hospital in patients with left bundle branch block and heart failure with poor LV function. The use of combined implantable defibrillators and atrio-biventricular pacemakers for patients with heart failure is likely to increase. One large trial (COMPANION) shows 21% reduction in all-cause mortality and all-cause admissions to hospital in the groups receiving biventricular pacing. ICD therapy in patients with NYHA (Grades 1 and 2, see Table 13.2, p. 688) has shown a reduction in mortality compared with placebo and amiodarone.

Cardiac transplantation

Cardiac transplantation has become the treatment of choice for younger patients with severe intractable heart failure, whose life-expectancy is less than 6 months. With careful recipient selection, the expected 1-year survival for patients following transplantation is over 90%, and is 75% at 5 years.

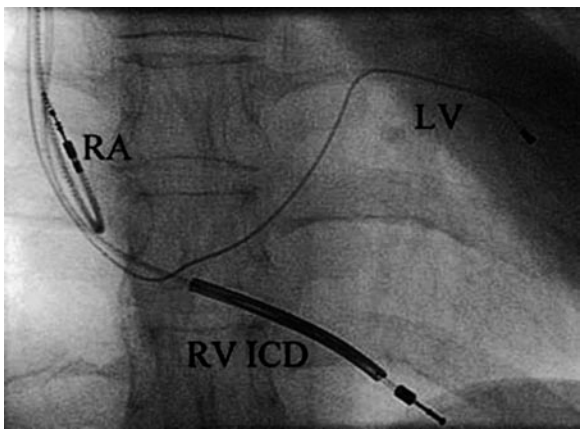


Fig. 13.62 Implanted biventricular pacing with ICD device.

Irrespective of survival, quality of life is dramatically improved for the majority of patients. The availability of heart transplantation is limited.

Heart allografts do not function normally. Cardiac denervation results in a high resting heart rate, loss of diurnal blood pressure variation and impaired renin–angiotensin–aldosterone regulation. Some patients develop 'stiff heart' syndrome, caused by rejection, denervation and ischaemic injury during organ harvest and implantation. Transplantation of an inappropriately small donor heart can also result in elevated right and left heart pressure.

The complications of heart transplantation are summarized in Table 13.24. Many (infection, malignancy, hypertension and hyperlipidaemia) are related to immunosuppression. Allograft coronary atherosclerosis is the major cause of long-term graft failure and is present in 30–50% of patients at 5 years. It is due to a 'vascular' rejection process in conjunction with hypertension and hyperlipidaemia.

There are specific contraindications to cardiac transplantation (Table 13.25); notably, high pulmonary vascular resistance is an absolute contraindication. Several options to transplantation are available: cardiomyoplasty (augmentation of left ventricular contraction by wrapping a latissimus dorsi muscle flap around the ventricle) and the Batista procedure (surgical ventricular size reduction and remodelling the geometry of the left ventricle). Both procedures have a high mortality and limited evidence of substantial benefit in the medium term.

Table 13.24 Complications of cardiac transplantation

Allograft rejection	**Allograft vascular disease**
'Humoral'	**Hypertension**
'Vascular'	**Hypercholesterolaemia**
'Cell-mediated'	**Malignancy**
Infections	
Early: nosocomial organisms – staphylococci, Gram-negative bacteria	
Late (2–6 months): opportunistic (toxoplasmosis, cytomegalovirus, fungi, *Pneumocystis*)	

Table 13.25 Contraindications for cardiac transplantation

Age > 60 years (some variations between centres)
Alcohol/drug abuse
Uncontrolled psychiatric illness
Uncontrolled infection
Severe renal/liver failure
High pulmonary vascular resistance
Systemic disease with multiorgan involvement
Treated cancer in remission but with <5 years' follow-up
Recent thromboembolism
Other disease with a poor prognosis

Acute heart failure

Acute heart failure (AHF) occurs with the rapid onset of symptoms and signs of heart failure secondary to abnormal cardiac function, causing elevated cardiac filling pressures. This causes severe dyspnoea and fluid accumulates in the interstition and alveolar spaces of the lung (pulmonary oedema). AHF has a poor prognosis with a 60-day mortality rate of nearly 10% and a rate of death or rehospitalization of 35% within 60 days. In patients with *acute pulmonary oedema* the in-hospital mortality rate is 12% and by 12 months this rises to 30%. Poor prognostic indicators include a high (>16 mmHg) pulmonary capillary wedge pressure (PCWP), low serum sodium concentration, increased left ventricular end-diastolic dimension on echo and low oxygen consumption.

The aetiology of AHF is similar to chronic heart failure:

- Patients with *ischaemic heart disease* present with an acute coronary syndrome or develop a complication of a myocardial infarct, e.g. papillary muscle rupture or ventricular septal defect requiring surgical intervention.
- Patients with *valvular heart disease* also present with AHF due to valvular regurgitation in endocarditis or prosthetic valve thrombosis. A thoracic aortic dissection may produce severe aortic regurgitation.
- Patients with *hypertension* present with episodes of 'flash' pulmonary oedema despite preserved left ventricular systolic function.
- In both *acute and chronic kidney disease* fluid overload and a reduced renal excretion will produce pulmonary oedema.
- *Atrial fibrillation* is frequently associated with AHF and may require emergency cardioversion.

Several clinical syndromes of AHF can be defined (Table 13.26). In a clinical environment both the Killip score (p. 758, based on a cardiorespiratory clinical assessment) and Forrester (based on right catheterization findings) classification are used to provide therapeutic and prognostic information.

Table 13.26 Clinical syndromes of acute heart failure (AHF)

Type	Clinical features
Acute decompensated heart failure	Mild features of heart failure, e.g. dyspnoea
Hypertensive AHF	High blood pressure, preserved left ventricular function, pulmonary oedema on CXR
Acute pulmonary oedema	Tachypnoea, orthopnoea, pulmonary crackles, oxygen saturation <90% on air, pulmonary oedema on CXR
Cardiogenic shock	Systolic blood pressure <90 mmHg, mean arterial pressure drop >30 mmHg, urine output <0.5 mL/kg/hour, heart rate >60 b.p.m.
High output heart failure	Warm peripheries, pulmonary congestion, blood pressure may be low, e.g. septic shock
Right heart failure	Low cardiac output, elevated jugular venous pressure, hepatomegaly, hypotension

Modified from the European Society of Cardiology.

Pathophysiology

The pathophyisology of AHF is similar to chronic heart failure with activation of the renin–angiotensin–aldosterone axis and sympathetic nervous system. In addition, prolonged ischaemia (e.g. in acute coronary syndromes) results in myocardial stunning (p. 740) that exacerbates myocardial dysfunction but may respond to inotropic support. If myocardial ischaemia persists the myocardium may exhibit hibernation (persistently impaired function due to reduced coronary blood flow) that may recover with successful revascularization.

Diagnosis

In a patient presenting with symptoms and signs of heart failure a structured assessment should result in the clinical diagnosis of AHF and direct initial treatment to stabilize the patient. Initial investigations performed in the emergency room should include:

- a *12-lead ECG* for acute coronary syndromes, left ventricular hypertrophy, atrial fibrillation, valvular heart disease, left bundle branch block
- a *chest X-ray* (cardiomegaly, pulmonary oedema, pleural effusion, non-cardiac disease)
- *blood investigations* (serum creatinine and electrolytes, full blood count, blood glucose, cardiac enzymes and troponin, CRP and D-dimer)
- *plasma BNP* or *NTproBNP* (BNP >100 pg/ml or NTproBNP >300 pg/ml) indicates heart failure
- transthoracic echocardiography should be performed without delay to confirm the diagnosis of heart failure (p. 704) and possibly identify the cause.

If the baseline investigations confirm AHF then treatment should be commenced.

Treatment

The goals of treatment in a patient with AHF include:

- immediate relief of symptoms and stabilization of haemodynamics (short-term benefits)
- reduction in length of hospital stay and hospital readmissions
- reduction in mortality from heart failure.

Patients with AHF should be managed in a high-care area with regular measurements of temperature, heart rate, blood pressure and cardiac monitoring. All patients require prophylactic anticoagulation with low molecular weigh heparin, e.g. enoxaparin 1 mg/kg s.c. daily.

Patients with haemodynamic compromise may require arterial lines (invasive blood pressure monitoring and arterial gases), central venous cannulation (intravenous medication, inotropic support, monitoring of central venous pressure) and pulmonary artery cannulation (calculation of cardiac output/index, peripheral vasoconstriction and pulmonary wedge pressure).

Initial therapy (Table 13.23) includes oxygen, diuretics (e.g. I.V. furosemide 50 mg) and vasodilator therapy (e.g. glyceryl trinitrate infusion 10–200 μg/min) if the blood pressure is maintained (systolic >85 mmHg). Inotropic support (see p. 907) with dobutamine, phosphodiesterase inhibitors or levosimendan can be added in patients who do not

FURTHER READING

Braunwald E. Biomarkers in heart failure. *New England Journal of Medicine* 2008; **358**: 2148–2158.

Neubauer S. The failing heart – an engine out of fuel. *New England Journal of Medicine* 2007; **356**: 1140–1151.

Nieminen MS, Böhm M, Cowie MR, et al. Executive summary of the guidelines on the diagnosis and treatment of acute heart failure: the Task Force on Acute Heart Failure of the European Society of Cardiology. *European Heart Journal* 2005; **26**: 384–416.

Swedberg K, Cleland J, Dargie H, et al. Guidelines for the diagnosis and treatment of chronic heart failure: executive summary (update 2005): The Task Force for the Diagnosis and Treatment of Chronic Heart Failure of the European Society of Cardiology. *European Heart Journal* 2005; **26**: 1115–1140.

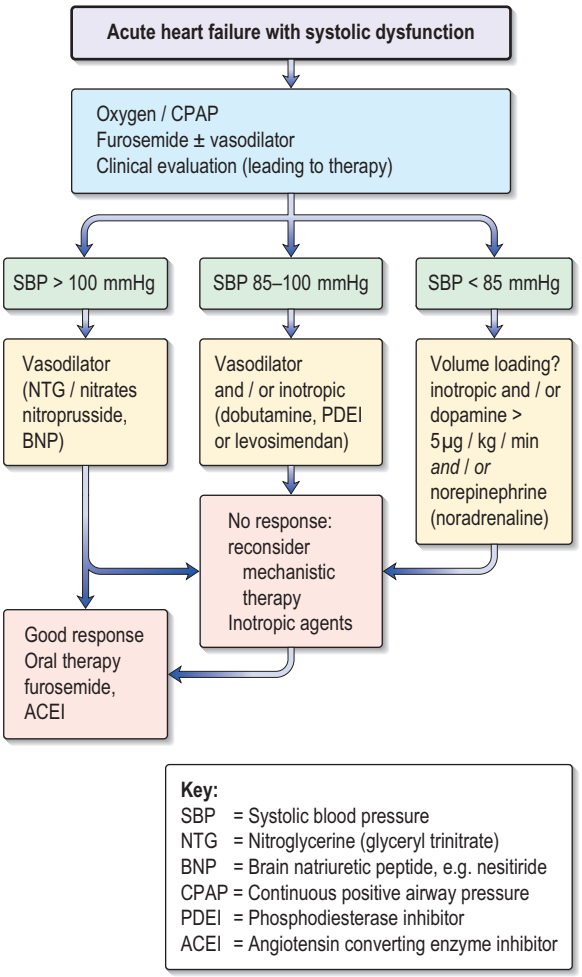


Fig. 13.63 **Algorithm for the management of acute heart failure with systolic dysfunction.** ACEI, angiotensin-converting enzyme inhibitor; BNP, brain natriuretic peptide, e.g. nesiritide; CPAP, continuous positive airway pressure; NTG, nitroglycerine; PDEI, phosphodiesterase inhibitor; SBP, systolic blood pressure. From Nieminen MS, Böhm M, Cowie MR et al. *European Heart Journal* 2005; **26**: 384–416, Figure 6, with permission from Oxford University Press and the European society of Cardiology.

respond to the initial therapy (Fig. 13.63). If blood pressure is low use noradrenaline (norepinephrine).

Patients with profound hypotension may require inotropes and vasopressors to improve the haemodynamic status and alleviate symptoms, but these have not been shown to improve mortality. Recently, non-invasive CPAP/NIPPV (see p. 916) has been shown to provide earlier improvement in dyspnoea and respiratory distress than standard oxygen via masks; mortality is however unaffected.

Mechanical assist devices

Mechanical assist devices can be used in patients who fail to respond to standard medical therapy but in whom there is either transient myocardial dysfunction with likelihood of recovery (e.g. post anterior myocardial infarction treated with coronary angioplasty) or a bridge is needed to cardiac surgery, including transplantation.

Ventricular assist devices (VAD) (Fig. 13.64). Ventricular assist devices are mechanical devices that replace or help the failing ventricles in delivering blood around the body. A

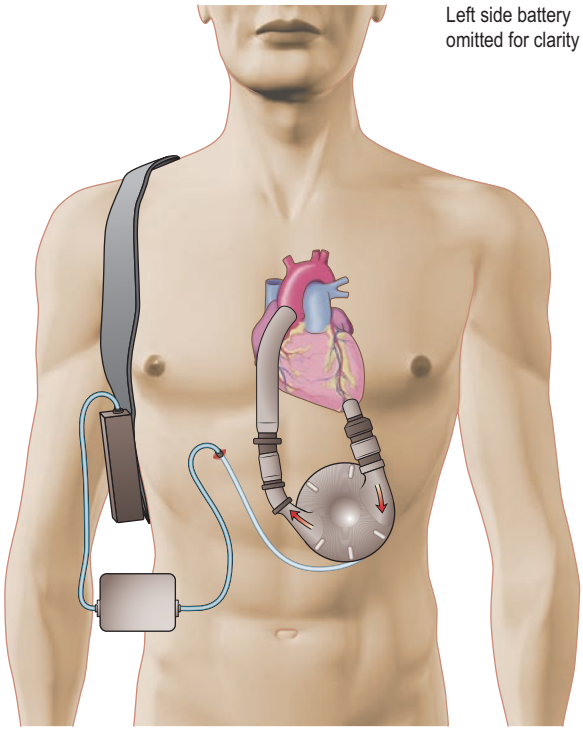


Fig. 13.64 **Left ventricular assist device.** From Moser DK, Riegel B. *Cardiac Nursing.* Philadelphia: Saunders, 2007: 981, with permission of Elsevier.

left ventricular assist device (LVAD) receives blood from the left ventricle and delivers it to the aorta; a right ventricular assist device (RVAD) receives blood from the right ventricle and delivers it to the pulmonary artery. The devices can be extracorporeal (suitable for short-term support) or intracorporeal (suitable for long-term support as a bridge to transplantation or as destination therapy in patients with end-stage heart failure not candidates for transplantation). The main problems with VADs include thromboembolism, bleeding, infection and device malfunction.

ISCHAEMIC HEART DISEASE

Myocardial ischaemia occurs when there is an imbalance between the supply of oxygen (and other essential myocardial nutrients) and the myocardial demand for these substances.

Coronary blood flow to a region of the myocardium may be reduced by a mechanical obstruction that is due to:

- atheroma
- thrombosis
- spasm
- embolus
- coronary ostial stenosis
- coronary arteritis (e.g. in SLE).

There can be a decrease in the flow of oxygenated blood to the myocardium that is due to:

- anaemia
- carboxyhaemoglobulinaemia

- hypotension causing decreased coronary perfusion pressure.

An increased demand for oxygen may occur owing to an increase in cardiac output (e.g. thyrotoxicosis) or myocardial hypertrophy (e.g. from aortic stenosis or hypertension).

Myocardial ischaemia most commonly occurs as a result of obstructive coronary artery disease (CAD) in the form of coronary atherosclerosis. In addition to this fixed obstruction, variations in the tone of smooth muscle in the wall of a coronary artery may add another element of dynamic or variable obstruction.

CAD is the largest single cause of death in the UK and many parts of the world. However, over the last decade, the mortality rate in the UK has fallen considerably. Each year there are approximately 60 deaths per 100 000 (giving a standardized mortality rate of about 200 per 100 000). It has been estimated that by 2010, 60% of the world's heart disease is expected to occur in India. Sudden cardiac death is a prominent feature of CAD. One in every six coronary attacks present with sudden death as the first, last and only symptom.

Process of coronary atherosclerosis

Coronary atherosclerosis is a complex inflammatory process characterized by the accumulation of lipid, macrophages and smooth muscle cells in intimal plaques in the large and medium-sized epicardial coronary arteries. The vascular endothelium plays a critical role in maintaining vascular integrity and homeostasis. Mechanical shear stresses (e.g. from morbid hypertension), biochemical abnormalities (e.g. elevated and modified LDL, diabetes mellitus, elevated plasma homocysteine), immunological factors (e.g. free radicals from smoking), inflammation (e.g. infection such as *Chlamydia pneumoniae* and *Helicobacter pylori*) and genetic alteration may contribute to the initial endothelial 'injury' or dysfunction, which is believed to trigger atherogenesis.

The *development of atherosclerosis* follows the endothelial dysfunction, with increased permeability to and accumulation of oxidized lipoproteins, which are taken up by macrophages at focal sites within the endothelium to produce lipid-laden foam cells. Macroscopically, these lesions are seen as flat yellow dots or lines on the endothelium of the artery and are known as '*fatty streaks*'. The 'fatty streak' progresses with the appearance of extracellular lipid within the endothelium ('*transitional plaque*'). Release of cytokines such as platelet-derived growth factor and transforming growth factor-β (TGF-β) by monocytes, macrophages or the damaged endothelium promotes further accumulation of macrophages as well as smooth muscle cell migration and proliferation. The proliferation of smooth muscle with the formation of a layer of cells covering the extracellular lipid separates it from the adaptive smooth muscle thickening in the endothelium. Collagen is produced in larger and larger quantities by the smooth muscle and the whole sequence of events cumulates as an '*advanced or raised fibrolipid plaque*'. The 'advanced plaque' may grow slowly and encroach on the lumen or become unstable, undergo thrombosis and produce an obstruction ('*complicated plaque*').

Two different mechanisms are responsible for thrombosis on the plaques (Fig. 13.65).

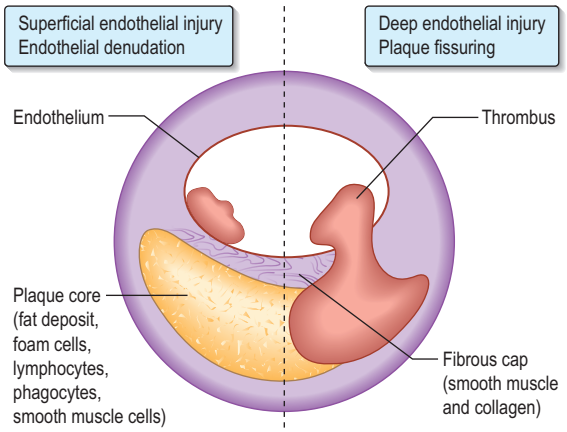


Fig. 13.65 **The mechanisms for the development of thrombosis on plaques.**

- The *first process* is superficial endothelial injury, which involves denudation of the endothelial covering over the plaque. Subendocardial connective tissue matrix is then exposed and platelet adhesion occurs because of reaction with collagen. The thrombus is adherent to the surface of the plaque.
- The *second process* is deep endothelial fissuring, which involves an advanced plaque with a lipid core. The plaque cap tears (ulcerates, fissures or ruptures), allowing blood from the lumen to enter the inside of the plaque itself. The core with lamellar lipid surfaces, tissue factor (which triggers platelet adhesion and activation) produced by macrophages and exposed collagen, is highly thrombogenic. Thrombus forms within the plaque, expanding its volume and distorting its shape. Thrombosis may then extend into the lumen.

A 50% reduction in luminal diameter (producing a reduction in luminal cross-sectional area of approximately 70%) causes a haemodynamically significant stenosis. At this point the smaller distal intramyocardial arteries and arterioles are maximally dilated (coronary flow reserve is near zero), and any increase in myocardial oxygen demand provokes ischaemia.

CAD gives rise to a wide variety of clinical presentations, ranging from relatively stable angina through to the acute coronary syndromes of unstable angina and myocardial infarction (Fig. 13.66). Figure 13.67 shows an actual plaque rupture.

Risk factors for coronary artery disease – primary and secondary prevention

CAD is an atherosclerotic disease that is multifactorial in origin, giving rise to the risk factor concept. Certain living habits promote atherogenic traits in genetically susceptible persons. A number of 'risk' factors are known to predispose to the condition (Table 13.27). Some of these, such as age, gender, race and family history, cannot be changed, whereas other major risk factors, such as serum cholesterol, smoking habits, diabetes and hypertension, can be modified.

Atherosclerotic disease manifest in one vascular bed is often advanced in other territories. Patients with intermittent

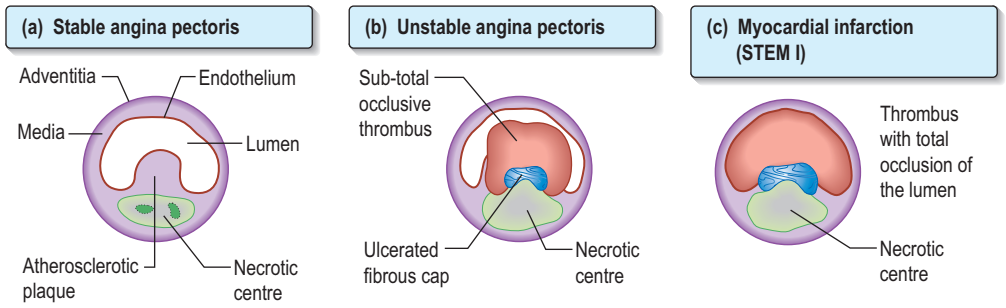


Fig. 13.66 **The mechanisms for the development of thrombosis on plaques. Relationship between the state of coronary artery vessel wall and clinical syndrome. (a)** Stable angina pectoris. **(b)** and **(c)** Acute coronary syndromes.

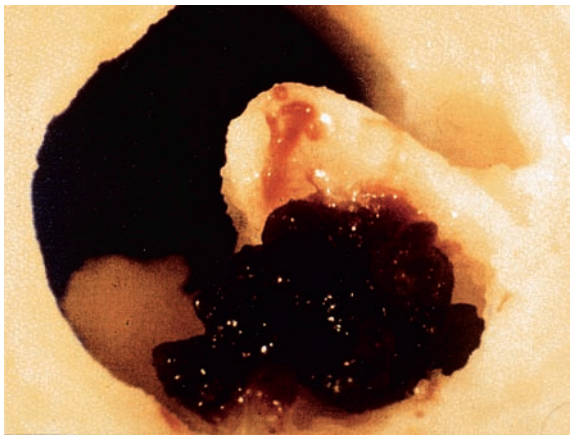


Fig. 13.67 **Acute coronary thrombus.** Cross-section (×30) of the epicardial coronary artery, demonstrating a rupture of the shoulder region of the plaque with a luminal thrombus.

Table 13.27 Risk factors for coronary disease

Fixed
Age
Male sex
Positive family history
Deletion polymorphism in the angiotensin-converting enzyme (ACE) gene (DD)
Potentially changeable
Hyperlipidaemia
Cigarette smoking
Hypertension
Diabetes mellitus
Lack of exercise
Blood coagulation factors – high fibrinogen, factor VII
C-reactive protein
Homocysteinaemia
Personality
Obesity
Gout
Soft water
Drugs, e.g. contraceptive pill, nucleoside analogues, COX-2 inhibitors, rosiglitazone
Heavy alcohol consumption

claudication have a two- to four-fold increased risk of CAD, stroke, or heart failure. Following initial myocardial infarction (MI), there is a three- to six-fold increase in the risk of heart failure and stroke. After stroke, the risk of heart failure and MI is increased twofold.

The disease can be asymptomatic in its most severe form, with one in three myocardial infarctions going unrecognized. Thirty to 40% of individuals who present with an acute coronary syndrome have had no prior warning symptom to suggest the presence of underlying disease.

Primary prevention can be defined as the prevention of the atherosclerotic disease process and *secondary prevention* as the treatment of the atherosclerotic disease process (i.e. treatment of the disease or its complications). The objective of prevention is to reduce the incidence of first or recurrent clinical events due to CAD, ischaemic stroke and peripheral artery disease.

Traditional risk factors

Age

CAD rates increase with age. Atherosclerosis is rare in childhood, except in familial hyperlipidaemia, but is often detectable in young men between 20 and 30 years of age. It is almost universal in the elderly in the West. Atheromatous lesions in the elderly are often complicated by calcification.

Gender

Men have a higher incidence of coronary artery disease than premenopausal women. However, after the menopause, the incidence of atheroma in women approaches that in men. The reasons for this gender difference are not clearly understood, but probably relate to the loss of the protective effect of oestrogen.

Family history

CAD is often found in several members of the same family. Because the disease is so prevalent and because other risk factors are familial, it is uncertain whether family history, per se, is an independent risk factor. A positive family history is generally accepted to refer to those in whom a first-degree relative has developed ischaemic heart disease before the age of 50 years.

Smoking (see p. 829)

In men, the risk of developing CAD is directly related to the number of cigarettes smoked. It is estimated that about 20% of deaths from CAD in men and 17% of deaths from CAD in women are due to smoking. Evidence suggests that each

person stopping smoking will reduce his/her own risk by 25%. The risk from smoking declines to almost normal after 10 years of abstention.

Diet and obesity (see p. 228)

Diets high in fats are associated with ischaemic heart disease, as are those with low intakes of antioxidants (i.e. fruit and vegetables). Supplementation with antioxidants has been shown to be unhelpful in RCTs (p. 224).

It is estimated that up to 30% of deaths from CAD are due to unhealthy diets. The dietary changes which would help to reduce rates of CAD include a reduction in fat, particularly saturated fat intake, a reduction in salt intake and an increase in carbohydrate intake. The consumption of fruit and vegetables should be increased by 50% to about 400 g per day, which is equivalent to at least five daily portions (see Box 5.2).

There is overwhelming evidence from clinical trials that modification of the diet has a significant impact on the risk of CVD in both the primary and secondary prevention settings.

Weight. Patients who are overweight and those who are obese have an increased risk of CAD. It is estimated that about 5% of deaths from CAD in men and that 6% of such deaths in women are due to obesity (a body mass index (BMI) of greater than 30 kg/m^2).

The adverse effect of excess weight is more pronounced when the fat is concentrated mainly in the abdomen. This is known as central obesity (visceral fat) and can be identified by a high waist/hip ratio.

Exercise. Reduction in weight by diet and exercise not only lowers the incidence of CVD but also diabetes/insulin resistance. It is estimated that about 36% of deaths from CAD in men and 38% of deaths from CAD in women are due to lack of physical activity. To produce the maximum benefit the activity needs to be regular and aerobic. Aerobic activity involves using the large muscle groups in the arms, legs and back steadily and rhythmically so that breathing and heart rate are significantly increased.

It is recommended that adults should participate in a minimum of 30 minutes of at least moderate intensity activity (such as brisk walking, cycling or climbing the stairs) on 5 or more days of the week.

Hypertension

Both systolic and diastolic hypertension are associated with an increased risk of CAD. Both drug treatment and lifestyle changes – particularly weight loss, an increase in physical activity and a reduction in salt and alcohol intake – can effectively lower blood pressure.

It is estimated that 14% of deaths from CAD in men and 12% of deaths from CAD in women are due to a raised blood pressure (defined as a systolic blood pressure of 140 mmHg or over, or a diastolic blood pressure of 90 mmHg or over) and that 6% of deaths from CAD in the UK could be avoided if the numbers of people who have high blood pressure were to be reduced by 50%.

Hyperlipidaemia (see p. 1063)

High serum cholesterol, especially when associated with a low value of high-density lipoproteins (HDL), is strongly associated with coronary atheroma. There is increasing evidence that high serum triglyceride (TG) is also independently linked with coronary atheroma.

Familial hypercholesterolaemia combined with hypertriglyceridaemia and remnant hyperlipidaemia are also associated with increased risk of coronary atherosclerosis.

Measurement of the fasting lipid profile (total cholesterol, low- and high-density lipoproteins and triglycerides) should be performed on all people with an increased risk of cardiac disease.

The risk of CAD is directly related to serum cholesterol levels. Serum cholesterol levels can be reduced by drugs, physical activity and by dietary changes, in particular a reduction in the consumption of saturated fat. It is estimated that 45% of deaths from CAD in men and 47% of deaths from CAD in women are due to a raised serum cholesterol level (in this case >5.2 mmol/L) and that 10% of deaths from CAD in the UK could be avoided if everyone in the population had a serum cholesterol level of less than 6.5 mmol/L.

Different guidelines give slightly different advice for managing high levels of serum cholesterol (hyperlipidaemia). The National Service Framework for coronary heart disease in England includes guidelines on the prevention of CAD in clinical practice and suggests a cholesterol target of less than 5.0 mmol/L for both primary and secondary prevention.

High-density lipoprotein cholesterol (HDL-cholesterol) is the fraction of cholesterol that removes cholesterol (via the liver) from the blood. Low levels of HDL-cholesterol are associated with an increased risk of CAD and a worse prognosis after a heart attack. Guidelines on HDL-cholesterol generally recommend treatment for those with concentrations below 1.0 mmol/L. HDL increases with exercise, alcohol in moderation, not smoking and when TG is lowered.

A 1% reduction in cholesterol levels reduces risks of CAD by 2–3%. Hyperlipidaemia can be treated as follows:

- *Statins*: 24–30% reduction in mortality in primary and secondary prevention will be achieved if a statin (pravastatin or simvastatin) is given. Up to 50% reduction is achieved if the dose of statin (e.g. atorvastatin) is titrated to achieve a target LDL of <2.6 mmol/L.
- *Fibrates* result in a significant reduction in CAD events in diabetics and patients with high TG and low HDL.
- *Diet*: the so-called Mediterranean diet (p. 210) has resulted in a 75% reduction in CAD events in post-myocardial infarction patients.

Angiographic studies have shown that lowering the serum cholesterol can slow the progression of coronary atherosclerosis, and can cause regression of disease. Large clinical trials have shown that lipid lowering, usually with a statin, can decrease total mortality and new coronary events, and reduce the need for revascularization. Management of hypercholesterolaemia is described in detail on page 1065.

Diabetes mellitus

Diabetes, an abnormal glucose tolerance or raised fasting glucose is strongly associated with vascular disease.

Diabetes substantially increases the risk of CAD. Men with type 2 diabetes have a two- to four-fold greater annual risk of CAD, with an even higher (three- to five-fold) risk in women with type 2 diabetes.

Diabetes not only increases the risk of CAD but also magnifies the effect of other risk factors for CAD such as raised cholesterol levels, raised blood pressure, smoking and obesity.

Other risk factors

Although there is general agreement on established cardiovascular risk factors, epidemiological research continues to identify or evaluate additional risk factors that contribute to the occurrence of atherosclerotic CVD and warrant further clarification.

Sedentary lifestyle

Lack of exercise is an independent risk factor for CAD equal to hypertension, hyperlipidaemia and smoking. Regular exercise probably protects against its development (see above).

Psychosocial well-being

Four different types of psychosocial factors have been found to be most consistently associated with an increased risk of CAD: work stress, lack of social support, depression (including anxiety) and personality (particularly hostility).

Alcohol

Moderate alcohol consumption (one or two drinks per day) is associated with a reduced risk of CAD. At high levels of intake – particularly in 'binges' – the risk of CAD is increased. It is currently advised that 'regular consumption of between three and four units a day by men' and 'between two and three units a day by women of all ages will not lead to any significant health risk'.

Genetic factors

A number of genetic factors have been linked with coronary artery disease. The angiotensin-converting enzyme (ACE) gene contains an insertion/deletion (I/D) polymorphism, the DD genotype of which has been associated with a predisposition to CAD and myocardial infarction.

Lipoprotein (a)

High plasma Lp(a) concentrations are associated with CAD and, although probably not an independent risk factor, elevated plasma Lp(a) increases the CAD risk associated with more traditional risk factors.

Coagulation factors

Serum fibrinogen is strongly, consistently and independently related to CAD risk. The pathophysiological mechanism by which fibrinogen levels mediate coronary disease risk is related to its effect on the coagulation cascade, platelet aggregation, endothelial function and smooth muscle cell proliferation and migration.

High levels of *coagulation factor VII* are also a risk factor. Polymorphisms of the factor VII gene may increase the risk of myocardial infarction.

Homocysteine, an amino acid regulated by vitamins B_{12}, B_6 and folate, is another factor that has been associated with CAD and atherosclerosis (see p. 224). Homocysteinaemia is a major risk factor in the pathogenesis of CAD and a strong predictor of mortality in this group. Plasma levels of homocysteine are influenced by a variety of genetic and non-genetic factors. The mechanism associating hyperhomocysteinaemia with atherosclerosis is its adverse effect on vascular endothelium. Folic acid in low doses may ameliorate this process.

C-reactive protein (CRP)

CRP is linked with future risk of coronary events independently of the traditional risk factors but its use as a marker for subclinical atherosclerosis and cardiovascular risk has been questioned.

Non-steroidal anti-inflammatory drugs (NSAIDs)

NSAIDs that are specific inhibitors of cyclo-oxygenase-2 (COX-2) have been shown to increase cardiovascular risk and mortality. Rofecoxib has already been withdrawn.

Prevention policy

The priorities for CVD prevention in clinical practice are:

- Patients with established CAD, PVD and cerebrovascular atherosclerotic disease.
- Asymptomatic individuals who are at high risk of developing atherosclerotic disease because of multiple risk factors resulting in a 10-year risk of >5% now (or if extrapolated to age 60) for developing a fatal event, i.e. those with markedly raised levels of single risk factors:
  - cholesterol >8 mmol/L
  - LDL cholesterol >6 mmol/L
  - BP >180/110 mmHg.
- All patients with diabetes.
- Close relatives of:
  - patients with early-onset atherosclerotic cardiovascular disease
  - asymptomatic individuals at a particular high risk.
- Other individuals encountered in routine clinical practice.

How to estimate total cardiovascular risk in asymptomatic people as a guide to prevention strategies

Patients with established CVD have declared themselves to be at high total risk of further vascular events. Therefore they require the most intense lifestyle intervention, and where appropriate drug therapies.

However, in the majority of asymptomatic, apparently healthy people, preventative actions should be guided in accordance with the total CVD risk level. Indeed, risk factor management decisions should usually not be based on considerations of a single modestly raised factor.

To evaluate candidates for the major cardiovascular events cost-effectively, multivariate risk profiles have been formulated; these facilitate targeting those at high risk for preventative measures.

The Joint British Societies, the European Society of Cardiology and the American Heart Association have emphasized the importance of these risk profiles for motivating as well as reassuring patients and in assisting in selecting therapy. They concluded that these scores direct healthcare professionals to look at the whole patient and to recognize the cumulative nature of risk factors (Fig. 13.68). However, not all practitioners agree with this approach (see pp. 748 and 1067).

National Service Framework (NSF)

Plans to reduce CAD-related death in the under-75-year age group by 40% by the year 2010 by the implementation of set standards have been published in the UK by the Department of Health. The NSF includes a nurse-led audited approach to reduce CAD by lowering saturated fat intake, increasing exercise and, most relevant, decreasing/stopping smoking. The hypertension treatment targets are 140/85 mmHg in patients at risk of or with established coronary artery disease

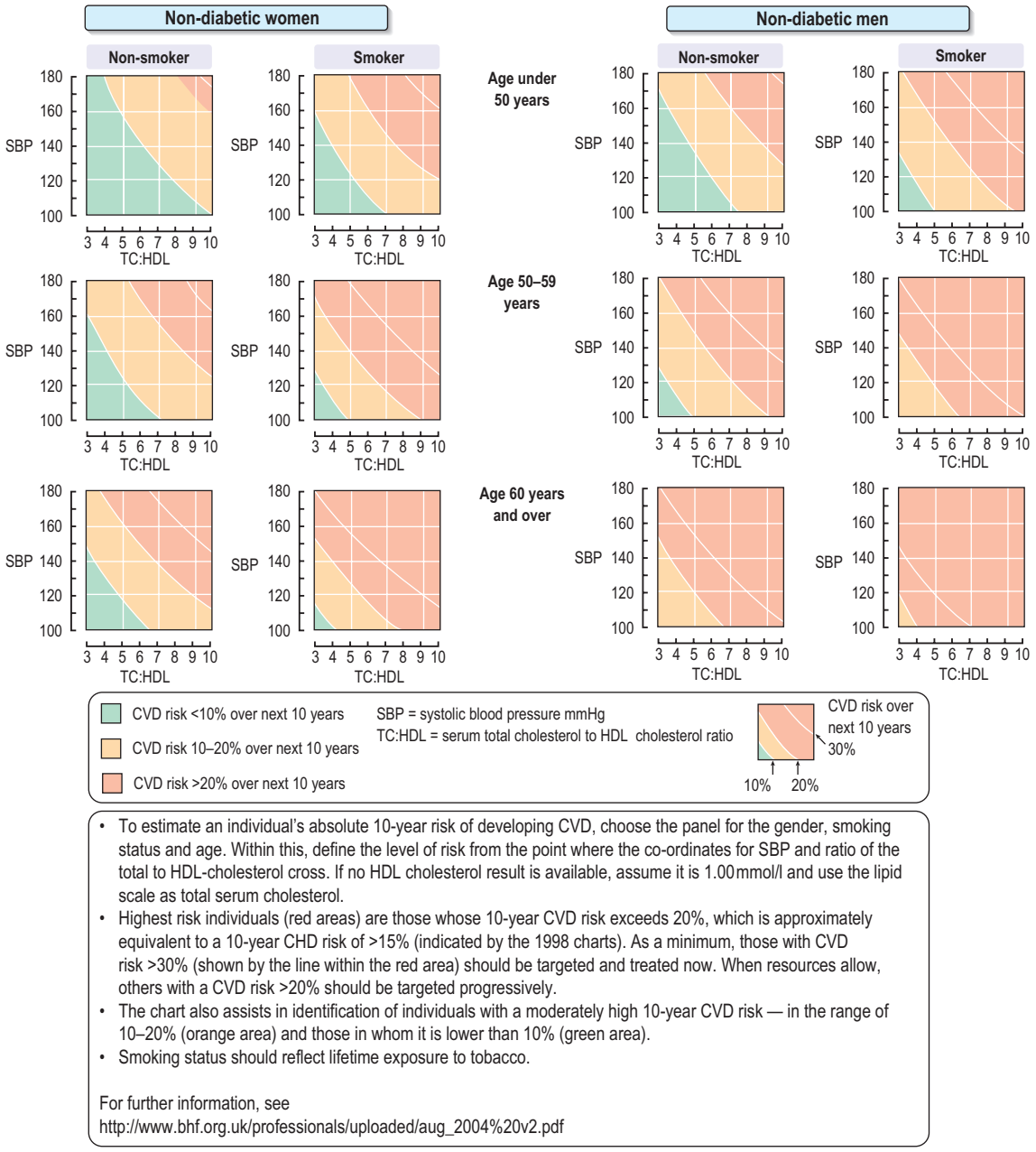


Fig. 13.68 **(a) Cardiovascular risk prediction charts for women. (b) Cardiovascular risk prediction charts for men.** Cardiovascular Risk Prediction Charts reproduced with permission from the University of Manchester Department of Medical Illustration.

and 130/80 mmHg in diabetics. The cholesterol target is either total cholesterol of <5.0 mmol/L (LDL-cholesterol <3 mmol/L) or a reduction of 30% (whichever is greater).

ANGINA (see also p. 688)

The diagnosis of angina is largely based on the clinical history. The chest pain is generally described as 'heavy', 'tight' or 'gripping'. Typically, the pain is central/retrosternal and may radiate to the jaw and/or arms. Angina can range from a mild ache to a most severe pain that provokes sweating and fear. There may be associated breathlessness. CAD is common, fatal and largely preventable. More than 1.4 million people in the UK suffer from angina. CAD accounts for about 3% of all hospital admissions in England. The prevalence of angina is approximately 2% with an incidence of new cases each year of approximately 1 per 1000.

Classical or exertional angina pectoris is provoked by physical exertion, especially after meals and in cold, windy weather, and is commonly aggravated by anger or excitement. The pain fades quickly (usually within minutes) with rest. Occasionally it disappears with continued exertion ('walking through the pain'). Whilst in some patients the pain occurs predictably at a certain level of exertion, in most patients the threshold for developing pain is variable.

Decubitus angina is angina that occurs on lying down. It usually occurs in association with impaired left ventricular function, as a result of severe coronary artery disease.

Nocturnal angina occurs at night and may wake the patient from sleep. It can be provoked by vivid dreams. It tends to occur in patients with critical coronary artery disease and may be the result of vasospasm.

Variant (Prinzmetal's) angina refers to an angina that occurs without provocation, usually at rest, as a result of coronary artery spasm. It occurs more frequently in women. Characteristically, there is ST segment elevation on the ECG during the pain. Specialist investigation using provocation tests (e.g. hyperventilation, cold-pressor testing or ergometrine challenge) may be required to establish the diagnosis. Arrhythmias, both ventricular tachyarrhythmias and heart block, can occur during the ischaemic episode.

Cardiac syndrome X refers to those patients with a good history of angina, a positive exercise test and angiographically normal coronary arteries. They form a heterogeneous group, and the syndrome is much more common in women than in men. Whilst they have a good prognosis, they are often highly symptomatic and can be difficult to treat. In women with this syndrome the myocardium shows an abnormal metabolic response to stress, consistent with the suggestion that the myocardial ischaemia results from abnormal dilator responses of the coronary microvasculature to stress. The prognostic and therapeutic implications are not known.

Unstable angina refers to angina of recent onset (<1 month), worsening angina or angina at rest, and will be discussed below in more detail in the section entitled 'Acute coronary syndrome' (see p. 752).

Examination and diagnosis

There are usually no abnormal findings in angina, although occasionally a fourth heart sound may be heard. Signs to suggest anaemia, thyrotoxicosis or hyperlipidaemia (e.g. lipid arcus, xanthelasma, tendon xanthoma) should be sought. It is essential to exclude aortic stenosis (i.e. slow-rising carotid impulse and ejection systolic murmur radiating to the neck) as a possible cause for the angina. The blood pressure should be taken to identify coexistent hypertension.

Investigations for angina

Resting ECG

This is usually normal between attacks. Evidence of old myocardial infarction (e.g. pathological Q waves), left ventricular hypertrophy or left bundle branch block may be present. During an attack, transient ST depression, T wave inversion or other changes of the shape of the T wave may appear.

Exercise ECG

Exercise testing can be very useful both in confirming the diagnosis of angina and in giving some indication as to the severity of the CAD. ST segment depression of ≥1 mm suggests myocardial ischaemia, particularly if typical chest pain occurs at the same time (Fig. 13.21).

The test has a specificity of 80% and a sensitivity of about 70% for CAD. A strongly positive test (within 6 minutes of starting the Bruce protocol) suggests 'prognostic' disease and helps to identify patients who should be offered coronary intervention. Exercise testing, however, can be misleading:

- A normal test does not exclude CAD (so-called false-negative test) although these patients, as a group, have a good prognosis.
- Up to 20% of patients with positive exercise tests are subsequently found to have no evidence of coronary artery disease (so-called false-positive test).

Cardiac scintigraphy

Myocardial perfusion scans (see p. 705), both at rest and after stress (i.e. exercise or dobutamine), are helpful. Redistribution of the contrast agent is a sensitive indicator of ischaemia and can be particularly useful in deciding if a stenosis seen at angiography is giving rise to ischaemia. A normal perfusion scan makes significant CAD unlikely.

Echocardiography

This can be used to assess ventricular wall involvement and ventricular function. Regional wall motion abnormalities at rest reflect previous ventricular damage. Stress echocardiography, although technically difficult, is useful, especially in women with coronary artery disease.

CT coronary angiography (see p. 706)

This is increasingly being used to diagnose coronary artery disease and exclude other causes, e.g. pulmonary embolism.

Cardiovascular magnetic resonance (MR)

This is being increasingly used and provides excellent details of the coronary anatomy (see Fig. 13.34, p. 708).

Coronary angiography

This is occasionally useful in patients with chest pain where the diagnosis is unclear. More often, the test is performed to delineate the exact coronary anatomy in patients being considered for revascularization (i.e. coronary artery bypass grafting or coronary angioplasty). Coronary angiography should be performed only when the benefit in terms of diagnosis and potential treatment outweighs the small risk of the procedure (a mortality rate of <1 in 1000 cases).

Lesions with complex morphology (irregular borders, overhanging edges, thrombus or ulceration) appear to identify a subgroup of stenoses associated with disease progression and adverse clinical outcomes.

Treatment of angina

General management

Patients should be informed as to the nature of their condition and reassured that the prognosis is good (annual mortality <2%). Underlying problems, such as anaemia or hyperthyroidism, should be treated. Management of coexistent conditions, such as diabetes and hypertension, should be optimized. Risk factors should be evaluated and steps made to correct them where possible; for example, smoking must be stopped, hypercholesterolaemia should be identified and treated (see below), weight loss, where appropriate, and regular exercise should be encouraged.

Choosing between medical therapy and revascularization (coronary artery bypass grafting and angioplasty) can be difficult and will depend on a number of factors including symptoms, angiographic anatomy and patient/physician preference. The various treatment options are not mutually exclusive and should be considered as complementary.

Medical treatment

Prognostic therapies

Aspirin reduces the risk of coronary events in patients with coronary artery disease. All patients with angina, therefore, should take aspirin (75 mg daily is probably adequate) unless contraindicated.

Lipid-lowering therapy (see p. 1065) should be used in patients with total cholesterol above 4.8 mmol/L (particularly if the LDL is >3.3 mmol/L and the HDL is <1.0 mmol/L), despite a low fat diet. If the triglycerides (TGs) are under 3.5 mmol/L, then one of the statins (HMG-CoA reductase inhibitors) should be used. If the TGs are above 3.5 mmol/L, a fibrate is indicated. If simple therapy fails to reduce the LDL adequately, then the patient should be referred to a lipidologist. Lipid-lowering therapy can be expected to prevent 20–30 deaths or myocardial infarcts per 1000 patient-years.

Hormone replacement therapy (HRT) is of no value in prevention of CAD.

Symptomatic treatment

Glyceryl trinitrate (GTN) used sublingually, either as a tablet or as a spray, gives prompt relief (peak action 4–8 minutes and lasts 20–30 minutes). It can be used prior to performing activities that the patient knows will provoke angina. Transdermal GTN preparations last up to 24 hours.

All but the most mildly affected patients will probably require regular prophylactic therapy. The choice of drugs is between beta-blockers, nitrates and calcium-channel blockers. There is no commonly accepted algorithm and treatment needs to be tailored to the individual patient. Some patients will require combination therapy, but there is little evidence that adding a third drug is of benefit. Patients not controlled adequately on medical therapy should be considered for revascularization (see below).

Beta-blockers reduce the heart rate (negative chronotropic effect) and the force of ventricular contraction (negative inotropic effect), both of which reduce myocardial oxygen demand, especially on exertion. They are the drugs of choice in patients with previous myocardial infarction because of their proven benefit in secondary prevention. Atenolol, 50–100 mg daily, is the most commonly prescribed. Metoprolol, 25–50 mg twice daily, is often used if renal function is impaired. Beta-blockers may aggravate coronary artery spasm.

Long-acting nitrates (e.g. isosorbide mononitrate) are particularly useful in patients who gain relief from sublingual GTN. They reduce venous return and hence intracardiac diastolic pressures, reduce the impedance to the emptying of the left ventricle and relax the tone of the coronary arteries. Once-daily preparations are available which have a smooth pharmacokinetic profile and avoid the problem of tolerance. Nitrates should be given with care to patients on other hypotensive agents. Sildenafil (or other PDE5 inhibitors) should not be given to patients taking nitrates.

Calcium-channel blockers block calcium flux into the cell and the utilization of calcium within the cell. They relax coronary arteries, cause peripheral vasodilatation and reduce the force of left ventricular contraction, thereby reducing the oxygen demand of the myocardium. The non-dihydropyridine calcium antagonists (e.g. diltiazem and verapamil) also reduce the heart rate and are particularly useful antianginal agents, but should be used with caution in combination with beta-blockers. Short-acting dihydropyridines (e.g. nifedipine) can cause reflex tachycardia when used alone. Case-control studies have suggested that high-dose nifedipine is associated with adverse outcome. Slow-release formulations and the third-generation agents (e.g. amlodipine) can be used once daily and have a smooth profile of action with no significant effect on the heart rate and no significant negative inotropic effect.

Nicorandil is a potassium-channel activator with a nitrate component; it has both arterial and venous vasodilating properties. Whilst not used as a first-line drug, it is used when there are contraindications to the above agents and in refractory unstable angina.

Ranolazine is a novel agent that interacts with sodium channels and can improve exercise tolerance and reduce the frequency of angina attacks in patients with ischaemic heart disease. However, ranolazine causes QT interval prolongation and should be reserved for patients who do not respond to other antianginal drugs (long-acting nitrates, calcium-channel blockers and beta-blockers).

Ivabradine is a selective and specific inhibitor of the cardiac pacemaker I_f current (see p. 682) that controls the spontaneous diastolic depolarization in the sinus node and regulates heart rate. It has recently been licensed for the symptomatic treatment of chronic stable angina pectoris in patients with normal sinus rhythm who have a contraindication or intolerance to beta-blockers. It is particularly helpful in a subgroup with heart rates >70 beats/min.

Percutaneous coronary intervention

Percutaneous transluminal coronary angioplasty (PTCA) is the process of dilating a coronary artery stenosis using an inflatable balloon introduced into the arterial circulation via the femoral, radial or brachial artery (Fig. 13.69). PTCA is an effective treatment that can reduce angina symptoms, reduce the requirement for antianginal drug therapy and increase exercise capacity. Coronary artery anatomy determines the success of PTCA. A discrete, soft lesion in a straight vessel without involving a bifurcation has the best outcome. Unfavourable lesions are occluded vessels, stenoses that are calcified, tortuous, long, or involve a bifurcation. The risks of PTCA include acute myocardial infarction (2%) and death (1%). The benefits of PTCA were initially reduced by coronary restenosis, but with the introduction of metallic stents (Fig. 13.70) the effectiveness increased. The STRESS and BENESTENT studies demonstrated superiority of percutaneous coronary intervention (PCI) with stenting over PTCA alone, with reduced need for repeat revascularization. Peripheral vascular complications were more frequent after PCI because of the use of heparin and warfarin therapy post stent insertion. The bleeding complications were significantly reduced with the use of ticlopidine and then clopidogrel as alternative antithrombotic agents to formal anticoagulation. Further improvements to patient outcome post PCI have occurred with the addition of the antiplatelet GPIIb/IIIa antagonists (tirofiban, eptifibatide and abciximab) that inhibit platelet function and reduce acute stent thrombosis.

A further development in PCI was the introduction of coated stents which are lined with substances that reduce coronary artery restenosis and offer the prospect of long-term vessel patency post PCI, especially in narrow vessels with long stenoses. The *Cypher stent* contains sirolimus, which is an immunosuppressant agent that reduces cellular proliferation. Biolimus (a lipophilic sirolimus analogue) eluting stents are being trialled; the drug dissolves completely

Ask the authors
www.studentconsult.com

(a) (b) (c)

Fig. 13.69 **Percutaneous transluminal coronary angioplasty (PTCA). (a)** Coronary angiography demonstrates a severe stenosis in the proximal left anterior descending artery. **(b)** During PTCA a soft guidewire is passed across the stenosis and then a balloon is expanded that dilates the stenosis. **(c)** Post-PTCA.

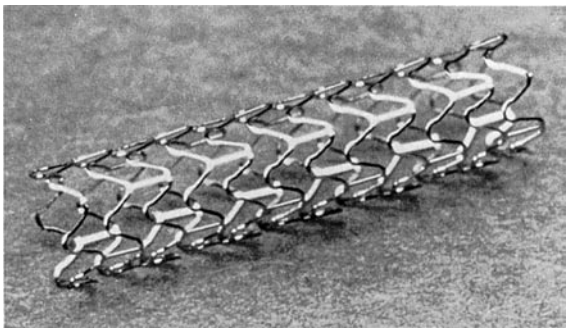


Fig. 13.70 **An intracoronary stent.**

leaving a bare-metal stent in 9 months. The *Taxus stent* contains paclitaxel which is a mitotic inhibitor drug that inhibits neointima formation. Both stents have demonstrated superiority over 'bare-metal' stenting in reducing coronary artery restenosis and the need for repeat revascularization at 6–12 months. However, recent reports have raised concerns about late-stent thrombosis (>6 months post insertion) in patients with drug-eluting stents leading to acute myocardial infarction and frequently death. It has been suggested that inadequate endothelialization of the stent leads to exposure of thrombus stimulating surface when the patient discontinues clopidogrel therapy, leading to recommendations that patients take prolonged dual therapy (aspirin and clopidogrel) and avoid discontinuing therapy within 6–12 months of implantation.

Coronary artery bypass grafting

With coronary artery bypass grafting (CABG) autologous veins or arteries are anastamosed to the ascending aorta and to the native coronary arteries distal to the area of stenosis (Fig. 13.71). Improved graft survival can be obtained with in-situ internal mammary and gastroepiploic arteries grafted onto the stenosed coronary artery. Three major randomized controlled trials compared CABG with medical therapy: the Coronary Artery Surgery Study (CASS), the Veterans Administration (VA) Co-operative Study and the European Coronary Surgery Study (ECSS). A meta-analysis has been performed that demonstrated that compared to medical therapy, CABG significantly improved angina symptoms, exercise capacity and reduced the need for antianginal therapy. In addition, CABG improves 10-year survival in patients with angina, and particularly patients with >50% left main stem stenosis and

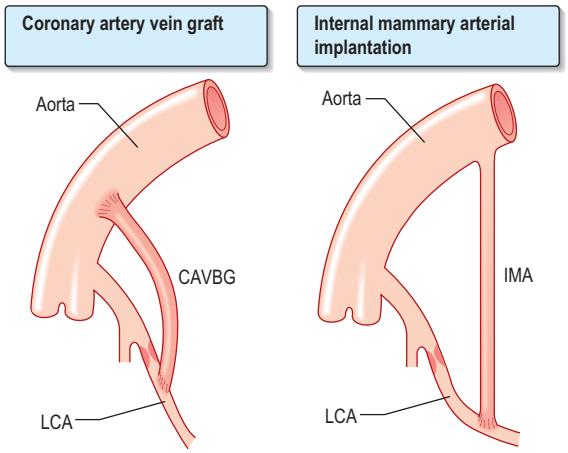


Fig. 13.71 **Relief of coronary obstruction by surgical techniques:** coronary artery vein bypass grafting (CAVBG) or internal mammary arterial implantation (IMA). In both of these examples, the graft bypasses a coronary obstruction in the left coronary artery (LCA).

triple-vessel disease with impaired left ventricular function. Operative mortality is well below 1% in patients with normal left ventricular function. Perioperative strokes occur in up to 2% of cases, and more subtle neurological deficits are common. Off-pump coronary surgery is now performed; results show that it is as safe as on-pump surgery and causes less myocardial damage, but the graft patency rate is lower. Minimally invasive operative procedures for bypass grafting ('MIDCAB') are being developed, including laparoscopic approaches, and may be of use in certain subgroups of patients (e.g. previous CABG and those with coexistent medical conditions which would increase the operative risks of 'full' CABG).

CABG versus PCI

Comparative trials between CABG and PCI have now been performed. In the multicentre ERACI II study, PCI patients had lower major adverse events (death, myocardial infarction, repeat revascularization procedures and stroke) in the first 30 days compared with the CABG patients (3.6% versus 12.3%) with a long-term follow-up (mean 18 months), patient survival was 96.9% in the PCI group versus 92.5% in the CABG group. However, the need for repeat revascularization was significantly greater in the PCI group than in the CABG group (16.8% versus 4.8%).

In another multicentre study, SoS (Stent or Surgery), the incidence of death or Q-wave myocardial infarction was similar for both groups (PCI 9% versus 10% CABG) at follow-up at a median of 2 years. However, there were fewer deaths in the CABG group (2% versus 5%) and there was less need for repeat revascularization (6% versus 21%). The choice of CABG versus PCI is influenced by coronary artery anatomy, left ventricular function, diabetes and patient choice. In asymptomatic patients revascularization cannot improve symptoms but will improve prognosis.

Patients with intractable angina

Some patients remain symptomatic despite medication and are not suitable for (further) revascularization.

- *Transmyocardial laser revascularization* (TMR), whereby a laser is used to form channels in the myocardium to allow direct perfusion of the myocardium from blood within the ventricular cavity, has been used in some centres but in controlled trials it has not been beneficial.
- *Spinal cord stimulation* (SCS) is accomplished using a flexible electrode in the epidural space at the mid-thoracic level. Stimulation via a pacemaker-like generator reduces angina and has been shown to reduce indices of ischaemia.
- Similarly, *transcutaneous electrical nerve stimulation* has been shown to reduce angina and increase exercise capacity in selected patients.

ACUTE CORONARY SYNDROMES

Acute coronary syndromes (ACS) include ST-elevation myocardial infarction (STEMI), non-ST-elevation myocardial infarction (NSTEMI) and unstable angina. Myocardial infarction (MI) occurs when cardiac myocytes die due to myocardial ischaemia, and can be diagnosed on the basis of appropriate clinical history, 12-lead ECG and elevated biochemical markers – troponin I and T, CK-MB. STEMI will be covered in the next section (p. 755).

FURTHER READING

Eagle KA, Lim MJ, Dabbous OH, et al, for the GRACE investigators. A validated prediction model for all forms of acute coronary syndrome. Estimating the risk of 6-month post-discharge death in an international registry. *Journal of the American Medical Association* 2004; **291**: 2727–2733.

Pathophysiology

The common mechanism to all ACS is rupture or erosion of the fibrous cap of a coronary artery plaque. This leads to platelet aggregation and adhesion, localized thrombosis, vasoconstriction and distal thrombus embolization. The presence of a rich lipid pool within the plaque and a thin fibrous cap are associated with an increased risk of rupture. Thrombus formation and the vasoconstriction produced by platelet release of serotonin and thromboxane A_2, results in myocardial ischaemia due to reduction of coronary blood flow.

Diagnosis

Clinical presentation

Patients with an ACS may complain of a new onset of *chest pain*, chest pain at rest, or a deterioration of pre-existing angina. However, some patients present with atypical features including indigestion, pleuritic chest pain or dyspnoea. Physical examination can detect alternative diagnoses such as aortic dissection, pulmonary embolism or peptic ulceration. In addition it can also detect adverse clinical signs such as hypotension, basal crackles, fourth heart sounds and cardiac murmurs.

Electrocardiogram

Although the 12-lead ECG may be normal in patients with an ACS, ST depression and T wave inversion are highly suggestive for an ACS, particularly if associated with anginal chest pain. The ECG should be repeated when the patient is in pain, and continuous ST-segment monitoring is recommended. With a STEMI, complete occlusion of a coronary vessel will result in persistent ST-elevation or left bundle branch block pattern, although transient ST elevation is seen with coronary vasospasm or Prinzmetal's angina.

Biochemical markers

- The measurement of the ***creatinine-kinase-MB*** level was until recently the standard marker for myocyte death used in ACS. However, the presence of low levels of CK-MB in the serum of normal individuals and in patients with significant skeletal muscle damage, has limited its accuracy. It can be used to determine reinfarction as levels drop back to normal after 36–72 hours.
- The ***cardiac troponin*** complex is made up of three distinct proteins (I, T and C) that are situated with tropomyosin on the thin actin filament that forms the skeleton of the cardiac myofilament. Troponin T attaches the complex to tropomyosin, troponin-C binds calcium during excitation–contraction coupling, and troponin-I inhibits the myosin binding site on the actin.

  The cardiac troponins are not detectable in normal people and so monoclonal antibody tests to cardiac-specific troponin I and cardiac-specific troponin T are highly sensitive markers of myocyte necrosis. If the initial troponin assay is negative, then it should be repeated 9–12 hours after admission. The troponin assay has prognostic information that can determine mortality risk in ACS (Fig. 13.72) and define which patients may benefit from aggressive medical therapy and early coronary revascularization.
- ***Myoglobin*** may be useful for a rapid diagnosis of an ACS as the levels become elevated very early in the time course of an MI, but because of the presence of myoglobin in skeletal muscle the test has poor specificity for ACS.

Risk stratification

Initial risk in ACS is determined by complications of the acute thrombosis. This may produce recurrent myocardial ischaemia, marked ST depression, dynamic ST changes, a raised troponin level and be demonstrated with coronary angiography.

Long-term risk is defined by clinical risk factors; age, prior myocardial infarction or bypass surgery, diabetes or heart failure. Biological markers such as C-reactive protein, fibrinogen, brain natriuretic peptide, modified albumin and serum creatinine, can be used to further stratify patient risk. Left ventricular dysfunction and the presence of left main or triple vessel disease significantly increase the future cardiovascular risk. Both the Thrombolysis In Myocardial Infarction (TIMI) (Table 13.28) score and the Global Registry of Acute Coronary Events (GRACE) prediction score can be used in patients with ACS to define risk. TIMI is shown in Table 13.28. The GRACE score is based on age, heart rate, systolic blood pressure, serum creatinine and the Killip score.

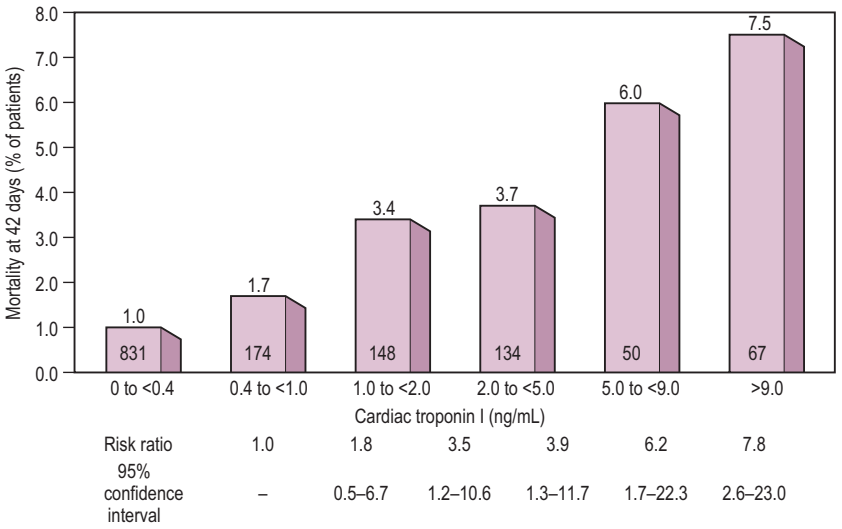


Fig. 13.72 **Relationship between cardiac troponin I levels and risk of death in patients with the acute coronary syndrome (ACS).** Numbers within histograms are the number of patients in each group. From Antman EM, Tanasijevic MJ, Thompson B et al. Cardiac-specific troponin I levels to predict the risk of mortality in patients with acute coronary syndromes. *New England Journal of Medicine* 1996; **335**: 1342–1349. Copyright © 1996 Massachusetts Medical Society. All rights reserved.

Table 13.28 The TIMI risk score in acute coronary syndrome (NSTEMI/UA)

Risk factor	Score
Age >65	1
More than three coronary artery disease risk factors – hypertension, hyperlipidaemia, family history, diabetes, smoking	1
Known coronary artery disease (coronary angiography stenosis >50%)	1
Aspirin use in the last 7 days	1
Severe angina (more than two episodes of rest pain in 24 hours)	1
ST deviation on ECG (horizontal ST depression or transient ST elevation >1 mm)	1
Elevated cardiac markers (CK–MB or troponin)	1

Total score	Rate of death/MI in 14 days (%)	Rate of death /MI/urgent revascularization (%)
0–1	3	4.75
2	3	8.3
3	5	13.2
4	7	19.9
5	12	26.2
6–7	19	40.9

Investigation and treatment

All patients require immediate management of their chest pain as outlined on p. 755 and in Table 13.29.

High-risk patients for progression to myocardial infarction or death require urgent coronary angiography. These patients include those with persistent or recurrent angina with ST changes ≥2 mm or deep negative T wave changes, clinical signs of heart failure or haemodynamic instability, life-threatening arrhythmias (VF, VT).

Patients with *immediate or high-risk* TIMI or GRACE scores, elevated troponins, dynamic ST or T wave changes, diabetes mellitus, renal dysfunction, reduced left ventricular function, early post-infarction angina, previous myocardial infarction, PCI within 6 months, or previous CABG, should have early (<72 hours) coronary angiography and interventions.

Low-risk patients can be managed with oral aspirin, clopidogral, beta-blockers and nitrates. These include patients with no recurrence of chest pain during observation, no signs of heart failure, normal ECG or minor T wave changes on arrival and at 6–12 hours, normal troponins on the initial assays and at 6–12 hours post admission. An exercise test should be performed – a negative result has a good prognosis and an early positive test should direct the patient to an invasive strategy. If the patient is unable to exercise satisfactorily, or if the baseline ECG is abnormal (e.g. left ventricular hypertrophy or LBBB), then dobutamine stress echocardiography or myocardial perfusion scintigraphy are recommended.

Antiplatelet agents

The platelet is a key part of the thrombosis cascade involved in ACS. Rupture of the atheromatous plaque exposes the circulating platelets to ADP (adenosine diphosphate), thromboxane A_2 (TxA_2), epinephrine (adrenaline), thrombin and collagen tissue factor. This causes platelet activation, with thrombin as an especially potent stimulant of such activity. Platelet activation stimulates the expression of glycoprotein (GP) IIb/IIIa receptors on the platelet surface. These receptor bridge fibrinogen between adjacent platelets, causing platelet aggregates (Fig 8.41).

Aspirin blocks the formation of thromboxane A_2 and so prevents platelet aggregation. In ACS patients 75–150 mg aspirin reduced the relative risk of death or myocardial infarction by about 35–50%.

Ticlopidine and clopidogrel are thienopyridines that inhibit ADP-dependent activation of the GPIIb/IIIa complex that allows platelet aggregates to form. Both ticlopidine and clopidogrel can be used in patients allergic to aspirin, but ticlopidine causes more systemic side-effects including neutropenia or thrombocytopenia. In the CURE study of 12562 ACS patients, 9 months of 75 mg clopidogrel reduced the primary end-point of cardiovascular death, myocardial infarction, or stroke from 11.4% to 9.3% ($P<0.0001$), compared to placebo. Prasugrel is also being used with fewer thrombotic episodes but no change in mortality.

Activated GP (glycoprotein) IIb/IIIa receptors on platelets bind to fibrinogen initiating platelet aggregation. Receptor antagonists have been developed that are powerful inhibitors

Table 13.29 Pharmacological therapy in acute coronary syndrome

Drug	Dose	Notes
Myocardial oxygenation		
Oxygen	35–50%	Check ABG in severe COPD
Antiplatelet		
Aspirin	150–300 mg chewable or soluble aspirin, then 75–10 mg p.o. daily	Caution if active peptic ulceration
Clopidogrel	300 mg p.o. loading dose, then 75 mg p.o. daily	Caution: increased risk of bleeding, avoid if CABG planned
Antithrombin		
Heparin	5000 units i.v. bolus, then 0.25 units/kg/hour	Measure anticoagulant effect with APTT at 6 hours
Low-molecular-weight heparins		
e.g. Enoxaparin	1 mg/kg s.c. × 2 daily	
Bivalirudin	750 μg/kg i.v. bolus, then 0.75 mg/kg/hour	
Glycoprotein IIB/IIIA inhibitors		
Abciximab	0.25 mg/kg i.v. bolus, then 0.125 μg/kg/min up to 10 μg/min i.v. × 12 hours	Indicated if coronary intervention likely within 24 hours
Eptifibatide	180 μg/kg i.v. bolus, then 2 μg/kg/min × 72 hours	Indicated in high-risk patients managed without coronary intervention or during PCI
Tirofiban	0.4 μg/kg/min for 30 minutes, then 0.1 μg/kg/min × 48–108 hours	Indicated in high-risk patients managed without coronary intervention or during PCI
Analgesia		
Diamorphine or morphine	2.5–5.0 mg i.v.	Prescribe with antiemetic, e.g. metoclopramide 10 mg i.v.
Myocardial energy consumption		
Atenolol	5 mg i.v. repeated after 15 minutes, then 25–50 mg p.o. daily	Avoid in asthma, heart failure, hypotension, bradyarrhythmias
Metoprolol	5 mg i.v. repeated to a maximum of 15 mg, then 25–50 mg p.o. × 2 daily	Avoid in asthma, heart failure, hypotension, bradyarrhythmias
Coronary vasodilation		
Glyceryl trinitrate	2–10 mg/hour i.v./buccal/sublingual	Maintain systolic BP >90 mmHg
Plaque stabilization/ventricular remodelling		
HMG-CoA reductase inhibitors (statins)		Combine with dietary advice and modification
Simvastatin	20–40 mg p.o.	
Pravastatin	20–40 mg p.o.	
Atorvastatin	80 mg p.o.	
ACE inhibitors		Monitor renal function
Ramipril	2.5–10 mg p.o.	
Lisinopril	5–10 mg p.o.	

of platelet aggregation. Abciximab is a monoclonal antibody that binds tightly and has a long half-life. Eptifibatide is a cyclic peptide that selectively inhibits GPIIb/IIIa receptors, but has a short half-life and wears off in 2–4 hours. Tirofiban is a small non-peptide that rapidly blocks the GPIIb/IIIa receptors and is reversible in 4–6 hours.

In the GUSTO-IV ACS study of 7800 patients, abciximab was administered but coronary intervention discouraged. At 30 days 8.2 % of abciximab patients and 8.0 % of placebo patients had reached the composite end-point of death or myocardial infarction. In the PRISM study of 3232 patients with angina, tirofiban reduced the 30-day death or myocardial infarction rate from 7.1% with placebo to 5.8%. Troponin-positive patients with diabetes scheduled to have coronary intervention benefit most from GPIIb/IIIa receptor antagonists.

Antithrombins

In ACS patients off aspirin, *unfractionated heparin (UFH)* produces a lower rate of refractory angina/myocardial infarction and death than placebo, and when used with aspirin reduces death and myocardial infarction from 10.3% to 7.9%. However, because of poor bioavailability and variable effects, frequent monitoring of APTT is necessary to ensure therapeutic levels. *Low molecular weight heparins* and in particular enoxaparin appear superior to UFH and can be given subcutaneously twice daily. Bivalirudin is a direct thrombin inhibitor that reversibly binds to thrombin and inhibits clot-bound thrombin. In the ACUITY trial bivalirudin appeared as effective as heparin plus GPIIb/IIIa inhibitors in reducing ischaemic events in patients pretreated with a thienopyridine and undergoing diagnostic angiography or percutaneous intervention, but with less bleeding.

Anti-ischaemia agents

In patients with no contraindications (asthma, AV-block, acute pulmonary oedema), beta-blockers are administered intravenously or orally, to reduce myocardial ischaemia by blocking circulating catecholamines. This will reduce the heart rate and blood pressure, reducing myocardial oxygen consumption. The dose can be titrated to produce a resting heart rate of 50–60 b.p.m. In patients with ongoing angina, nitrates should be given either sublingually or intravenously. They effectively reduce preload and produce coronary vasodilation. However, tolerance can become a problem and patients should be weaned off intravenous administration should if symptoms resolve.

Plaque stabilization/remodelling

HMG-CoA reductase inhibitor drugs (statins) and ACE-inhibitors are routinely administered to patients with ACS. These agents may produce plaque stabilization, improve vascular and myocardial remodelling, and reduce future cardiovascular events. Starting the drugs whilst the patient is still in hospital increases the likelihood of patients receiving secondary drug therapy.

Coronary intervention

Coronary revascularization is recommended in high-risk patients with ACS. Coronary stenting may stabilize the disrupted coronary plaque; in the BENESTENT II trial it was demonstrated that stenting was superior to PTCA in reducing angiographic restenosis rates. Sub-group analysis of patients with unstable angina in the EPIC, EPILOG and CAPTURE trials, confirmed the benefit of GPIIb/IIIa inhibitors at reducing the complication rate during PCI. The PCI-CURE study demonstrated that pretreatment with clopidogrel reduces the rate of cardiovascular death and MI. The current rate of CABG in ACS is low (5.4%). The mortality rates with CABG are greater in the high risk group patients, particularly with a recent myocardial infarction. Single vessel lesions are usually treated with PCI, unless the anatomy is unfavourable. Conversely in patients with left main stem or triple vessel disease with impaired left ventricular function are best managed with surgery. Two studies have compared a conservative versus an invasive strategy in the modern era. In the FRISC-II study 2457 high-risk ACS patients were randomized to PTCA or CABG at 4 and 8 days, respectively, or a conservative approach with intervention only for severe angina. Revascularization within 10 days was performed in 71% of the invasive arm versus 9% of the conservative arm. After 1 year there was a significant reduction in total mortality (2.2% versus 3.9%) in the invasive arm, as well a significant reduction in MI (8.6% versus 11.6%). In addition, the rate of angina or readmission was reduced by 50%. In the TACTICS study of 2220 high-risk ACS patients, similar findings were obtained with the rate of death or MI reduced from 9.5% to 7.3% by an invasive strategy. Patients with a troponin T >0.01 ng/ml obtained benefit, but not those who were troponin T negative.

Post ACS

After the initial management of the ACS (diagnosis, treatment, investigation, revascularization) risk factor modification is achieved to reduce future cardiovascular events:

- Patients should stop cigarette smoking and if necessary be referred to a smoking cessation clinic. In addition, patients should maintain optimal weight, daily exercise and have a healthy diet.
- Hypertension should be treated to a level <130/85.
- In patients with diabetes, glycaemic control should be very tight.
- Low-fat diets should be combined with HMG-CoA reductase inhibitors to reduce LDL cholesterol.
- Medication on discharge should include aspirin, clopidogrel, beta-blocker and an ACE-inhibitor, with a GTN spray for symptomatic relief of angina.

ST ELEVATION MYOCARDIAL INFARCTION (STEMI)

Myocardial infarction occurs when cardiac myocytes die due to prolonged myocardial ischaemia. The diagnosis can be made in patients with an appropriate clinical history together with findings from repeated 12-lead ECGs and elevated biochemical markers – troponin I and T, CK-MB.

Pathophysiology

Rupture or erosion of a vulnerable coronary artery plaque can produce prolonged occlusion of a coronary artery leading to myocardial necrosis within 15–30 minutes. The subendocardial myocardium is initially affected but with continued ischaemia the infarct zone extends through to the subepicardial myocardium, producing a transmural Q-wave myocardial infarction. Early reperfusion may salvage regions of the myocardium, reducing future mortality and morbidity.

The 1-month mortality in patients with a myocardial infarction may be as high as 50% in the community, with 50% of deaths occurring in the first 2 hours of the event. In the pre-thrombolytic era the in-hospital mortality rate was nearly 20% but with modern therapy it may be as low as 6–7% at 1 month. Several risk factors can be identified that predict death rate at 30 days (TIMI STEMI score – Table 13.30).

Diagnosis

Symptoms and signs

Any patient presenting with severe chest pain lasting more than 20 minutes may be suffering from a myocardial infarction. The pain does not usually respond to sublingual GTN, and opiate analgesia is required. The pain may radiate to the left arm, neck or jaw. However, in some patients, particularly elderly or diabetic patients, the symptoms may be atypical and include dyspnoea, fatigue, pre-syncope or syncope. Autonomic symptoms are common and on examination the patient is pale and clammy, with marked sweating. In addition, the pulse is thready with significant hypotension, bradycardia or tachycardia.

Electrocardiography

An ECG in patients with chest pain should be performed on admission to A&E. The baseline ECG is rarely normal, but if so should be repeated every 15 minutes while the patient remains in pain. Continuous cardiac monitoring is required because of the high likelihood of significant cardiac arrhythmias. ECG changes (Table 13.31) are usually confined to the ECG leads that 'face' the infarction. The presence of new ST

Table 13.30 TIMI risk score in ST elevation myocardial infarction (STEMI)

Risk factor	Score
Age >65	2
Age >75	3
History of angina	1
History of hypertension	1
History of diabetes	1
Systolic BP <100	3
Heart rate >100	2
Killip II–IV	2
Weight >67 kg	1
Anterior MI or LBBB	1
Delay to treatment >4 hours	1

Total score	Risk of death at 30 days (%)
0	0.8
1	1.6
2	2.2
3	4.4
4	7.3
5	12.4
6	16.1
7	23.4
8	26.8
9–16	35.9

elevation (due to opening of the K^+ channels) ≥0.2 mV at the J-point in leads V_1–V_3, and ≥0.1 mV in other leads, suggests anterior MI (Fig. 13.73). An inferior wall MI is diagnosed when ST elevation is seen in leads II, III and AVF (Fig. 13.74). Lateral MI produces changes in leads I, AVL and V_5/V_6. In patients with a posterior MI, there may be ST depression in leads V_1–V_3 with a dominant R wave, and ST elevation in lead V_5/V_6. New LBBB or presumed new LBBB is compatible with coronary artery occlusion requiring urgent reperfusion therapy. The evolution of the ECG during the course of STEMI is illustrated in Figure 13.75.

Table 13.31 Typical ECG changes in myocardial infarction (STEMI)

Infarct site	Leads showing ST elevation
Anterior:	
Small	V_3–V_4
Extensive	V_2–V_5
Anteroseptal	V_1–V_3
Anterolateral	V_4–V_6, I, AVL
Lateral	I, AVL
Inferior	II, III, AVF
Posterior	V_1, V_2 (reciprocal)
Subendocardial	Any lead
Right ventricle	VR_4

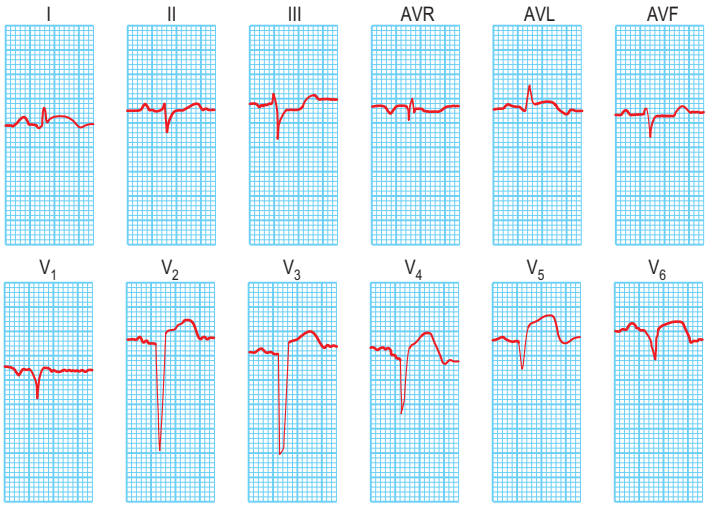


Fig. 13.73 **An acute anterolateral myocardial infarction** shown by a 12-lead ECG. Note the ST segment elevation in leads I, AVL and V_2–V_6. The T wave is inverted in leads I, AVL and V_3–V_6. Pathological Q waves are seen in leads V_2–V_6.

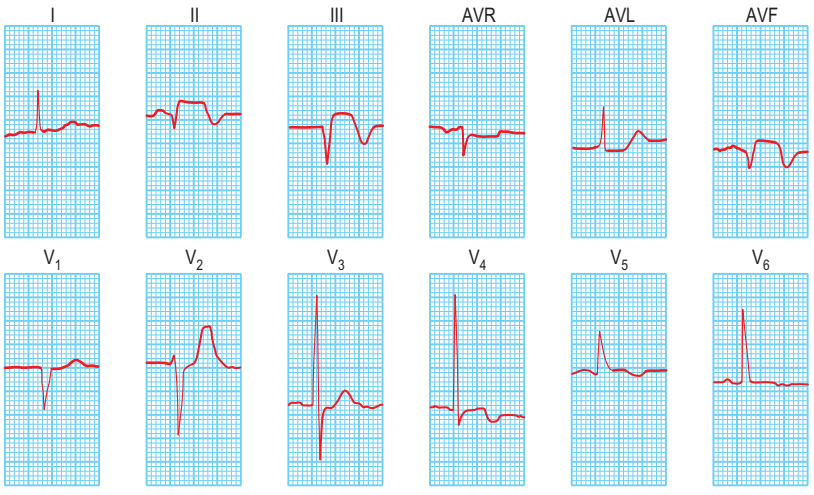


Fig. 13.74 **An acute inferior wall myocardial infarction** shown by a 12-lead ECG. Note the raised ST segment and Q waves in the inferior leads (II, III and AVF). The additional T wave inversion in V_4 and V_5 probably represents anterior wall ischaemia.

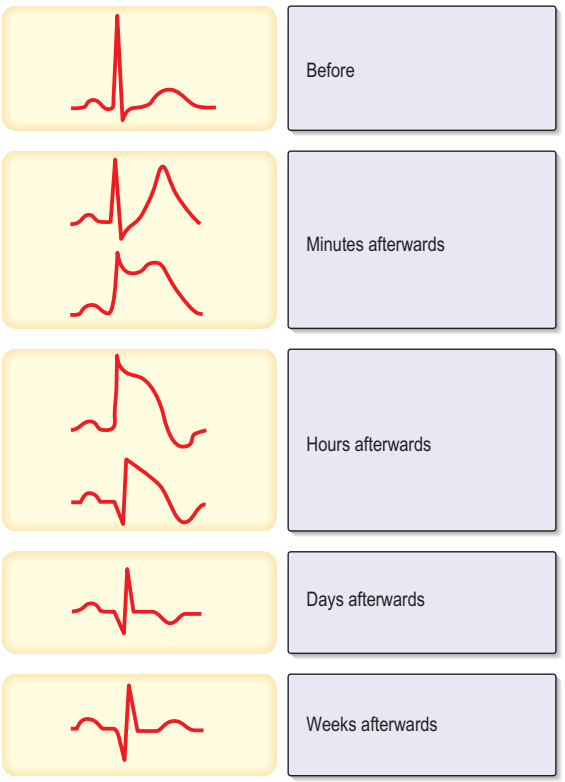


Fig. 13.75 **Electrocardiographic evolution of myocardial infarction (STEMI).** After the first few minutes the T waves become tall, pointed and upright and there is ST segment elevation. After the first few hours the T waves invert, the R wave voltage is decreased and Q waves develop. After a few days the ST segment returns to normal. After weeks or months the T wave may return to upright but the Q wave remains.

Investigations

Blood samples should be taken for cardiac troponin I or T levels or CK-MB level according to local hospital protocol, although treatment should not be deferred until the results are available. Full blood count, serum electrolytes, glucose and lipid profile should be obtained. Transthoracic echocardiography (TTE) may be helpful to confirm a myocardial infarction, as wall-motion abnormalities are detectable early in STEMI. TTE may detect alternative diagnoses such as aortic dissection, pericarditis or pulmonary embolism.

Early medical management

Accident and emergency

Rapid triage for chest pain (N.B. time is muscle):

- aspirin 150–300 mg chewed and clopidogrel 300 mg oral gel
- sublingual glyceryl trinitrate 0.3–1 mg. Repeat
- oxygen – nasal cannula 2–4 L/min (Fig. 15.21)
- brief history/risk factors. Examination
- intravenous access + blood for markers (plus FBC, biochemistry, lipids, glucose)
- 12-lead ECG
- intravenous opiate, e.g. diamorphine (or morphine) 2.5–5 mg + antiemetic, e.g. metoclopramide 10 mg
- beta-blocker (if no contraindication) for ongoing chest pain, hypertension, tachycardia
- if primary PCI available (see p. 750) give GP IIb/IIIa inhibitor. Alternatively give thrombolysis (see below).

Pre-hospital treatment, including thrombolysis, can be given by trained healthcare professionals under strict guidelines.

Fibrinolysis

Fibrinolytic agents enhance the breakdown of occlusive thromboses by the activation of plasminogen to form plasmin. The initial thrombolytic agent used in clinical trials was streptokinase. This agent is derived from bacteria, which can lead to the development of neutralizing antibodies that limit its repeated use. In the ISIS 2 study of acute myocardial infarction, patients were randomized to receive a 1-hour intravenous infusion of 1.5 million units of streptokinase, or 1 month of 160 mg/day enteric-coated aspirin, both active treatments, or neither. Both streptokinase and aspirin were significantly better than placebo in reducing vascular mortality at 5 weeks (9.2% versus 12.0%, and 9.4% versus 11.8%, respectively). The combination of streptokinase and aspirin was significantly better than either agent alone.

The GISSI 2 and ISIS 3 studies showed no significant difference in the effectiveness of streptokinase and tissue plasminogen activator (t-PA) or anistreplase. In addition there was no mortality benefit from subcutaneous unfractionated heparin versus no heparin.

In the GUSTO trial, accelerated t-PA over 90 minutes with intravenous heparin reduced death by 10 in every 1000 patients versus streptokinase, although there were three more strokes in the t-PA group versus streptokinase and subcutaneous heparin.

In the meta-analysis of fibrinolytics (FTT), fibrinolysis within 6 hours of STEMI or LBBB MI, prevented 30 deaths in every 1000 patients treated. Between 7 and 12 hours 20 in every 1000 deaths were prevented. After 12 hours the benefits are limited, and there is evidence to suggest less benefit for older patients, possibly because of the increased risk of strokes.

Prompt reperfusion therapy (door to needle time <30 minutes) will reduce the death rate following myocardial infarction. Double bolus r-PA (reteplase) and single bolus TNK-t-PA (tenecteplase) facilitate rapid administration of fibrinolytic therapy and can be used for pre-hospital thrombolysis. In patients who fail to reperfuse by 60–90 minutes as demonstrated by 50% resolution of the ST segment elevation, rethrombolysis or referral for rescue coronary angioplasty is recommended.

Aspirin therapy should be prescribed with fibrinolysis, but there is little additional benefit in combining clopidogrel or abciximab therapy in patients with STEMI/new LBBB. Heparin is recommended with t-PA or tenecteplase, but not with streptokinase. Enoxaparin (low molecular weight heparin) appears to be superior to unfractionated i.v. heparin in patients receiving TNK-t-PA, with less reocclusion and better late patency (ASSENT 3).

The contraindications to thrombolysis are provided in Table 13.32.

Percutaneous coronary intervention (PCI)

PCI performed within 90 minutes is the preferred reperfusion therapy in interventional cardiology centres that have the expertise available. In the PAMI (Primary Angioplasty in Myocardial Infarction) trial, patients with a myocardial infarction who presented within 12 hours of the onset of STEMI were

Table 13.32 Contraindications to thrombolysis

Absolute contraindications

Haemorrhagic stroke or stroke of unknown origin at any time

Ischaemic stroke in preceding 6 months

Central nervous system damage or neoplasms

Recent major trauma/surgery/head injury (within preceding 3 weeks)

Gastrointestinal bleeding within the last month

Known bleeding disorder

Aortic dissection

Relative contraindications

Transient ischaemic attack in preceding 6 months

Oral anticoagulant therapy

Pregnancy or within 1 week postpartum

Non-compressible punctures

Traumatic resuscitation

Refractory hypertension (systolic blood pressure > 180 mmHg)

Advanced liver disease

Infective endocarditis

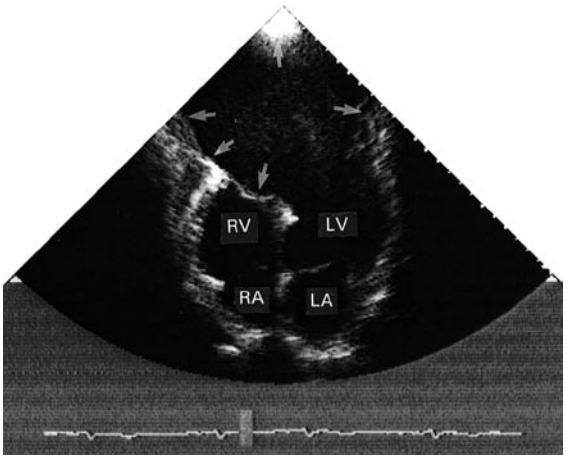


Fig. 13.76 **2-D Echocardiogram (apical four-chamber view) showing a very large apical left ventricular aneurysm** (arrows).The relatively static blood in the aneurysm produces a swirling 'smoke' effect. This aneurysm was successfully resected surgically. LV, left ventricle; LA, left atrium; RV, right ventricle; RA, right atrium.

randomized to primary PTCA or t-PA followed by conservative care. At 2-year follow-up the primary PTCA group had less recurrent ischaemia, lower re-intervention rates and reduced hospital readmission rates. Primary PTCA produced a combined end-point of death or re-infarction of 14.9% compared to 23% for t-PA. PCI with thrombus aspiration has recently been shown to result in better reperfusion and clinical outcomes.

The DANAMI 2 study investigated if rapid transfer of patients with STEMI for primary angioplasty in an interventional centre was superior to thrombolysis. Patients within 12 hours of a high-risk STEMI (>4 mm elevation) received front-loaded t-PA, primary angioplasty at a local centre, or primary angioplasty at an interventional centre after transfer. Primary PCI significantly reduced the death rate in both local and transferred patients as compared to thrombolytic therapy (8.0% versus 13.7%). The majority of the benefits with primary PCI are obtained by a reduction in recurrent myocardial infarction.

Coronary stenting in primary PCI reduces the need for repeat target vessel revascularization but did not appear to reduce mortality rates. However, one recent study using drug-eluting stents showed a decreased 2-year mortality rate.

The use of abciximab in STEMI patients undergoing primary angioplasty reduces immediate outcome (death, myocardial infarction, urgent revascularization), but this benefit is minimal by 6 months.

PCI following thrombolysis was discouraged but a recent trial suggests it is safe and improves the 1-year clinical outcome.

Recent randomized trials have compared a strategy of thrombolysis (in hospitals without PCI capability) versus transfer to a PCI centre and have demonstrated a significant reduction in the combined end-point of death, reinfarction and stroke (although with a non-significant reduction in mortality) with transfer. The strategy of transfer for primary PCI is appropriate if the intervention can be performed within 90 minutes of presentation and this is now optimal therapy.

Coronary artery bypass surgery

Cardiac surgery is usually reserved for the complications of myocardial infarction, such as ventricular septal defect or mitral regurgitation.

Complications of myocardial infarction

Heart failure

Cardiac failure post STEMI is a poor prognostic feature that necessitates medical and invasive therapy to reduce the death rate. The Killip classification is used to assess patients with heart failure post MI:

- Killip I – no crackles and no third heart sound
- Killip II – crackles in <50% of the lung fields or a third heart sound
- Killip III – crackles in >50% of the lung fields
- Killip IV – cardiogenic shock.

Mild heart failure may respond to intravenous furosemide 40–80 mg i.v., with GTN administration if the blood pressure is satisfactory. Oxygen is required, with regular oxygen monitoring. ACE inhibitors can be given in <24–48 hours if the blood pressure is satisfactory. Patients with severe heart failure may require Swan–Ganz catheterization to determine the pulmonary wedge pressure. Intravenous inotropes such as dopamine or dobutamine are used in patients with severe heart failure. If the patient is in cardiogenic shock, then revascularization ± intra-aortic balloon pump insertion may be required.

Myocardial rupture and aneurysmal dilatation

Rupture of the free wall of the left ventricle is usually an early, catastrophic and fatal event. The patient will have a haemodynamic collapse, then an electromechanical cardiac arrest. A subacute rupture may allow for pericardiocentesis followed by the surgical repair of the rupture. Aneurysmal dilatation of the infarcted myocardium (Figs 13.17 and 13.76) is a late complication that may require surgical repair.

Ventricular septal defect (VSD)

A VSD may occur in 1–2.0% of patients with STEMI, and may be associated with delayed or failed fibrinolysis. However, mortality is very high with a 12-month unoperative mortality of 92%. An intra-aortic balloon pump (IABP) and coronary angiography may allow for patient optimization prior to surgery.

Mitral regurgitation

Severe mitral regurgitation can occur early in the course of STEMI. Three mechanisms may be responsible for the mitral regurgitation, and a transoesophageal echocardiogram (TOE) may be necessary to confirm the aetiology:

- severe left ventricular dysfunction and dilatation, causing annular dilatation of the valve and subsequent regurgitation
- myocardial infarction of the inferior wall, producing dysfunction of the papillary muscle that may respond to coronary intervention
- myocardial infarction of the papillary muscles, producing sudden severe pulmonary oedema and cardiogenic shock (IABP, coronary angiography and early surgery may improve patient survival).

Cardiac arrhythmias

Ventricular tachycardia and ventricular fibrillation are common in STEMI, particularly with reperfusion. Cardiac arrest requires defibrillation. Ventricular tachycardia should be treated with intravenous beta-blockers (metoprolol 5 mg, esmolol 50–200 µg/kg/min), lidocaine 50–100 mg, or amiodarone 900–1200 mg per 24 hours. If the patient is hypotensive, synchronized cardioversion may be performed. Ensure that the serum potassium is above 4.5 mmol/L. Refractory ventricular tachycardia or fibrillation may respond to magnesium 8 mmol/L over 15 minutes i.v.

Atrial fibrillation occurs frequently and treatment with beta-blockers and digoxin may be required. Cardioversion is possible but relapse is frequent.

Bradyarrhythmias can be treated initially with i.v. atropine 0.5 mg repeated up to six times in 4 hours. Temporary transcutaneous or transvenous pacemaker insertion may be required in patients with symptomatic heart block.

Conduction disturbances

These are common following MI. AV block may occur during acute MI, especially of the inferior wall (the right coronary artery usually supplies the SA and AV nodes). Heart block, when associated with haemodynamic compromise, may need treatment with atropine or a temporary pacemaker. Such blocks may last for only a few minutes, but frequently continue for several days. Permanent pacing may need to be considered if heart block persists for over 2 weeks.

Post-MI pericarditis and Dressler's syndrome

See section below entitled 'Pericardial disease' (p. 794).

Post-MI lifestyle modification

After recovery from an acute myocardial infarction patients should be encouraged to participate in a cardiac rehabilitation programme that provides education and information appropriate to the patients' requirements. An exercise programme forms part of the rehabilitation.

- Patients should be encouraged to eat a Mediterranean-style diet and to consume >7 g of omega 3 fatty acids/week from oily fish or >1 g daily of omega-3-acid ethyl esters
- Patients should maintain alcohol consumption within safe limits (≤21 units/week for men or ≤14 units for women) and to avoid binge drinking.
- Patients should be physically active for 20–30 minutes/day.
- Patients should stop smoking.
- Overweight and obese patients should be offered advice and support to achieve and maintain a healthy weight.
- Patients with hypertension should be treated to <140/90 or <130/80 if chronic kidney disease or diabetes.
- Patients with diabetes should be treated to maintain HbA1c <7%.

Post-MI drug therapy and assessment

Extensive clinical trial evidence has been gathered in post-myocardial infarction patients, demonstrating that a range of pharmaceuticals are advantageous in reducing mortality over the following years. Therefore, post MI most patients should be taking most of the following medications:

- Aspirin 75–100 mg/day.
- A beta-blocker to maintain heart rate <60 b.p.m., e.g. metoprolol 50 mg twice daily.
- ACE inhibitors, e.g. ramipril 2.5 mg twice daily, titrated to maximum tolerated or target dose (if intolerant of ACE inhibitors use ARB, e.g. valsartan 20 mg twice daily titrated to maximum tolerated or target dose)
- Statins, e.g. simvastatin 20–80 mg/day.
- Clopidogrel 75 mg/day for 9–12 months should be added in moderate–high risk patients with non-ST elevation acute coronary syndrome (NST-ACS).
- Aldosterone antagonist, e.g. eplerenone 25 mg/day, should be given in patients post MI with clinical evidence of heart failure and reduced ejection fraction (renal function and potassium levels should be monitored).

If primary angioplasty has not been performed, then it is necessary to identify residual ischaemia and viability, to determine the need for coronary angiography. Uncomplicated patients with no angina during the hospital stay should have a low-level exercise test prior to discharge followed by a formal ETT 6 weeks later. A positive test usually suggests diagnostic/therapeutic coronary angiography/stenting. Alternatively, nuclear scintigraphy or dobutamine stress echocardiography can be used at 5 days to determine the amount of viable myocardium and the extent of myocardial ischaemia. Echocardiography should be performed to guide therapy and to determine baseline ejection fraction. In a patient with an ejection fraction < 30%, ventricular tachycardia on a Holter monitor may warrant an ICD.

SUDDEN CARDIAC DEATH

(see also Table 13.7)

Sudden cardiac death is generally defined as death due to cardiac causes which occurs within 6 hours of the onset of

FURTHER READING

Anderson JL, Adams CD, Antman EM, et al. ACC/AHA 2007 Guidelines for the management of patients with unstable angina/non-ST-elevation myocardial infarction: Executive summary. *Circulation* 2007; **116**: 803–877.

Bassand JP, Hamm CW, Ardissino D, et al, for The Task Force for the Diagnosis and Treatment of Non-ST-Segment Elevation Acute Coronary Syndromes of the European Society of Cardiology. Guidelines for the diagnosis and treatment of non-ST-segment elevation acute coronary syndromes. *European Heart Journal* 2007; **28**: 1598–1660.

Hansson SK. Inflammation, atherosclerosis and coronary artery disease. *New England Journal of Medicine* 2005; **352**: 1685–1695.

Xavier D, Pais P, Devereauz PJ, et al. Treatment and outcome of acute coronary syndromes in India (CREATE). *Lancet* 2008; **371**: 1435–1442.

Table 13.33 Causes of sudden cardiac death

Coronary artery disease
Acute myocardial infarction – STEMI
Chronic ischaemic heart disease
Following coronary artery bypass surgery
After successful resuscitation for cardiac arrest
Congenital anomaly of coronary arteries
Coronary artery embolism
Coronary arteritis

Non-coronary artery disease
Hypertrophic cardiomyopathy
Dilated cardiomyopathy (ischaemic or idiopathic)
Arrhythmogenic right ventricular cardiomyopathy
Congenital long QT syndrome
Brugada's syndrome
Valvular heart disease (aortic stenosis, mitral valve prolapse) ± infective endocarditis
Cyanotic heart disease (tetralogy of Fallot, transposition)
Acyanotic heart disease (ventricular septal defect, patent ductus arteriosus)

Table 13.34 Rheumatic valvular lesions

Valve involved	Percentage of cases
Mitral valve alone	50
Mitral and aortic valves	40
Mitral, aortic and tricuspid valves	5
Aortic valve alone	2
All other combinations	3

Table 13.35 Complications of mitral stenosis

Atrial fibrillation
Systemic embolization
Pulmonary hypertension
Pulmonary infarction
Chest infections
Infective endocarditis (rare)
Tricuspid regurgitation
Right ventricular failure

symptoms. The causes of sudden cardiac death tend to mirror, in risk factors and prevalence, the predominant cardiac causes of death in a given population. In developed countries, the majority (80%) are estimated to be due to coronary artery disease, a further 10–15% to cardiomyopathies and 5% to valvular heart disease. Post-mortem studies have revealed that coronary atheroma is present in 80–90% of cases. Forty per cent of all deaths due to coronary atherosclerosis occur suddenly in this way, although up to 50% of patients dying suddenly because of coronary atheroma have no preceding history of coronary disease. The majority of cases of sudden death due to atheroma appear to be due to fatal ventricular arrhythmia, sometimes triggered by acute myocardial ischaemia. Other less common causes of sudden cardiac death are listed in Table 13.33. Patients presenting with a cardiac arrest require immediate cardiopulmonary resuscitation as outlined in Figure 13.36. In those patients who survive a cardiac arrest, an implantable cardioverter–defibrillator is often required, to prevent further cardiac arrest. Antiarrhythmic drugs such as amiodarone may be used as an alternative to an implantable cardioverter–defibrillator but are less effective.

VALVULAR HEART DISEASE

MITRAL STENOSIS

Almost all mitral stenosis is due to rheumatic heart disease.

At least 50% of sufferers have a history of rheumatic fever or chorea. The single most common valve lesion due to rheumatic fever is pure mitral stenosis (50%) (Table 13.34). The mitral valve is affected in over 90% of those with rheumatic valvular heart disease. Rheumatic mitral stenosis is much more common in women. The pathological process results after some years in valve thickening, cusp fusion, calcium deposition, a narrowed (stenotic) valve orifice and progressive immobility of the valve cusps.

Other causes include:

- Lutembacher's syndrome, which is the combination of acquired mitral stenosis and an atrial septal defect
- a rare form of congenital mitral stenosis
- in the elderly, a syndrome similar to mitral stenosis, which develops because of calcification and fibrosis of the valve, valve ring and subvalvular apparatus (chordae tendineae)
- carcinoid tumours metastasizing to the lung, or primary bronchial carcinoid.

Pathophysiology

When the normal valve orifice area of 5 cm^2 is reduced to approximately 1 cm^2, severe mitral stenosis is present. In order that sufficient cardiac output will be maintained, the left atrial pressure increases and left atrial hypertrophy and dilatation occur. Consequently, pulmonary venous, pulmonary arterial and right heart pressures also increase. The increase in pulmonary capillary pressure is followed by the development of pulmonary oedema. This is partially prevented by alveolar and capillary thickening and pulmonary arterial vasoconstriction (reactive pulmonary hypertension). Pulmonary hypertension leads to right ventricular hypertrophy, dilatation and failure. Right ventricular dilatation results in tricuspid regurgitation. Mitral stenosis is frequently associated with complications (Table 13.35).

Symptoms

Usually there are no symptoms until the valve orifice is moderately stenosed (i.e. has an area of 2 cm^2). In Europe this does not usually occur until several decades after the first attack of rheumatic fever, but children of 10–20 years of age in the Middle or Far East may have severe calcific mitral stenosis.

Because of pulmonary venous hypertension and recurrent bronchitis, progressively *severe dyspnoea* develops. A cough productive of blood-tinged, frothy sputum is quite common, and occasionally frank haemoptysis may occur. The development of pulmonary hypertension eventually leads to *right heart failure* and its symptoms of weakness, fatigue and abdominal or lower limb swelling.

The large left atrium favours *atrial fibrillation*, giving rise to symptoms such as palpitations. Atrial fibrillation may result in *systemic emboli*, most commonly to the cerebral vessels resulting in neurological sequelae, but mesenteric, renal and peripheral emboli are also seen. Clinical pulmonary embolism as a result of mitral stenosis associated with atrial fibrillation is less commonly seen, but it is likely that subclinical pulmonary emboli occur.

Signs (see Clinical memo in Fig. 13.77)

Face

Severe mitral stenosis with pulmonary hypertension is associated with the so-called mitral facies or malar flush. This is a bilateral, cyanotic or dusky pink discoloration over the upper cheeks that is due to arteriovenous anastomoses and vascular stasis.

Pulse

Mitral stenosis may be associated with a small-volume pulse which is usually regular early on in the disease process when most patients are in sinus rhythm. However, as the severity of the disease progresses, many patients develop atrial fibrillation resulting in an irregularly irregular pulse. The development of atrial fibrillation in these patients often causes a dramatic clinical deterioration.

Jugular veins

If right heart failure develops, there is obvious distension of the jugular veins. If pulmonary hypertension or tricuspid stenosis is present, the *a* wave will be prominent provided that atrial fibrillation has not supervened.

Palpation

There is a tapping impulse felt parasternally on the left side. This is the result of a palpable first heart sound combined with left ventricular backward displacement produced by an enlarging right ventricle. A sustained parasternal impulse due to right ventricular hypertrophy may also be felt.

Auscultation

Auscultation (Fig. 13.77) reveals a loud first heart sound if the mitral valve is pliable, but it will not occur in calcific mitral stenosis. As the valve suddenly opens with the force of the increased left atrial pressure, an 'opening snap' will be heard. This is followed by a low-pitched 'rumbling' mid-diastolic murmur best heard with the bell of the stethoscope held lightly at the apex with the patient lying on the left side. If the patient is in sinus rhythm, the murmur becomes louder at the end of diastole as a result of atrial contraction (pre-systolic accentuation).

The severity of mitral stenosis is judged clinically on the basis of several criteria:

- The presence of pulmonary hypertension implies that mitral stenosis is severe. Pulmonary hypertension is recognized by a right ventricular heave, a loud pulmonary component of the second heart sound, eventually with signs of right-sided heart failure, such as oedema and hepatomegaly. Pulmonary hypertension results in pulmonary valvular regurgitation that causes an early diastolic murmur in the pulmonary area known as a Graham Steell murmur.
- The closeness of the opening snap to the second heart sound is proportional to the severity of mitral stenosis.
- The length of the mid-diastolic murmur is proportional to the severity.
- As the valve cusps become immobile, the loud first heart sound softens and the opening snap disappears. When pulmonary hypertension occurs, the pulmonary component of the second sound is increased in

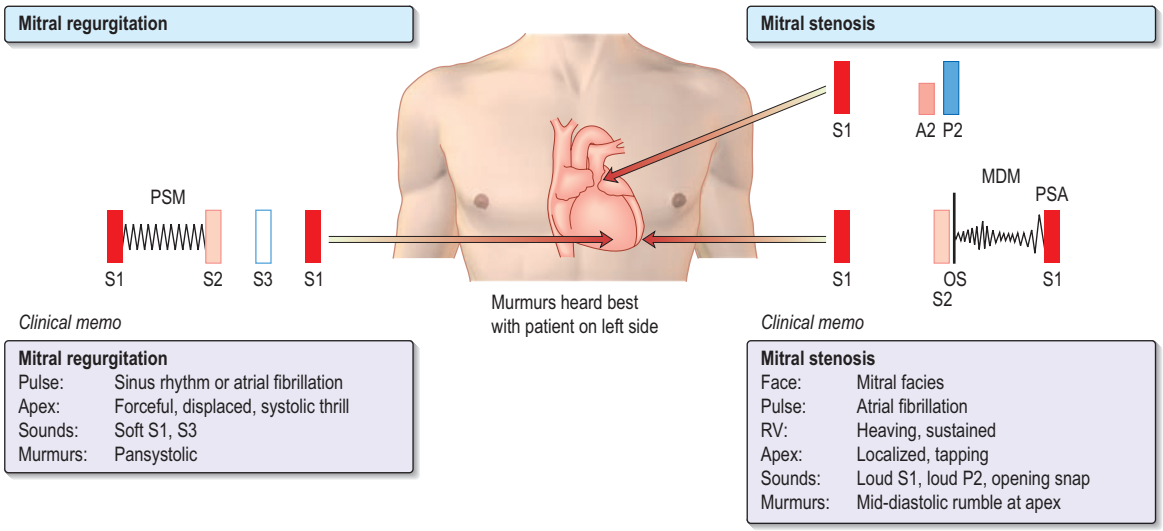


Fig. 13.77 Features associated with mitral regurgitation and mitral stenosis. A2, aortic component of the second heart sound; MDM, mid-diastolic murmur; OS, opening snap; P2, pulmonary component of the second heart sound; PSA, presystolic accentuation; PSM, pansystolic murmur; S1, first heart sound; S2, second heart sound; S3, third heart sound.

intensity and the mitral diastolic murmur may become quieter because of the reduction of cardiac output.

Investigations

Chest X-ray

The chest X-ray usually shows a generally small heart with an enlarged left atrium (Fig. 13.14, p. 693). Pulmonary venous hypertension is usually also present. Late in the course of the disease a calcified mitral valve may be seen on a penetrated or lateral view. The signs of pulmonary oedema or pulmonary hypertension may also be apparent when the disease is severe.

Electrocardiogram

In sinus rhythm the ECG shows a bifid P wave owing to delayed left atrial activation (Fig. 13.78). However, atrial fibrillation is frequently present. As the disease progresses, the ECG features of right ventricular hypertrophy (right axis deviation and perhaps tall R waves in lead V1) may develop (Fig. 13.79).

Imaging

Echocardiogram (Fig. 13.28)

Two-dimensional echocardiography allows assessment of the mitral valve apparatus and calculation of mitral valve area, thus providing a useful guide in determining whether balloon valvotomy or valve replacement is the treatment of choice in patients symptomatic on medical therapy. Two-dimensional echocardiography also determines left atrial and right ventricular size and function. Continuous wave (CW) Doppler may also be used to measure mitral valve area and provides an estimate of pulmonary artery pressure through measurement of the degree of tricuspid regurgitation. In most cases, echocardiography alone is sufficient to judge the severity of mitral stenosis such that decisions regarding surgery can be made.

Transoesophageal echocardiography *(TOE)* is performed to detect the presence of left atrial thrombus (p. 702).

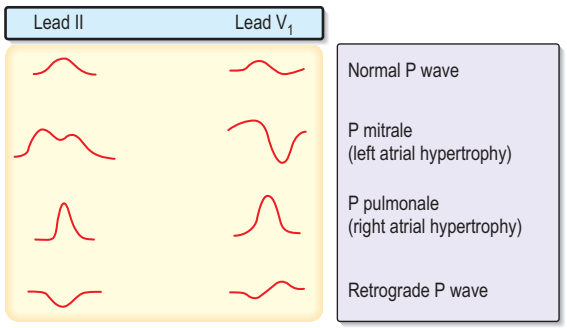


Fig. 13.78 **A bifid P wave as seen on the ECG in mitral stenosis (P mitrale).** Also shown for comparison are other P wave abnormalities.

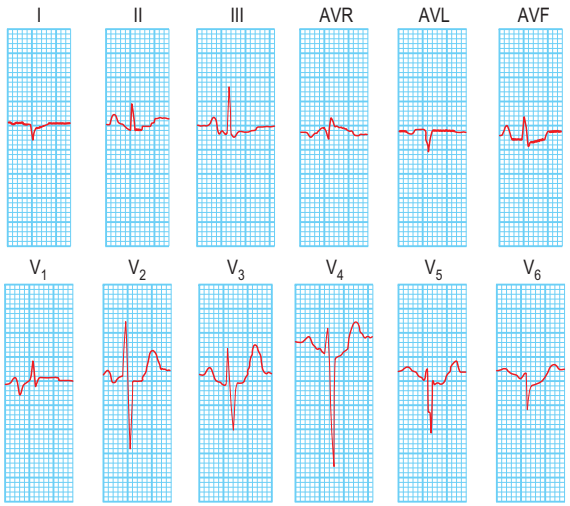


Fig. 13.79 **Severe mitral stenosis** shown by a 12-lead ECG. Note the right axis deviation (frontal plane axis = + 120°), the left atrial conduction abnormality (large terminal negative component of the P wave in V_1) and the right ventricular hypertrophy (R wave in V_1 and right axis deviation).

Cardiac magnetic resonance (CMR; see p. 707)

This accurately shows mitral valve anatomy.

Cardiac catheterization is now seldom required and is only used if coexisting cardiac problems (e.g. mitral regurgitation or coronary artery disease) are suspected. The typical findings in mitral stenosis are a diastolic pressure that is higher in the left atrium than in the left ventricle. This gradient of pressure is usually proportional to the degree of the stenosis.

Treatment

Mild mitral stenosis may need no treatment other than prompt therapy of attacks of bronchitis. Infective endocarditis in pure mitral stenosis is uncommon. Early symptoms of mitral stenosis such as mild dyspnoea can usually be treated with low doses of diuretics. The onset of atrial fibrillation requires treatment with digoxin and anticoagulation to prevent atrial thrombus and systemic embolization. If pulmonary hypertension develops or the symptoms of pulmonary congestion persist despite therapy, surgical relief of the mitral stenosis is advised. There are four operative measures.

Trans-septal balloon valvotomy

A catheter is introduced into the right atrium via the femoral vein. The interatrial septum is then punctured and the catheter advanced into the left atrium and across the mitral valve. A balloon is passed over the catheter to lie across the valve, and then inflated briefly to split the valve commissures. The procedure is performed under local anaesthesia in the cardiac catheter laboratory. As with other valvotomy techniques, significant regurgitation may result, necessitating valve replacement (see below). This procedure is ideal for patients with pliable valves in whom there is little involvement of the subvalvular apparatus and in whom there is minimal mitral regurgitation. The procedure cannot be performed when there is heavy calcification or more than mild mitral regurgitation. The presence of thrombus in the left atrium is a contraindication to balloon valvotomy; therefore, transoesophageal echocardiography must be performed prior to this technique in order that left atrial thrombus can be excluded.

Closed valvotomy

This operation is advised for patients with mobile, non-calcified and non-regurgitant mitral valves. The fused cusps are forced apart by a dilator introduced through the apex of

the left ventricle and guided into position by the surgeon's finger inserted via the left atrial appendage. Cardiopulmonary bypass is not needed for this operation. Closed valvotomy may produce a good result for 10 years or more. The valve cusps often re-fuse and eventually another operation may be necessary.

Open valvotomy

This operation is often preferred to closed valvotomy. The cusps are carefully dissected apart under direct vision. Cardiopulmonary bypass is required. Open dissection reduces the likelihood of causing traumatic mitral regurgitation.

Mitral valve replacement

Replacement of the mitral valve is necessary if:

- mitral regurgitation is also present
- there is a badly diseased or badly calcified stenotic valve that cannot be reopened without producing significant regurgitation
- there is moderate or severe mitral stenosis and thrombus in the left atrium despite anticoagulation.

Artificial valves (see p. 769) may work successfully for more than 20 years. Anticoagulants are generally necessary to prevent the formation of thrombus, which might obstruct the valve or embolize.

MITRAL REGURGITATION

There are many causes of mitral valve regurgitation but rheumatic heart disease (50%) and a prolapsing mitral valve are the most common. Other causes include:

- aortic valve disease
- acute rheumatic fever
- myocarditis
- dilated cardiomyopathy
- hypertensive heart disease
- ischaemic heart disease
- infective endocarditis – mitral regurgitation may result from destruction of the mitral valve leaflets
- hypertrophic cardiomyopathy – left ventricular contraction is disorganized
- rheumatic autoimmune disease – systemic lupus erythematosus (SLE)
- collagen abnormalities – Marfan's syndrome and Ehlers–Danlos syndrome
- degeneration of the valve cusps or mitral annular calcification
- rupture of the chordae tendineae (due to myocardial infarction, infective endocarditis or trauma) – results in acute and very severe mitral regurgitation
- drugs, e.g. centrally acting appetite suppressants, e.g. fenfluramine; dopamine agonists, e.g. cabergoline.

Pathophysiology

Regurgitation into the left atrium produces left atrial dilatation but little increase in left atrial pressure if the regurgitation is long-standing, as the regurgitant flow is accommodated by the large left atrium. With acute mitral regurgitation the normal compliance of the left atrium does not allow much dilatation and the left atrial pressure rises. Thus, in acute mitral regurgitation the left atrial '*v*' wave is greatly increased and pulmonary venous pressure rises to produce pulmonary oedema. Since a proportion of the stroke volume is regurgitated, the stroke volume increases to maintain the forward cardiac output and the left ventricle therefore enlarges.

Symptoms

Mitral regurgitation can be present for many years and the cardiac dimensions greatly increased before any symptoms occur. The increased stroke volume is sensed as a 'palpitation'. *Dyspnoea and orthopnoea* develop owing to pulmonary venous hypertension occurring as a direct result of the mitral regurgitation and secondarily to left ventricular failure. *Fatigue and lethargy* develop because of the reduced cardiac output. In the late stages of the disease the *symptoms of right heart failure* also occur and eventually lead to congestive cardiac failure. *Cardiac cachexia* may develop. Thromboembolism is less common than in mitral stenosis, but *subacute infective endocarditis* is much more common.

Signs (see Clinical memo in Fig. 13.77)

The physical signs of uncomplicated mitral regurgitation are:

- laterally displaced (forceful) diffuse apex beat and a systolic thrill (if severe)
- soft first heart sound, owing to the incomplete apposition of the valve cusps and their partial closure by the time ventricular systole begins
- pansystolic murmur, owing to the occurrence of regurgitation throughout the whole of systole, being loudest at the apex but radiating widely over the precordium and into the axilla
- prominent third heart sound, owing to the sudden rush of blood back into the dilated left ventricle in early diastole (sometimes a short mid-diastolic flow murmur may follow the third heart sound).

The signs related to atrial fibrillation, pulmonary hypertension, and left and right heart failure develop later in the disease. The onset of atrial fibrillation has a much less dramatic effect on symptoms than in mitral stenosis.

Investigations

Chest X-ray

The chest X-ray may show left atrial and left ventricular enlargement. There is an increase in the CTR, and valve calcification is seen.

Electrocardiogram

The ECG shows the features of left atrial delay (bifid P waves) and left ventricular hypertrophy (Fig. 13.80) as manifested by tall R waves in the left lateral leads (e.g. leads I and V_6) and deep S waves in the right-sided precordial leads, (e.g. leads V_1 and V_2). (Note that SV_1 plus RV_5 or $RV_6 > 35$ mm indicates left ventricular hypertrophy.) Left ventricular hypertrophy occurs in about 50% of patients with mitral regurgitation. Atrial fibrillation may be present.

Echocardiogram

The echocardiogram shows a dilated left atrium and left ventricle. There may be specific features of chordal or papillary muscle rupture. CW Doppler can determine the velocity of the regurgitant jet.

The echocardiogram is not as definitive in mitral regurgitation as in mitral stenosis. However, useful information

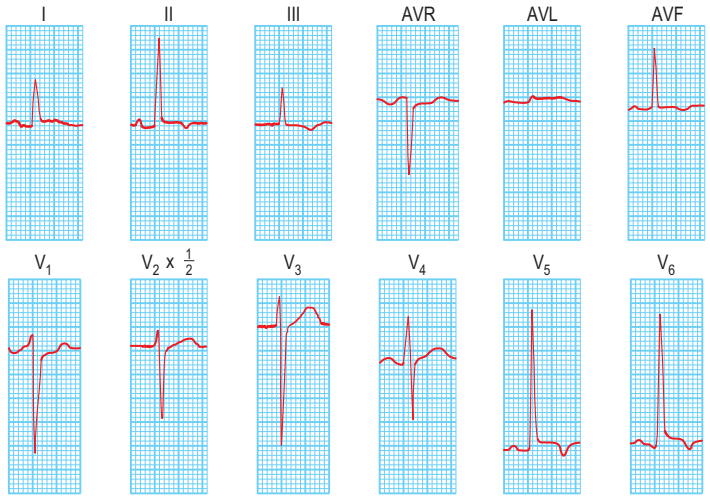


Fig. 13.80 **Left ventricular hypertrophy** shown in a 12-lead ECG. Note the size of the S wave seen in V_1 (21 mm); S in V_1 + R in V_6 = >35 mm.

regarding the severity of the condition can be obtained indirectly by observing the dynamics of ventricular function. Transoesophageal echocardiography (TOE) helps to identify structural valve abnormalities before surgery. Intraoperative TOE can also aid assessment of the efficacy of valve repair.

Cardiac catheterization

This demonstrates a prominent left atrial systolic pressure wave, and when contrast is injected into the left ventricle it is seen regurgitating into an enlarged left atrium during systole. CMR has largely replaced catheterization.

Treatment

Mild mitral regurgitation in the absence of symptoms can be managed conservatively by following the patient with serial echocardiograms. Prophylaxis against endocarditis is discussed in Chapter 4 (see p. 93). Any evidence of progressive cardiac enlargement generally warrants early surgical intervention by either mitral valve repair or replacement. The advantages of surgical intervention are diminished in more advanced disease. In patients who are not considered appropriate for surgical intervention, or in whom surgery will be considered at a later date, management usually involves treatment with ACE inhibitors, diuretics and possibly anticoagulants. Sudden torrential mitral regurgitation, as seen with chordal or papillary muscle rupture or infective endocarditis, necessitates emergency mitral valve replacement.

Prolapsing (billowing) mitral valve

This is also known as Barlow's syndrome or floppy mitral valve. It is due to excessively large mitral valve leaflets, an enlarged mitral annulus, abnormally long chordae or disordered papillary muscle contraction. Histology may demonstrate myxomatous degeneration of the mitral valve leaflets. It is more commonly seen in young women than in men or older women and it has a familial incidence. Its cause is unknown but it is associated with Marfan's syndrome, thyrotoxicosis, rheumatic or ischaemic heart disease. It also occurs in association with atrial septal defect and as part of hypertrophic cardiomyopathy. Mild mitral valve prolapse is so common that it should be regarded as a normal variant.

Pathophysiology

During ventricular systole, a mitral valve leaflet (most commonly the posterior leaflet) prolapses into the left atrium. This may result in abnormal ventricular contraction, papillary muscle strain and some mitral regurgitation. Usually the syndrome is not haemodynamically serious. Thromboembolism occurs.

Symptoms

Atypical chest pain is the most common symptom. Usually the pain is left submammary and stabbing in quality. Sometimes it is substernal, aching and severe. Rarely, it is similar to typical angina pectoris. *Palpitations* may be experienced because of the abnormal ventricular contraction or because of the atrial and ventricular arrhythmias that are commonly associated with mitral valve prolapse. Sudden cardiac death due to fatal ventricular arrhythmias is a very rare but recognized complication.

Signs

The most common sign is a mid-systolic click, which is produced by the sudden prolapse of the valve and the tensing of the chordae tendineae that occurs during systole. This may be followed by a late systolic murmur owing to some regurgitation. With more regurgitation, the murmur becomes pansystolic mitral regurgitation.

Investigations

The diagnosis is confirmed on two-dimensional echocardiography, which typically shows posterior movement of one or both mitral valve cusps into the left atrium during systole.

Treatment

Usually, beta-blockade is effective for the treatment of the atypical chest pain and palpitations. Sometimes more specific antiarrhythmic drug treatment is necessary. When a prolapsing mitral valve is associated with significant mitral regurgitation and atrial fibrillation, anticoagulation is advised to prevent thromboembolism. Mitral valve prolapse associated with severe mitral regurgitation has a risk of sudden cardiac death. Mitral valve repair is indicated in all such cases. Very occasionally, mitral valve replacement rather

than repair may be necessary for severe regurgitation when there is severe prolapse of both leaflets. Prophylaxis against endocarditis is discussed in Chapter 4 (see p. 93).

AORTIC STENOSIS

Congenital aortic valve stenosis develops progressively because of turbulent blood flow through a congenitally abnormal (usually bicuspid) aortic valve. Most congenitally abnormal aortic valves occur in men and will calcify later in life.

Rheumatic fever results in progressive fusion, thickening and calcification of a previously normal three-cusped aortic valve. In rheumatic heart disease the aortic valve is affected in about 40% of cases and there is usually associated mitral valve disease.

Calcific valvular disease is the commonest cause of aortic stenosis and mainly occurs in the elderly. This is an inflammatory process involving macrophages and T lymphocytes with initially thickening of the subendothelium with adjacent fibrosis. The lesions contain lipoproteins which calcify, increasing leaflet stiffness and reducing systolic opening.

Valvular aortic stenosis should be distinguished from other causes of obstruction to left ventricular emptying (Fig. 13.81), which include:

- supravalvular obstruction – a congenital fibrous diaphragm above the aortic valve often associated with mental retardation and hypercalcaemia (William's syndrome)
- hypertrophic cardiomyopathy – septal muscle hypertrophy obstructing left ventricular outflow
- subvalvular aortic stenosis – a congenital condition in which a fibrous ridge or diaphragm is situated immediately below the aortic valve.

Pathophysiology

Obstructed left ventricular emptying leads to increased left ventricular pressure and compensatory left ventricular hypertrophy. In turn, this results in relative ischaemia of the left ventricular myocardium, and consequent angina, arrhythmias and left ventricular failure. The obstruction to left ventricular emptying is relatively more severe on exercise. Normally, exercise causes a many-fold increase in cardiac output, but when there is severe narrowing of the aortic valve orifice the cardiac output can hardly increase. Thus, the blood pressure falls, coronary ischaemia worsens, the myocardium fails and cardiac arrhythmias develop. Left ventricular systolic function is typically preserved in patients with aortic stenosis (cf. aortic regurgitation).

Symptoms

There are usually no symptoms until aortic stenosis is moderately severe (when the aortic orifice is reduced to one-third of its normal size). At this stage, *exercise-induced syncope*, *angina* and *dyspnoea* develop. When symptoms occur, the prognosis is poor – on average, death occurs within 2–3 years if there has been no surgical intervention.

Signs (see Clinical memo in Fig. 13.82)

Pulse

The carotid pulse is of small volume and is slow-rising or plateau in nature (see p. 690).

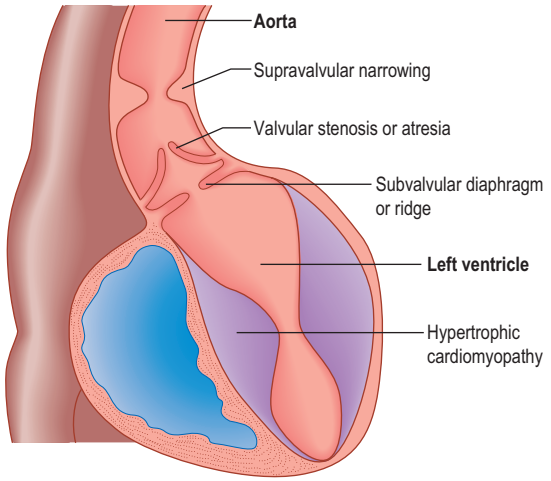
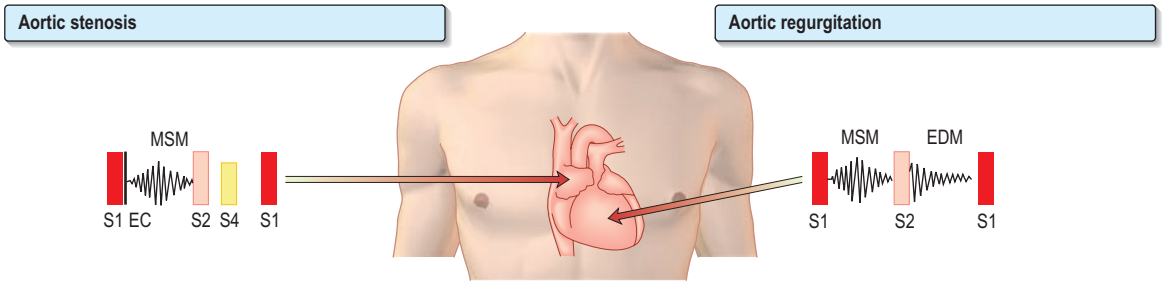


Fig. 13.81 Several forms of left ventricular outflow tract obstruction.



Clinical memo

Aortic stenosis

Pulse:	Sinus rhythm, low volume, slow rising
Aortic area:	Systolic thrill
Apex:	Not displaced, sustained
Sounds:	Ejection click, soft A2, S4
Murmurs:	Systolic, low pitched, ejection, radiating to carotids

Murmurs heard best with patient leaning forwards and breath held in expiration

Clinical memo

Aortic regurgitation

Pulse:	Sinus rhythm, large volume, collapsing
Blood pressure:	Wide pulse pressure
Apex:	Displaced, diffuse, forceful
Murmurs:	(1) High pitched, early diastolic at LSE (2) Ejection systolic at base and into neck (high flow) (3) Mid-diastolic rumble at apex (Austin-Flint) (not shown)

Fig. 13.82 Features of aortic stenosis and aortic regurgitation. EC, ejection click; EDM, early diastolic murmur; MSM, mid-systolic murmur; S1, first heart sound.

Precordial palpation

The apex beat is not usually displaced because hypertrophy (as opposed to dilatation) does not produce noticeable cardiomegaly. However, the pulsation is sustained and obvious. A double impulse is sometimes felt because the fourth heart sound or atrial contraction ('kick') may be palpable. A systolic thrill may be felt in the aortic area.

Auscultation

The most obvious auscultatory finding in aortic stenosis is an ejection systolic murmur that is usually 'diamond-shaped' (crescendo–decrescendo). The murmur is usually longer when the disease is more severe, as a longer ejection time is needed. The murmur is usually rough in quality and best heard in the aortic area. It radiates into the carotid arteries and also the precordium. The intensity of the murmur is not a good guide to the severity of the condition because it is lessened by a reduced cardiac output. In severe cases, the murmur may be inaudible.

Other findings include:

- systolic ejection click, unless the valve has become immobile and calcified
- soft or inaudible aortic second heart sound when the aortic valve becomes immobile
- reversed splitting of the second heart sound (splitting on expiration) (see p. 692)
- prominent fourth heart sound, which is caused by atrial contraction, is heard unless coexisting mitral stenosis prevents this.

Investigations

Chest X-ray

The chest X-ray usually reveals a relatively small heart with a prominent, dilated, ascending aorta. This occurs because turbulent blood flow above the stenosed aortic valve produces so-called 'post-stenotic dilatation'. The aortic valve may be calcified. The CTR increases when heart failure occurs.

Electrocardiogram

The ECG shows left ventricular hypertrophy and left atrial delay. A left ventricular 'strain' pattern due to 'pressure overload' (depressed ST segments and T wave inversion in leads orientated towards the left ventricle, i.e. leads I, AVL, V_5 and V_6) is common when the disease is severe. Usually, sinus rhythm is present, but ventricular arrhythmias may be recorded.

Echocardiogram

The echocardiogram readily demonstrates the thickened, calcified and immobile aortic valve cusps. Left ventricular hypertrophy may also be seen. The gradient across the valve can be estimated by CW Doppler, provided the left ventricular function is reasonable (Fig. 13.27, p. 702).

Cardiac catheterization

Cardiac catheterization can be used to document the systolic pressure difference (gradient) between the aorta and the left ventricle (Fig. 13.83) and assess left ventricular function. This is rarely necessary since all of this information can be gained non-invasively with echocardiography and CMR. Coronary angiography is necessary before recommending surgery.

Treatment

In patients with aortic stenosis, symptoms are a good index of severity and all symptomatic patients should have *aortic valve replacement*. Asymptomatic patients should be under regular review for assessment of symptoms and echocardiography. Antibiotic prophylaxis against infective endocarditis is discussed in Chapter 4 (see p. 93).

Provided that the valve is not severely deformed or heavily calcified, critical aortic stenosis in childhood or adolescence can be treated by valvotomy (performed under direct vision by the surgeon or by balloon dilatation using X-ray visualization). This produces temporary relief from the obstruction. Aortic valve replacement will usually be needed a few years later. Balloon dilatation (valvuloplasty) has been tried in adults, especially in the elderly, as an alternative to surgery. Generally results are poor and such treatment is reserved for patients unfit for surgery or as a 'bridge' to surgery (as systolic function will often improve).

Percutaneous valve replacement

A novel treatment for patients unsuitable for surgical aortic valve replacement is transcatheter implantation with a balloon expandable stent valve. In a recent study valve implantation was successful (86%) with a procedural mortality of 2% and 30-day mortality of 12%. Further larger and randomized studies with long-term follow-up are required.

AORTIC REGURGITATION

The most common causes of aortic regurgitation are rheumatic fever and infective endocarditis complicating a previ-

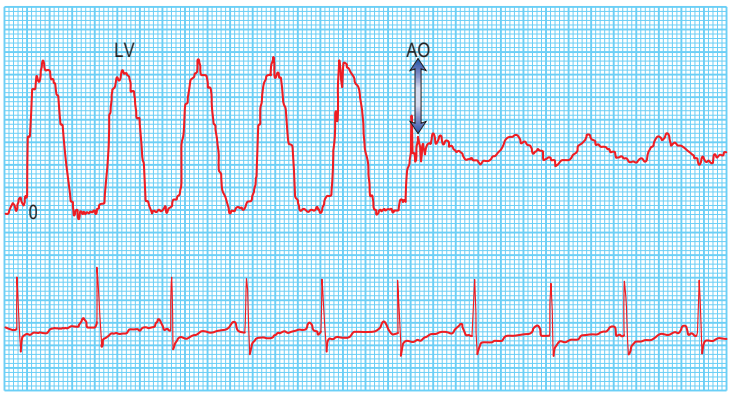


Fig. 13.83 **ECG and pressure trace as a cardiac catheter is withdrawn** from the left ventricle (LV) to the aorta (AO). Note that the peak systolic pressure changes from 250 to 130 mmHg (arrow). The 120 mmHg peak-to-peak systolic gradient indicates severe aortic valvular stenosis.

www.studentconsult.com
Ask the authors

Table 13.36 Causes and associations of aortic regurgitation

Acute aortic regurgitation	Chronic aortic regurgitation
Acute rheumatic fever	Rheumatic heart disease
Infective endocarditis	Syphilis
Dissection of the aorta	Arthritides: Reiter's syndrome Ankylosing spondylitis Rheumatoid arthritis
Ruptured sinus of Valsalva aneurysm	Hypertension (severe)
Failure of prosthetic valve	Bicuspid aortic valve
	Aortic endocarditis
	Marfan's syndrome
	Osteogenesis imperfecta

ously damaged valve. This can be a congenitally abnormal valve (e.g. a bicuspid valve) or one damaged by rheumatic fever. There are numerous other causes and associations (Table 13.36). The majority of patients with aortic regurgitation are men (75%), but rheumatic aortic regurgitation occurs more commonly in women.

Pathophysiology

Aortic regurgitation is reflux of blood from the aorta through the aortic valve into the left ventricle during diastole. If net cardiac output is to be maintained, the total volume of blood pumped into the aorta must increase, and consequently the left ventricular size must enlarge. Because of the aortic run-off during diastole, diastolic blood pressure falls and coronary perfusion is decreased. In addition, the larger left ventricular size is mechanically less efficient so that the demand for oxygen is greater and cardiac ischaemia develops.

Symptoms

In aortic regurgitation, significant symptoms occur late and do not develop until *left ventricular failure* occurs. As with mitral regurgitation, a common symptom is '*pounding of the heart*' because of the increased left ventricular size and its vigorous pulsation. *Angina pectoris* is a frequent complaint. Varying grades of *dyspnoea* occur depending on the extent of left ventricular dilatation and dysfunction. *Arrhythmias* are relatively uncommon.

Signs (see Clinical memo in Fig. 13.82)

The signs of aortic regurgitation are many and are due to the hyperdynamic circulation, reflux of blood into the left ventricle and the increased left ventricular size.

The pulse is bounding or collapsing (see p. 690). The following signs, which are rare, also indicate a hyperdynamic circulation:

- Quincke's sign – capillary pulsation in the nail beds
- De Musset's sign – head nodding with each heart beat
- Duroziez's sign – a to-and-fro murmur heard when the femoral artery is auscultated with pressure applied distally (if found, it is a sign of severe aortic regurgitation)
- pistol shot femorals – a sharp bang heard on auscultation over the femoral arteries in time with each heart beat.

The apex beat is displaced laterally and downwards and is forceful in quality. On auscultation, there is a high-pitched early diastolic murmur best heard at the left sternal edge in the fourth intercostal space with the patient leaning forward and the breath held in expiration.

Because of the volume overload there is commonly an ejection systolic flow murmur. The regurgitant jet can impinge on the anterior mitral valve cusp, causing a mid-diastolic murmur (Austin Flint).

Investigations

Chest X-ray

The chest X-ray features are those of left ventricular enlargement and possibly of dilatation of the ascending aorta. The ascending aortic wall may be calcified in syphilis, and the aortic valve may be calcified if valvular disease is responsible for the regurgitation.

Electrocardiogram

The ECG appearances are those of left ventricular hypertrophy due to 'volume overload' – tall R waves and deeply inverted T waves in the left-sided chest leads, and deep S waves in the right-sided leads. Normally, sinus rhythm is present.

Echocardiogram

The echocardiogram demonstrates vigorous cardiac contraction and a dilated left ventricle. The aortic root may also be enlarged. Diastolic fluttering of the mitral leaflets or septum occurs in severe aortic regurgitation (producing the Austin Flint murmur). The regurgitant jet can be detected by CW Doppler.

Cardiac catheterization

During cardiac catheterization, injection of contrast medium into the aorta (aortography) will outline aortic valvular abnormalities and allow assessment of the degree of regurgitation.

Treatment

The underlying cause of aortic regurgitation (e.g. syphilitic aortitis or infective endocarditis) may require specific treatment. The treatment of aortic regurgitation usually requires aortic valve replacement but the timing of surgery is critical.

Because symptoms do not develop until the myocardium fails and because the myocardium does not recover fully after surgery, operation is performed before significant symptoms occur. The timing of the operation is best determined according to haemodynamic, echocardiographic or angiographic criteria.

Both mechanical prostheses and tissue valves are used. Tissue valves are preferred in the elderly and when anticoagulants must be avoided, but are contraindicated in children and young adults because of the rapid calcification and degeneration of the valves.

Antibiotic prophylaxis against infective endocarditis (see p. 93) is sometimes necessary if a prosthetic valve replacement has been performed.

TRICUSPID STENOSIS

This uncommon valve lesion, which is seen much more often in women than in men, is usually due to rheumatic heart

disease and is frequently associated with mitral and/or aortic valve disease. Tricuspid stenosis is also seen in the carcinoid syndrome.

Pathophysiology

Tricuspid valve stenosis results in a reduced cardiac output, which is restored towards normal when the right atrial pressure increases. The resulting systemic venous congestion produces hepatomegaly, ascites and dependent oedema.

Symptoms

Usually, patients with tricuspid stenosis complain of symptoms due to associated left-sided rheumatic valve lesions. The *abdominal pain* (due to hepatomegaly) and swelling (due to ascites), and *peripheral oedema* that occur are relatively severe when compared with the degree of *dyspnoea*.

Signs

If the patient remains in sinus rhythm, which is unusual, there is a prominent jugular venous *a* wave. This pre-systolic pulsation may also be felt over the liver. There is usually a rumbling mid-diastolic murmur, which is heard best at the lower left sternal edge and is louder on inspiration. It may be missed because of the murmur of coexisting mitral stenosis. A tricuspid opening snap may occasionally be heard.

Hepatomegaly, abdominal ascites and dependent oedema may be present.

Investigations

Chest X-ray

On the chest X-ray there may be a prominent right atrial bulge.

Electrocardiogram

The enlarged right atrium may be manifested on the ECG by peaked, tall P waves (>3 mm) in lead II.

Echocardiogram

The echocardiogram may show a thickened and immobile tricuspid valve, but this is not so clearly seen as an abnormal mitral valve.

Cardiac catheterization

This demonstrates a diastolic pressure gradient between the right atrium and the right ventricle. Contrast injection will demonstrate a large right atrium.

Treatment

Medical management consists of diuretic therapy and salt restriction. Tricuspid valvotomy is occasionally possible, but tricuspid valve replacement is often necessary. Other valves usually also need replacement because tricuspid valve stenosis is rarely an isolated lesion.

TRICUSPID REGURGITATION

Functional tricuspid regurgitation may occur whenever the right ventricle dilates, e.g. in cor pulmonale, myocardial infarction or pulmonary hypertension.

Organic tricuspid regurgitation may occur with rheumatic heart disease, infective endocarditis, carcinoid syndrome, Ebstein's anomaly (a congenitally malpositioned tricuspid valve) and other congenital abnormalities of the atrioventricular valves.

Symptoms and signs

The valvular regurgitation gives rise to high right atrial and systemic venous pressure. Patients may complain of the symptoms of *right heart failure* (see p. 736).

Physical signs include a large jugular venous '*cv*' wave and a palpable liver that pulsates in systole. Usually a right ventricular impulse may be felt at the left sternal edge, and there is a blowing pansystolic murmur, best heard on inspiration at the lower left sternal edge. Atrial fibrillation is common.

Treatment

Functional tricuspid regurgitation usually disappears with medical management. Severe organic tricuspid regurgitation may require operative repair of the tricuspid valve (annuloplasty or plication). Very occasionally, tricuspid valve replacement may be necessary. In drug addicts with infective endocarditis of the tricuspid valve, surgical removal of the valve is recommended to eradicate the infection. This is usually well tolerated in the short term. The insertion of a prosthetic valve for this condition is considered on page 769.

PULMONARY STENOSIS

This is usually a congenital lesion, but it may rarely result from rheumatic fever or from the carcinoid syndrome. Congenital pulmonary stenosis may be associated with an intact ventricular septum or with a ventricular septal defect (Fallot's tetralogy).

Pulmonary stenosis may be valvular, subvalvular or supravalvular. Multiple congenital pulmonary arterial stenoses are usually due to infection with rubella during pregnancy.

Symptoms and signs

The obstruction to right ventricular emptying results in right ventricular hypertrophy which in turn leads to right atrial hypertrophy. Severe *pulmonary obstruction* may be incompatible with life, but lesser degrees of obstruction give rise to *fatigue*, *syncope* and the symptoms of *right heart failure*. Mild pulmonary stenosis may be asymptomatic.

The physical signs are characterized by a harsh midsystolic ejection murmur, best heard on inspiration, to the left of the sternum in the second intercostal space. This murmur is often associated with a thrill. The pulmonary closure sound is usually delayed and soft. There may be a pulmonary ejection sound if the obstruction is valvular. A right ventricular fourth sound and a prominent jugular venous *a* wave are present when the stenosis is moderately severe. A right ventricular heave (sustained impulse) may be felt.

Investigations

Chest X-ray

The chest X-ray usually shows a prominent pulmonary artery owing to post-stenotic dilatation.

Electrocardiogram
The ECG demonstrates both right atrial and right ventricular hypertrophy, although it may sometimes be normal even in severe pulmonary stenosis.

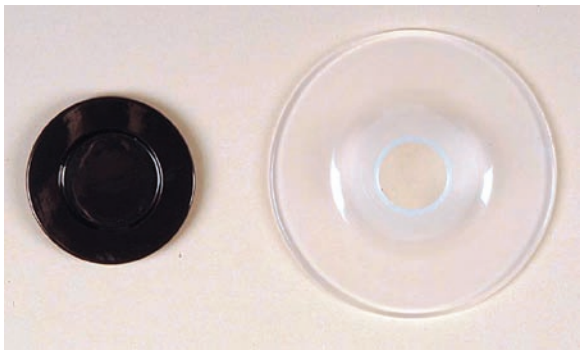
Echocardiogram
Doppler is the investigation of choice.

Cardiac catheterization
The passage of a catheter through the right heart allows the level and degree of the stenosis to be established by measuring the systolic pressure gradient.

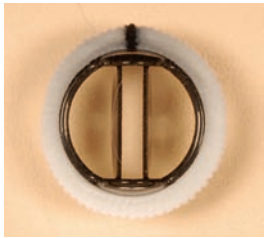
Treatment
Treatment of severe pulmonary stenosis requires pulmonary valvotomy (balloon valvotomy or direct surgery).

PULMONARY REGURGITATION
This is the most common acquired lesion of the pulmonary valve. It results from dilatation of the pulmonary valve ring, which occurs with pulmonary hypertension. It is characterized by a decrescendo diastolic murmur, beginning with the pulmonary component of the second sound, that is difficult to distinguish from the murmur of aortic regurgitation. Pulmonary regurgitation usually causes no symptoms and treatment is rarely necessary.

PROSTHETIC VALVES
There is no ideal replacement for our own normally functioning, native heart valves. There are two options for valve prostheses: mechanical or tissue (bioprosthetic) (Fig. 13.84).



(a)



(b)

(c)

Fig. 13.84 Prosthetic valves. **(a)** Bjork–Shiley mechanical prosthetic valve. **(b)** St Jude double tilting disc. **(c)** Aortic valve tissue prosthesis (aortic view).

The valves consist of two basic components: an opening to allow blood to flow through and an occluding mechanism to regulate the flow. Mechanical prostheses rely on artificial concluders: a ball and cage (Starr–Edwards), tilting disc (Bjork–Shiley) or double tilting disc (St Jude). Tissue prostheses are derived from human (homograft), or from porcine or bovine (xenograft) origin. A valve replacement from within the same patient (i.e. pulmonary to aortic valve position) is termed an autograft.

Mechanical versus tissue valves
Mechanical valves, being artificial structures, are more durable than their tissue counterparts, which tend to degenerate after 10 years. However, artificial structures are more thrombogenic. Mechanical valves require formal anticoagulation for the lifetime of the prosthesis. Tissue valves only require anticoagulation for a limited postoperative period whilst the suture lines endothelialize; it can then be discontinued unless another risk factor for thromboembolism (e.g. atrial fibrillation) persists. The potential lifetime of the prosthetic valve and the patient's tolerance of long-term anticoagulation can inform the choice of valve to be implanted. On auscultation, tissue valve heart sounds are comparable to those of a native valve. Mechanical valve heart sounds are generally louder and both opening and closing sounds can be heard.

Complications
All prostheses carry a risk of infection. Prosthetic valve endocarditis is associated with significant morbidity and mortality; prevention is the cornerstone of management. Patient education about antibiotic prophylaxis is vital and this should be re-inforced at clinic visits. Any procedure which results in a breach of the body's innate defences (i.e. dental treatment, catheter insertion) increases the risk of exposing the prosthesis to a bacteraemia. This must be borne in mind when managing a patient with a prosthetic heart valve and steps taken to minimize the risk involved. The prosthetic valve occluding mechanism can be interrupted by vegetations, but also by thrombosis and calcification, resulting in either stenosis or regurgitation. The prosthesis can become detached from the valve ring resulting in a *para-prosthetic leak*. Evidence of structural failure can be detected by simple auscultation, with echocardiography as the initial investigation of choice. Transthoracic echocardiography is non-invasive, but scattering of echoes by mechanical valves makes their assessment difficult. Transoesophageal echocardiography provides alternative views and higher image resolution, making it the investigation of choice when prosthetic valve endocarditis is suspected.

INFECTIVE ENDOCARDITIS
Infective endocarditis is an endovascular infection of cardiovascular structures, including cardiac valves, atrial and ventricular endocardium, large intrathoracic vessels and intracardiac foreign bodies, e.g. prosthetic valves, pacemaker leads and surgical conduits. The annual incidence in

Valvular heart disease
Tricuspid regurgitation
Pulmonary stenosis
Pulmonary regurgitation
Prosthetic valves
Infective endocarditis

FURTHER READING

Enrique-Sarano M, Tajik AJ. Aortic regurgitation. *New England Journal of Medicine* 2004; **351**: 1539–1546.

Hayek E, Gring KN, Griffin BP. Mitral valve prolapse. *Lancet* 2005; **365**: 507–518.

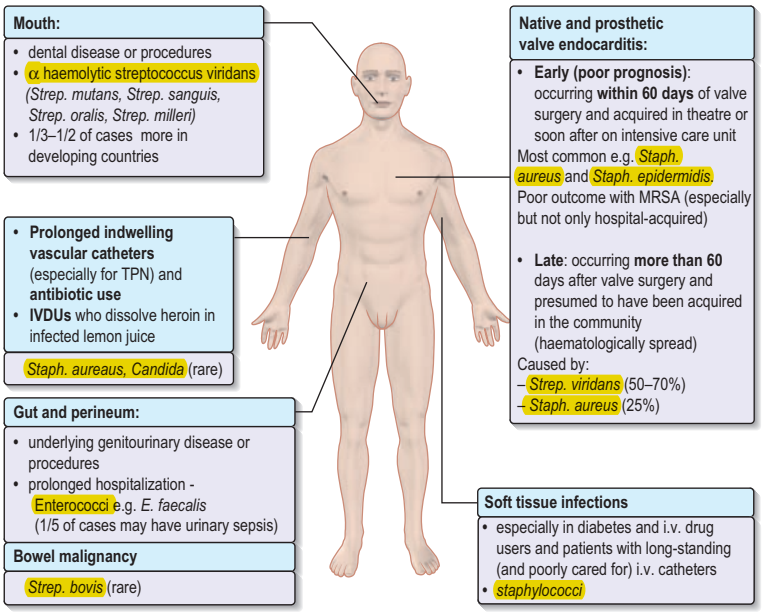


Fig. 13.85 **Infective endocarditis – aetiology and sources of infection.**

the UK is 6–7 per 100 000, but it is more common in developing countries. Without treatment the mortality approaches 100% and even with treatment there is a significant morbidity and mortality.

Aetiology

Endocarditis is usually the consequence of two factors: the presence of organisms in the bloodstream and abnormal cardiac endothelium facilitating their adherence and growth.

Bacteraemia may occur due to patient-specific reasons (poor dental hygiene, intravenous drug use, soft tissue infections) or be associated with diagnostic or therapeutic procedures (dental treatment, intravascular cannulae, cardiac surgery, or permanent pacemakers). Although bacteraemia may occur there is no good evidence that it leads to infective endocarditis (p. 773).

Damaged endocardium promotes platelet and fibrin deposition which allows organisms to adhere and grow, leading to an infected vegetation. Valvular lesions may create non-laminar flow, and jet lesions from septal defects or a patent ductus arteriosus result in abnormal vascular endothelium. Aortic and mitral valves are most commonly involved in infective endocarditis apart from intravenous drug users in whom right sided lesions are more common.

Organisms

Common organisms and the sources of infection are shown in Figure 13.85.

Rare causes

These include the HACEK group of organisms which tend towards a more insidious course (see Box 13.1).

Culture negative endocarditis

This accounts for 5–10% of endocarditis cases. The usual cause is prior antibiotic therapy (good history taking is vital) but some cases are due to a variety of fastidious organisms that fail to grow in normal blood cultures. These include *Coxiella burnetti* (the cause of Q fever), *Chlamydia* species, *Bartonella* species (organisms that cause trench fever and cat scratch disease) and *Legionella*.

Table 13.37 Clinical features of infective endocarditis

	Approximate %
General	
Malaise	95
Clubbing	10
Cardiac	
Murmurs	90
Cardiac failure	50
Arthralgia	25
Pyrexia	90
Skin lesions	
Osler's nodes	15
Splinter haemorrhages	10
Janeway lesions	5
Petechiae	50
Eyes	
Roth spots	5
Conjunctival splinter haemorrhages	Rare
Splenomegaly	40
Neurological	
Cerebral emboli	20
Mycotic aneurysm	10
Renal	
Haematuria	70

Clinical presentation (Table 13.37)

The clinical presentation of infective endocarditis is dependent on the organism and the presence of predisposing cardiac conditions. Infective endocarditis may occur as an acute, fulminating infection but also occurs as a chronic or subacute illness with low-grade fever and non-specific

symptoms. A high index of clinical suspicion is required to identify patients with infective endocarditis and certain criteria should alert the physician.

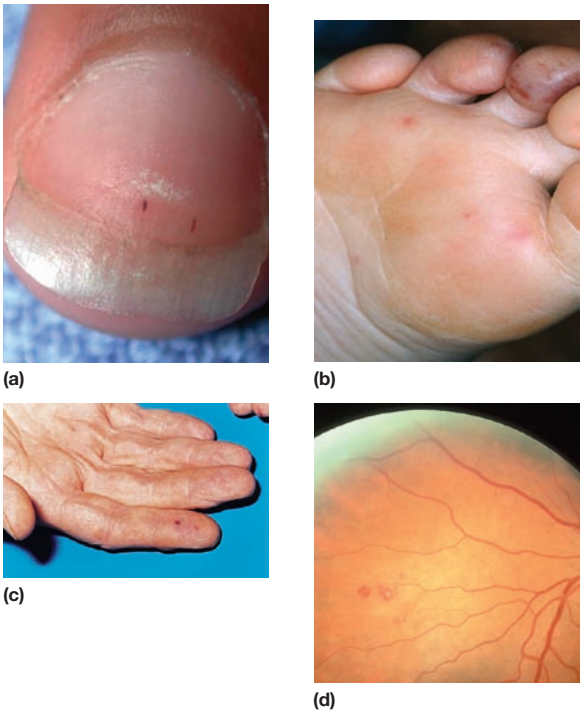
High clinical suspicion:

- new valve lesion/(regurgitant) murmur
- embolic event(s) of unknown origin
- sepsis of unknown origin
- haematuria, glomerulonephritis and suspected renal infarction
- 'fever' plus:
  - prosthetic material inside the heart
  - other high predisposition for infective endocarditis, e.g. i.v. drug abuse
  - newly developed ventricular arrhythmias or conduction disturbances
  - first manifestation of congestive cardiac failure
  - positive blood cultures (with typical organism)
  - cutaneous (Osler, Janeway) or ophthalmic (Roth) manifestations (Fig. 13.86)
  - peripheral abscesses (renal, splenic, spine) of unknown origin
  - predisposition and recent diagnostic / therapeutic interventions known to result in significant bacteraemia.

Low clinical suspicion. Fever plus none of the above.

Diagnostic criteria

The criteria for the clinical diagnosis of endocarditis have been established – the modified Duke criteria (Box 13.1).



(a) (b) (c) (d)

Fig. 13.86 **Infective endocarditis. (a)** Splinter haemorrhages. **(b)** Janeway lesions. **(c)** Osler's nodes. **(d)** Roth spots. (a,b) From Moser DK, Riegel B. *Cardiac Nursing.* Philadelphia: Saunders, 2007: 1127, with permission from Elsevier. (c) From Forbes CD, Jackson WF. *Color Atlas and Text of Clinical Medicine*. St Louis: Mosby, 2003: 232, copyright Elsevier. (d) Courtesy of Professor Ian Constable.

Investigations

Investigations are required to confirm the diagnosis of infective endocarditis; to identify the organism to ensure appropriate therapy; and to monitor the patient's response to therapy.

Box 13.1 Modified Duke criteria for endocarditis*

After Raoult D, Abbara S, Jassal DS, Kradin RL. Case records of the Massachusetts General Hospital. Case 5-2007. A 53-year-old man with a prosthetic aortic valve and recent onset of fatigue, dyspnea, weight loss, and sweats. *New England Journal of Medicine* 2007; **356**: 715. Copyright © 2007 Massachusetts Medical Society. All rights reserved.

Major criteria

- A positive blood culture for infective endocarditis, as defined by the recovery of a typical microorganism from two separate blood cultures in the absence of a primary focus (viridans streptococci, *Abiotrophia* species and *Granulicatella* species; *Streptococcus bovis*, HACEK group†, or community-acquired *Staphylococcus aureus* or enterococcus species), or
- A persistently positive blood cultures, defined as the recovery of a microorganism consistent with endocarditis from either blood samples obtained more than 12 hours apart or all three or a majority of four or more separate blood samples, with the first and last obtained at least 1 hour apart, or
- A positive serological test for Q fever, with an immunofluorescence assay showing phase 1 IgG antibodies at a titre >1:800, or
- Echocardiographic evidence of endocardial involvement:
  - an oscillating intracardiac mass on the valve or supporting structures, in the path or regurgitant jets, or on implanted material in the absence of an alternative anatomical explanation; or
  - an abscess; or
  - new partial dehiscence of prosthetic valve; or
- New valvular regurgitation.

Minor criteria

- Predisposition: predisposing heart condition or intravenous drug use.
- Fever: temperature ≥38°C (100.4°F).
- Vascular phenomena: major arterial emboli, septic pulmonary infarcts, mycotic aneurysm, intracranial haemorrhage, conjunctival haemorrhages, Janeway's lesion.
- Immunologic phenomena: glomerulonephritis, Osler's nodes, Roth's spots, rheumatoid factor.
- Microbiological evidence: a positive blood culture but not meeting a major criterion as noted above, or serological evidence of an active infection with an organism that can cause infective endocarditis.‡
- Echocardiogram: findings consistent with infective endocarditis but not meeting a major criterion as noted above.

*The diagnosis of infective endocarditis is definite when: (a) a microorganism is demonstrated by culture of a specimen from a vegetation, an embolism or an intracardiac abscess; (b) active endocarditis is confirmed by histological examination of the vegetation or intracardiac abscess; (c) two major clinical criteria, one major and three minor criteria, or five minor criteria are met.

†HACEK denotes *Haemophilus* species, *Actinobacillus actinomycetemcomitans*, *Cardiobacterium hominis*, *Eikenella corrodens* and *Kingella kingae*.

‡Excluded from this criterion is a single positive blood culture for coagulase-negative staphylococci or other organisms that do not cause endocarditis. Serological tests for organisms that cause endocarditis include tests for *Brucella*, *Coxiella burnetii*, *Chlamydia*, *Legionella* and *Bartonella* species.

Microbiology

- **Blood cultures** are the *key diagnostic investigation* in infective endocarditis. At least three sets of samples (i.e. six bottles) should be taken from different venepuncture sites. Do not use indwelling catheter lines. Liaise with the microbiology department. The yield of any test is increased by the amount of information given about the subject.
- **Serological tests** can be sent when the diagnosis is suspected and the blood cultures are negative. They aid diagnosis in cases where the organisms will not grow in standard blood cultures (i.e. *Coxiella, Bartonella, Legionella* and *Chlamydia*).

Other laboratory tests

- **Full blood count**. A mild normochromic normocytic anaemia and polymorphonuclear leucocytosis are common. Thrombocytopenia or thrombocytosis can occur.
- **Urea and electrolytes**. Renal dysfunction is common in sepsis.
- **Liver biochemistry**. Serum alkaline phosphatase may be increased.
- **Inflammatory markers**. C-reactive protein (CRP) and erythrocyte sedimentation rate are increased in any infection. CRP may be useful in monitoring response to therapy (and any relapse).
- **Urine**. Proteinuria and haematuria occur frequently.
- **Polymerase chain reaction (PCR)** may be useful in culture-negative infective endocarditis.

Electrocardiogram

This may show evidence of myocardial infarction (emboli) or conduction defects. New atrioventricular block is suggestive of abscess formation. Patients with suspected infective endocarditis therefore should have an ECG on presentation and repeated regularly during their admission depending on their clinical course.

Chest X-ray

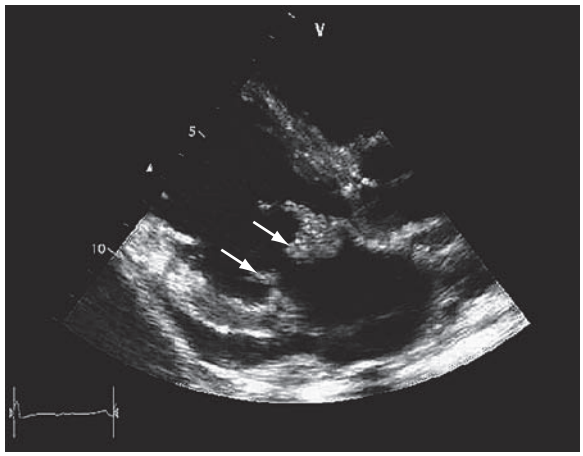
This may show evidence of heart failure or, in right-sided endocarditis, multiple pulmonary emboli and/oral abscesses. The combination of sepsis and pulmonary infiltrates on chest X-ray should alert the clinician to the possibility of right-sided endocarditis.

Echocardiography

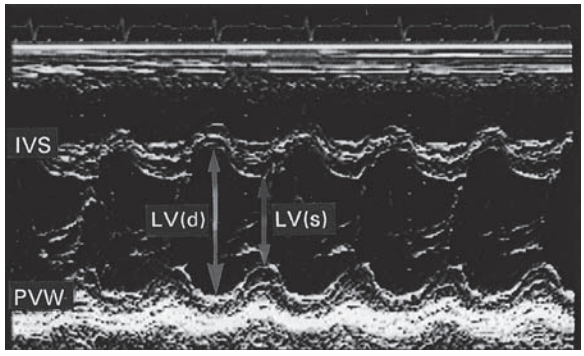
Transthoracic echocardiography (TTE) is rapid, non-invasive and has high specificity for visualizing vegetations (Fig. 13.87), although sensitivity is 60–75%. It is also useful in documenting valvular dysfunction and other local complications, such as aortic root abscesses.

Transoesophageal echocardiography (TOE) has a higher sensitivity (>90%) and specificity for abscess formation because of the close physical proximity of the transducer to the aortic root. TOE also enhances the visualization of prosthetic valves and is recommended for all cases of suspected prosthetic valve endocarditis.

Echocardiography is an extremely useful tool if used appropriately. A negative echocardiogram does not exclude a diagnosis of endocarditis. It is not an appropriate screening test for patients with just a fever or an isolated positive blood culture, where there is a low pre-test probability of endocarditis.



(a)



(b)

Fig. 13.87 **Infective endocarditis. (a) Two-dimensional echocardiogram (long-axis view) showing vegetations** (arrows) attached to both the anterior and posterior leaflets of the mitral valve in a patient with infective endocarditis. **(b) M-mode echocardiogram of the left ventricle** demonstrating hyperdynamic contraction associated with volume overloading from severe mitral regurgitation in a patient with endocarditis. IVS, interventricular septum; LV(d) and LV(s), diastolic and systolic left ventricular dimensions; PVW, posterior ventricular wall.

Treatment

The location of the infection means that prolonged courses of antibiotics are usually required in the treatment of infective endocarditis. The combination of antibiotics may be synergistic in eradicating microbial infection and minimizing resistance. Blood cultures should be taken prior to empirical antibiotic therapy (but this should not delay therapy in unstable patients. Antibiotic treatment should continue for 4–6 weeks. Typical therapeutic regimens are shown in Table 13.38 but specific therapy should be sought from the local microbiology department according to the organism identified and current sensitivities. Serum levels of gentamicin and vancomycin need to be monitored to ensure adequate therapy and prevent toxicity. In patients with penicillin allergy one of the glycopeptide antibiotics, vancomycin or teicoplanin, can be used. Penicillins, however, are fundamental to the therapy of bacterial endocarditis; allergies therefore seriously compromise the choice of antibiotics. It is essential to confirm the nature of a patient's allergy to ensure that the appropriate treatment is not withheld needlessly. Anaphy-

FURTHER READING

Baddour LM, Wilson WR, Bayer AS, et al. Infective endocarditis: diagnosis, antimicrobial therapy, and management of complications: a statement for healthcare professionals from the Committee on Rheumatic Fever, Endocarditis, and Kawasaki Disease, Council on Cardiovascular Disease in the Young, and the Councils on Clinical Cardiology, Stroke, and Cardiovascular Surgery and Anesthesia, American Heart Association: endorsed by the Infectious Diseases Society of America. *Circulation* 2005; **111**: e394–434.

Table 13.38 Antibiotics in endocarditis (adapted from British Society for Antimicrobial Chemotherapy (BSAC) guidelines)

Clinical situation	Suggested antibiotic regimen to start (all given i.v.)
Clinical endocarditis, culture results awaited, no suspicion of staphylococci	Penicillin 1.2 g 4-hourly, gentamicin 80 mg 12-hourly
Suspected staphylococcal endocarditis (IVDU, recent intravascular devices or cardiac surgery, acute infection)	Vancomycin 1 g 12-hourly, gentamicin 80–120 mg 8-hourly
Streptococcal endocarditis (penicillin sensitive)	Penicillin 1.2 g 4-hourly, gentamicin 80 mg 12-hourly
Enterococcal endocarditis (no high-level gentamicin resistance)	Ampicillin/amoxicillin 2 g 4-hourly, gentamicin 80 mg 12-hourly
Staphylococcal endocarditis*	Vancomycin 1 g 12-hourly, OR Flucloxacillin 2 g 4-hourly, OR Benzylpenicillin 1.2 g 4-hourly, PLUS gentamicin 80–120 mg 8-hourly

Note:
1. Monitor vancomycin and gentamicin levels, and adjust if necessary.
2. Choice of antibiotic for staphylococci depends on sensitivities.
3. Optimum choice of therapy needs close liaison with Microbiology/Infectious Diseases.
All antibiotics given i.v.
IVDA, intravenous drug abuse.
*NB: MRSA can affect valves.

laxis would be much more influential in antibiotic choice than a simple gastrointestinal disturbance.

Persistent fever

Most patients with infective endocarditis should respond within 48 hours of initiation of appropriate antibiotic therapy. This is evidenced by a resolution of fever, reduction in serum markers of infection and relief of systemic symptoms of infection. Failure of this to occur needs to be taken very seriously. The following should be considered:

- perivalvular extension of infection and possible abscess formation
- drug reaction (the fever should promptly resolve after drug withdrawal)
- nosocomial infection (i.e. venous access site, urinary tract infection)
- pulmonary embolism (secondary right-sided endocarditis or prolonged hospitalization).

In such cases, samples for culture should be taken from all possible sites and evidence sought for the above causes. Changing antibiotic dosage or regimen should be avoided unless there are positive cultures or a drug reaction is suspected. Emergence of bacterial resistance is uncommon. Close liaison with microbiology is recommended and a cardiothoracic surgical opinion should be sought.

Surgery

Decisions about surgical intervention in patients with infective endocarditis should be made after joint consultation between the cardiologist and cardiothoracic surgeon, taking into account patient-specific (age, non-cardiac morbidities, presence of prosthetic material or cardiac failure) and infective endocarditis features (infective organism, vegetation size, presence of perivalvular infection, systemic embolization).

Prevention (p. 93)

The value of antibiotic prophylactic therapy and the optimal regimen have never been established in randomized controlled trials and hence their widespread use has been questioned. The recent NICE guidelines (Table 4.9) suggest no prophylaxis unless there is an infection at the site of gastrointestinal or genitourinary tract procedures in patients at risk of endocarditis; the adverse effects of antibiotics outweigh the risks.

CONGENITAL HEART DISEASE

A congenital cardiac malformation occurs in about 1% of live births. There is an overall male predominance, although some individual lesions (e.g. atrial septal defect and persistent ductus arteriosus) occur more commonly in females. As a result of improved medical and surgical management, more children with congenital cardiac disease are surviving into adolescence and adulthood. Thus there is a need for an increased awareness amongst general physicians and cardiologists of the problems posed by these individuals.

Aetiology

The aetiology of congenital cardiac disease is often unknown, but recognized associations include:

- maternal prenatal rubella infection (persistent ductus arteriosus, and pulmonary valvular and arterial stenosis)
- maternal alcohol abuse (septal defects)
- maternal drug treatment and radiation
- genetic abnormalities (e.g. the familial form of atrial septal defect and congenital heart block)
- chromosomal abnormalities (e.g. septal defects and mitral and tricuspid valve defects are associated with Down's syndrome (trisomy 21) or coarctation of the aorta in Turner's syndrome (45, XO)).

Classification (see Table 13.39)

Symptoms and signs

Congenital heart disease should be recognized as early as possible, as the response is usually better the earlier treatment is initiated. Some symptoms, signs and clinical problems are common in congenital heart disease:

- *Central cyanosis* occurs because of right-to-left shunting of blood or because of complete mixing of systemic and pulmonary blood flow. In the latter case, e.g. Fallot's tetralogy, the abnormality is described as cyanotic congenital heart disease.
- *Pulmonary hypertension* results from large left-to-right shunts. The persistently raised pulmonary flow leads to

Table 13.39 Classification of congenital heart disease

Acyanotic	Cyanotic
With shunts	**With shunts**
Atrial septal defect	Fallot's tetralogy
Ventricular septal defect	Transposition of the great vessels
Patent ductus arteriosus	Severe Ebstein's anomaly
Partial anomalous venous drainage	
	Without shunts
Without shunts	Severe pulmonary stenosis
Coarctation of the aorta	Tricuspid atresia
Congenital aortic stenosis	Pulmonary atresia
	Hypoplastic left heart

Table 13.40 Common congenital lesions

	Percentage of congenital lesions	Occurrence in first-degree relatives (%)
Ventricular septal defect	39	4
Atrial septal defect	10	2
Persistent ductus arteriosus	10	4
Pulmonary stenosis	7	
Coarctation of the aorta	7	2
Aortic stenosis	6	4
Fallot's tetralogy	6	4
Others	15	

the development of increased pulmonary artery vascular resistance and consequent pulmonary hypertension. This is known as the Eisenmenger's reaction (or the Eisenmenger's complex when due specifically to a ventricular septal defect). The development of pulmonary hypertension significantly worsens the prognosis.

- *Clubbing of the fingers* occurs in congenital cardiac conditions associated with prolonged cyanosis.
- *Paradoxical embolism* of thrombus from the systemic veins to the systemic arterial system may occur when a communication exists between the right and left heart. There is therefore an increased risk of cerebrovascular accidents and also abscesses (as with endocarditis).
- *Polycythaemia* can develop secondary to chronic hypoxaemia, leading to a hyperviscosity syndrome and an increased thrombotic risk, e.g. strokes.
- *Growth retardation* is common in children with cyanotic heart disease.
- *Syncope* is common when severe right or left ventricular outflow tract obstruction is present. Exertional syncope, associated with deepening central cyanosis, may occur in Fallot's tetralogy. Exercise increases resistance to pulmonary blood flow but reduces systemic vascular resistance. Thus, the right-to-left shunt increases and cerebral oxygenation falls.
- *Squatting* is the posture adopted by children with Fallot's tetralogy. It results in obstruction of venous return of desaturated blood and an increase in the peripheral systemic vascular resistance. This leads to a reduced right-to-left shunt and improved cerebral oxygenation.

Presentation

Adolescents and adults with congenital heart disease present with specific common problems related to the long-standing structural nature of these conditions and any surgical treatment:

- endocarditis (particularly in association with otherwise innocuous lesions such as small VSDs or bicuspid aortic valve that can give up to 10% lifetime risk)
- progression of valvular lesions (calcification and stenosis of congenitally deformed valves, e.g. bicuspid aortic valve)
- atrial and ventricular arrhythmias (often quite resistant to treatment)
- sudden cardiac death
- right heart failure (especially when surgical palliation results in the right ventricle providing the systemic supply)
- end-stage heart failure (rarely managed by heart or heart–lung transplantation).

Genetic counselling

These conditions necessitate active follow-up of adult patients. Pregnancy is normally safe except if pulmonary hypertension or vascular disease is present, when the prognosis for both mother and fetus is poor.

Table 13.40 lists the most common congenital lesions and their occurrence in first-degree relatives.

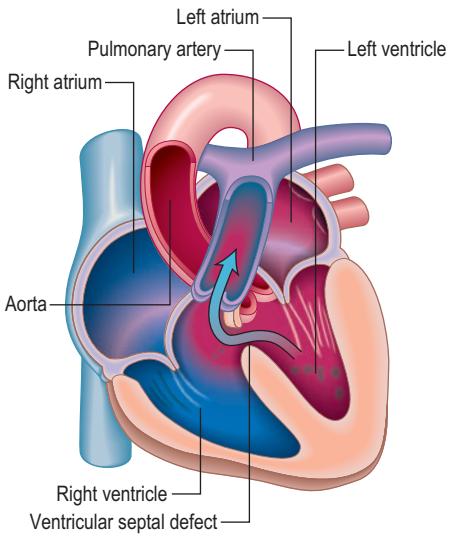
Genetic factors should be considered in all patients presenting with congenital heart disease. For example, parents with a child suffering from Fallot's tetralogy stand a 4% chance of conceiving another child with the disease, and so fetal ultrasound screening of the mother during pregnancy is essential. Parents with congenital heart disease are also more likely to have affected offspring. Fathers have a 2% risk while mothers have a higher risk (around 5%). Individual families can exhibit even higher risks of recurrence.

Ventricular septal defect (VSD)

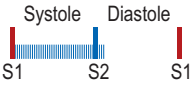
VSD is the most common congenital cardiac malformation (1 in 500 live births). It occurs as an isolated abnormality or in association with other anomalies. Left ventricular pressure is higher than right ventricular pressure; blood therefore moves from left to right and pulmonary blood flow increases. When pulmonary blood flow is very large, progressive obliteration of the pulmonary vasculature eventually causes the pulmonary arterial pressure to equal the systemic pressure (Eisenmenger's complex). Consequently, the shunt is reduced or reversed (becoming right-to-left) and central cyanosis develops.

Clinical features (Fig. 13.88)

Small VSDs ('maladie de Roger') are asymptomatic and usually close spontaneously, with 90% no longer patent by 10 years of age. Unfortunately there is a future risk of the development of aortic regurgitation or endocarditis even after spontaneous closure. A *moderate VSD* produces fatigue and dyspnoea with cardiac enlargement and a prominent apex beat. There is often a palpable systolic thrill at the lower left



Heart sounds and murmurs



Harsh pansystolic murmur heard at the left sternal edge, accompanied by a systolic thrill ± left parasternal heave. May be signs of pulmonary hypertension

Pathophysiology

- Left-to-right shunt
- When the left ventricle contracts, it ejects some blood into the aorta and some across the ventricular septal defect into the right ventricle and pulmonary artery
- **Small** VSDs ('maladie de Roger'): loud and sometimes long systolic murmur
- **Moderate** VSDs: loud 'tearing' pansystolic murmur
- **Large** VSDs: cause pulmonary hypertension and soft murmur Eisenmenger's complex may result

Fig. 13.88 **Ventricular septal defect:** Pathophysiology and auscultatory findings.

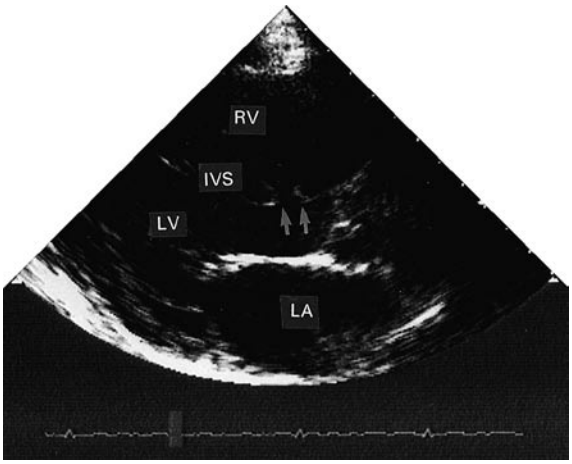


Fig. 13.89 **2-D Echocardiogram (long-axis view) showing a ventricular septal defect** (arrows). Colour Doppler would provide graphic demonstration of the left-to-right shunt. IVS, interventricular septum; LA, left atrium; LV, left ventricle; RV, right ventricle.

sternal edge. A loud 'tearing' pansystolic murmur is heard at the same position. A large VSD eventually causes pulmonary hypertension.

Investigations

A small VSD produces no abnormal X-ray or ECG findings.

- **Chest X-ray** shows a prominent pulmonary artery owing to increased pulmonary blood flow in larger defects. In Eisenmenger's complex the radiological signs of pulmonary hypertension (i.e. 'pruned' pulmonary arteries) can be seen. Cardiomegaly occurs when a moderate or a large VSD is present. ECG shows features of both left and right ventricular hypertrophy.
- **2-D echocardiography and CW Doppler** (Fig. 13.89) can assess the size and location of the VSD, and its haemodynamic consequences.
- **CMR** is increasingly being used.

Treatment

Moderate and large VSDs should be surgically repaired before the development of severe pulmonary hypertension. Infective endocarditis prophylaxis is discussed on page 93.

Atrial septal defect (ASD)

Clinical features

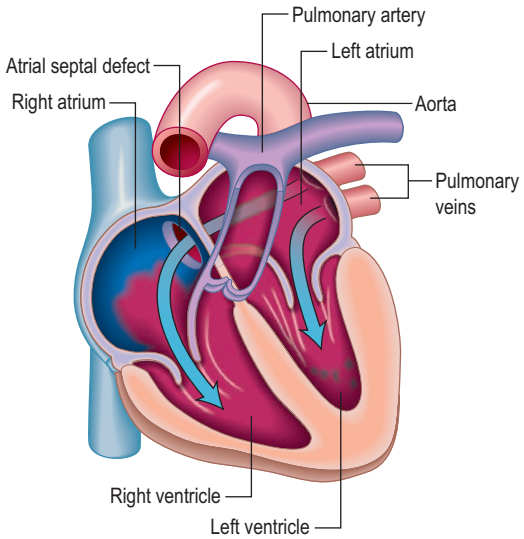
This condition is often first diagnosed in adults and represents one-third of all adult congenital heart disease. It is two to three times more common in women than in men. There are two main types of ASD, *ostium secundum* (75%) and *ostium primum* (15%). The common form is the ostium secundum defect, which involves the fossa ovalis in the atrial mid-septum. (This should not be confused with the *patent foramen ovale* (PFO), which is a normal variant and not a true septal defect. PFO is usually asymptomatic but is associated with paradoxical emboli and an increased incidence of embolic stroke.) Communication at the level of the atria allows left-to-right shunting of blood. Because the pulmonary vascular resistance is low and the right ventricle is easily distended (i.e. it is compliant), there is a considerable increase in right heart output. Above the age of 30 years there may be an increase in pulmonary vascular resistance, which gives rise to *pulmonary hypertension*. Atrial arrhythmias, particularly *atrial fibrillation*, are common at this stage.

Most children with ASD are asymptomatic, although they are prone to *pulmonary infection*. Some complain of *dyspnoea* and *weakness*. Palpitations due to atrial arrhythmias are not uncommon. Right heart failure and atrial fibrillation may develop to become the initial presentation in adult life.

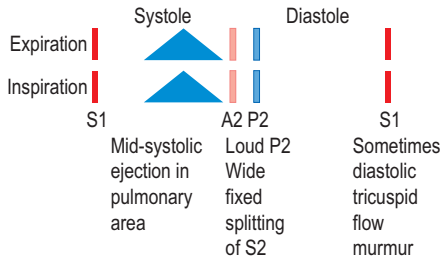
The physical signs of ASD reflect the volume overloading of the right ventricle (Fig. 13.90). A right ventricular heave can usually be felt.

Investigations

- **Chest X-ray** reveals a prominent pulmonary artery and pulmonary plethora. There may be noticeable right ventricular enlargement.



Heart sounds and murmurs



Pathophysiology

- Left-to-right shunt through the defect in the interatrial septum
- Dilatation of the pulmonary artery resulting from the shunt
- The murmur is produced by increased flow across pulmonary valve and an increased stroke volume

Fig. 13.90 **Atrial septal defect:** Pathophysiology and auscultatory findings.

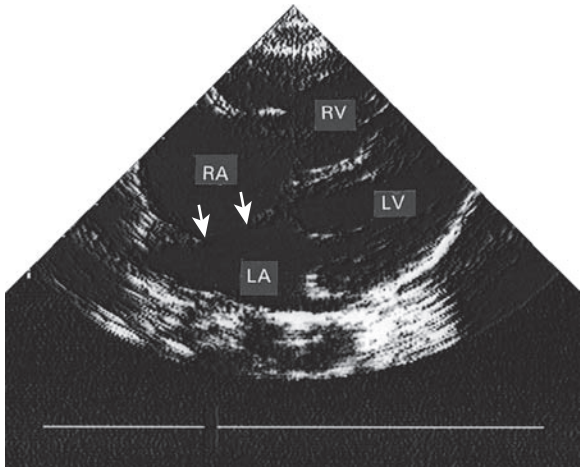


Fig. 13.91 **Ostium secundum atrial septal defect** (arrows) in a young girl, shown by a 2-D echocardiogram subcostal four-chamber view (similar to Fig. 13.89, but rotated clockwise). Colour Doppler can demonstrate the left-to-right shunt. LA, left atrium; LV, left ventricle; RA, right atrium; RV, right ventricle.

Fig. 13.92 **Angiographic appearance of a fully deployed ASD closure device**. The device bridges the ASD and wedges against the surfaces of the right and left atrial septa, occluding flow. The metal object in frame is the distal end of a transoesophageal echocardiography probe. Courtesy of Dr D Ward, St George's, University of London.

- **ECG** usually shows some degree of right bundle branch block (because of dilatation of the right ventricle) and right axis deviation. Sometimes the ASD is part of a major developmental abnormality involving the ventricular septum and the mitral and tricuspid valves. In this case there is left axis deviation on the ECG.
- **Echocardiogram** is usually abnormal if a significant defect is present. Indirect evidence includes right ventricular hypertrophy and pulmonary arterial dilatation, and abnormal motion of the interventricular septum. Subcostal views may demonstrate the ASD (Fig. 13.91). Flow disturbance can be assessed by colour Doppler. A 3-D echocardiogram is shown in Fig. 13.25.
- **CMR** can provide additional information.

Treatment

A significant ASD (i.e. a pulmonary flow that is more than 50% increased when compared with systemic flow) should be repaired before the age of 10 years or as soon as possible if first diagnosed in adulthood. There is a dilemma with regard to whether less severe shunts should be closed when diagnosed in adulthood. Anecdotal evidence suggests that the shunt might be progressive and that low-risk closure would be appropriate.

There is a good result from surgery with osteum secundum ASD unless pulmonary hypertension has developed. Angiographic closure is now possible with significantly lower risk using a transcatheter clamshell device (Fig. 13.92).

A PFO discovered in a patient with an otherwise unexplained thrombotic stroke is now closed in this way to prevent paradoxical thromboembolism. In high-risk groups, e.g. deep-sea divers, such closures may be undertaken even though the patients are asymptomatic. Uncorrected ASDs do not require antibiotic prophylaxis for endocarditis. If there is an accompanying valvular lesion, however, prophylaxis may be indicated.

Patent ductus arteriosus (PDA)

The ductus arteriosus connects the pulmonary artery at its bifurcation to the descending aorta immediately distal to the subclavian artery. In fetal life the ductus diverts blood away from the unexpanded, and hence high-resistance, pulmonary circulation into the systemic circulation, where the blood is reoxygenated as it passes through the placenta. At birth, the high oxygen in the lungs and the reduced pulmonary vascular resistance trigger closure of the duct. If the duct is malformed (i.e. it does not contain sufficient elastic tissue), it will not close. This is more common in females and is sometimes associated with maternal rubella. Premature babies are often born with persistent ducts that are anatomically normal but are immature in that they lack the mechanism to close. Other associations include continual prenatal hypoxaemia and high-altitude environments.

Because aortic pressure exceeds pulmonary artery pressure throughout the cardiac cycle, a persistent duct produces continuous aorta-to-pulmonary artery shunting. This leads to an increased pulmonary venous return to the left heart and an increased left ventricular volume load. If the shunt is large, this results in severe left heart failure and pulmonary hypertension. One-third of individuals with an unrepaired ductus die from heart failure, pulmonary hypertension or endocarditis by the age of 40; two-thirds by the age of 60.

Clinical features

There are often no symptoms until later in life when heart failure or infective endocarditis develops.

The characteristic physical sign is a continuous 'machinery' murmur (Fig. 13.93). The peripheral pulse is large in volume ('bounding') because of the increased left heart blood flow and the decompression of the aorta into the pulmonary artery.

Investigations

The aorta and pulmonary arterial system are usually prominent on X-ray, although a small ductus shows no abnormality. There is both a left atrial abnormality and left ventricular hypertrophy on the ECG. With the development of Eisenmenger's reaction, right ventricular hypertrophy may be seen. The echocardiogram shows a dilated left atrium and left ventricle. Right heart changes are apparent in late disease. CMR can be useful.

Treatment

Premature infants with a persistent duct are treated medically with indometacin, which inhibits prostaglandin production and stimulates duct closure. In other cases the duct can be ligated surgically or angiographically with very little risk. Surgery should be performed as soon as possible and not later than the age of 5 years. Closure is inappropriate if pulmonary hypertension is severe.

Coarctation of the aorta

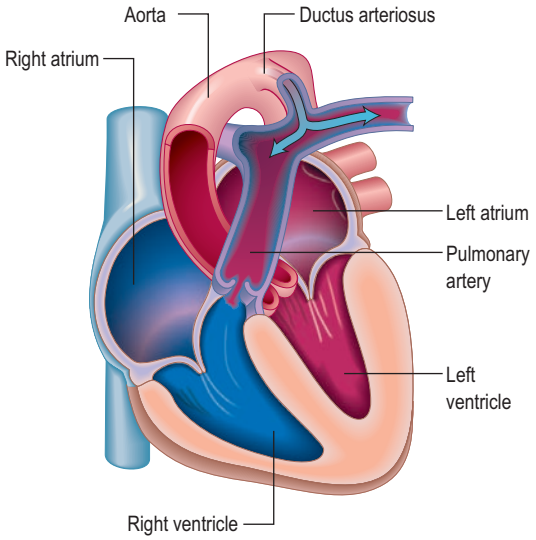
A coarctation of the aorta is a narrowing of the aorta at, or just distal to, the insertion of the ductus arteriosus, i.e. distal to the left subclavian artery (Fig. 13.94). Rarely it can occur proximal to the left subclavian. It occurs twice as commonly in men as in women. It is also associated with Turner's syndrome (p. 1007). In 80% of cases the aortic valve is bicuspid (and potentially stenotic or endocarditic). Other associations include patent ductus arteriosus, ventricular septal defect, mitral stenosis or regurgitation and circle of Willis aneurysms.

Severe narrowing of the aorta encourages the formation of a collateral arterial circulation involving the periscapular and intercostal arteries. Decreased renal perfusion can lead to the development of systemic hypertension that persists even after surgical correction.

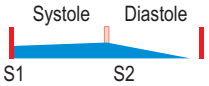
Clinical features

Coarctation of the aorta is often asymptomatic for many years. *Headaches* and *nosebleeds* (due to hypertension), and *claudication* and cold legs (due to poor blood flow in the lower limbs) may be present.

Physical examination reveals hypertension in the upper limbs, and weak, delayed (radiofemoral delay) pulses in the legs. If coarcation is present in the aorta, proximal to the left subclavian artery, there will be asynchronious radial pulses



Heart sounds and murmurs



Continuous 'machinery' murmur best heard below the left clavicle in the first interspace or over the first rib.
A thrill can often be felt

Pathophysiology

- Left-to-right shunt
- Some of the blood from the aorta crosses the ductus arteriosus and flows into the pulmonary artery
- Murmur is produced by the turbulent aortic-to-pulmonary artery shunting in both systole and diastole
- Dilatation of the pulmonary artery, left atrium and left ventricle
- As pulmonary hypertension (Eisenmenger's reaction) develops, the murmur becomes quieter, may be confined to systole or may even disappear causing central cyanosis

Fig. 13.93 **Patent ductus arteriosus:** Pathophysiology and auscultatory findings.

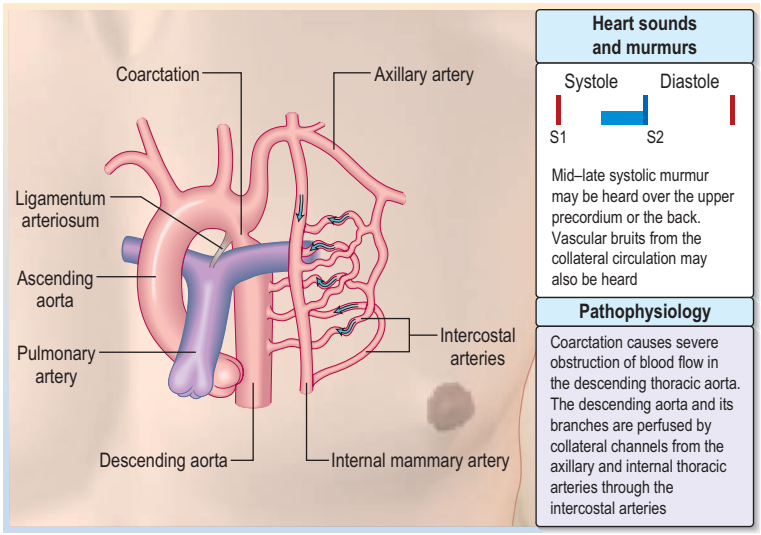


Fig. 13.94 **Coarctation of the aorta:** Pathophysiology and auscultatory findings.

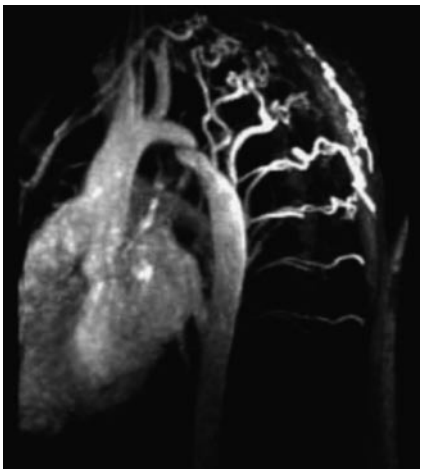


Fig. 13.95 **Cardiac magnetic resonance angiography image of severe aortic coarctation in an adult.**

in right and left arms. Poor peripheral pulses are seen in severe cases.

For heart sounds and murmurs in coarctation of aorta, see Figure 13.94.

Investigations

Chest X-ray may reveal a dilated aorta indented at the site of the coarctation. This is manifested by an aorta (seen in the upper right mediastinum) shaped like a 'figure 3'. In adults, tortuous and dilated collateral intercostal arteries may erode the undersurfaces of the ribs ('rib notching').

ECG demonstrates left ventricular hypertrophy. Echocardiography sometimes shows the coarctation and other associated anomalies. CT and CMR (Fig. 13.95) scanning can accurately demonstrate the coarctation and quantify flow.

Treatment

Treatment is usually indicated if the pressure gradient across the coarctation is greater than 30 mmHg. This involves surgical excision of the coarctation and end-to-end anastomosis of the aorta. If the coarctation is extensive, prosthetic vascular grafts may be needed. When surgery is performed in early childhood, hypertension usually resolves completely. However, when the operation is performed on adolescents or adults, the hypertension persists in 70% because of previous renal damage. There is also an increased risk of accelerated atherosclerosis and strokes in these individuals. Balloon dilatation is used in some centres for either primary disease or post-surgical recurrence, although there is a higher incidence of aneurysm formation and post-dilatation recurrence.

Surgical correction in childhood gives a good 25-year survival rate of 83%. If this is delayed until adulthood (20–40 years), the 25-year survival rate drops to 75%. If coarctation is left uncorrected, however, only 25% of patients are alive at 50, while cardiac failure ensues in two-thirds of surviving patients over 40.

Cyanotic congenital heart disease: Fallot's tetralogy

Most children with cyanotic congenital heart disease do not survive the neonatal period. Fallot's is the most common cyanotic anomaly in those who do survive and is commonest amongst adults. Transposition of the great vessels is more common in the neonatal period but is more likely to be fatal. Other cyanotic congenital heart diseases are shown in Table 13.39.

Fallot's tetralogy consists of the four features shown in Figure 13.96.

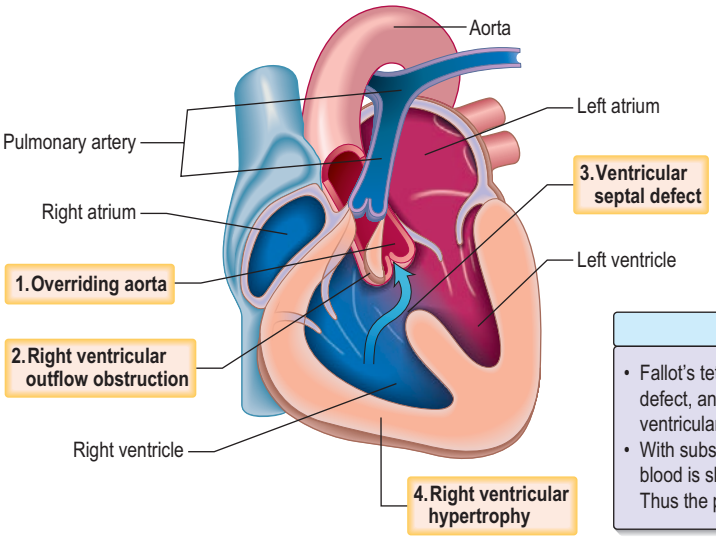
The level of the right ventricular outflow obstruction may be subvalvular, valvular or supravalvular. The most common obstruction is subvalvular, either alone (50%) or in combination with valvular stenosis (25%).

This combination of lesions leads to a high right ventricular pressure and right-to-left shunting of blood through the VSD. Thus the patient is centrally cyanosed.

Clinical features

Children with this condition may present with *dyspnoea* or *fatigue*, or with hypoxic episodes on exertion (Fallot's spells) – *deep cyanosis* and possible *syncope*. These can even result in seizures, cerebrovascular events or sudden death. *Squatting* is common.

Adults tend not to suffer 'spells' but fatigue easily with dyspnoea on exertion. *Erythrocytosis*, secondary to chronic



Heart sounds and murmurs

Systole Diastole

S1 S2

Systolic ejection murmur often associated with a thrill in the second left interspace close to the sternum. Single second heart sound because the pulmonary component is too soft to be heard

Pathophysiology

- Fallot's tetralogy is characterized by a large ventricular septal defect, an aorta that overrides the left and right ventricles, right ventricular outflow tract obstruction, and right ventricular hypertrophy.
- With substantial obstruction of the right ventricular outflow tract, blood is shunted through the ventricular septal defect from right to left. Thus the patient is centrally cyanosed

Fig. 13.96 **Fallot's tetralogy:** Pathophysiology and auscultatory findings.

hypoxaemia, commonly results in thrombotic strokes. *Endocarditis* is common.

A parasternal sustained heave is evident. Central cyanosis is commonly present from birth, and *finger clubbing* and *polycythaemia* are obvious after about 12 months. Growth is usually retarded.

Investigations

- **Chest X-ray** shows a large right ventricle and a small pulmonary artery (classically described as 'boot-shaped').
- **ECG** reveals right ventricular hypertrophy, and the echocardiogram demonstrates discontinuity between the aorta and the anterior wall of the ventricular septum.
- **CMR** is used (p. 706).
- **Cardiac catheterization** is performed to evaluate the size and degree of the right ventricular outflow obstruction.

Treatment

Complete surgical correction of this combination of lesions is possible even in infancy. As adults, however, these individuals are at increased risk of right ventricular failure and ventricular arrhythmias due to the trauma from past corrective surgery. Less damaging procedures are used presently, the long-term results of which are unknown as yet.

Occasionally a palliative procedure – an anastomosis between a subclavian artery and a pulmonary artery (Blalock shunt) – is performed on very young infants or the premature, in order to increase blood supply to the lungs. Without intervention, 66% survive to 1 year, while only 11% survive to 20 years.

Fallot's spells may need treatment with beta-blockade or, when severe, with diamorphine to relax the right ventricular outflow obstruction. Antibiotic prophylaxis for endocarditis is discussed in Chapter 4 (see p. 93).

Transposition of the great arteries (TGA)

TGA occurs in 2–3/10 000 live births with a male/female ratio of 2:1. In nearly 50% of cases it is associated with a VSD.

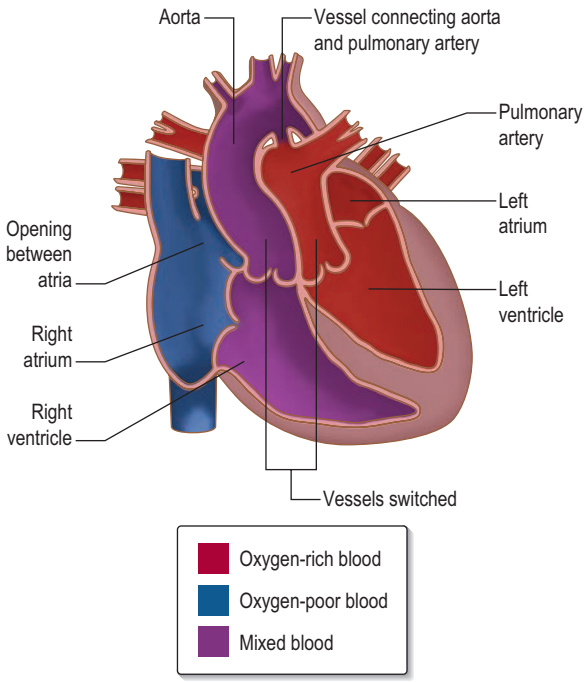


Fig. 13.97 **Transposition of great arteries.** AO, Aorta; LA, left atrium; LV, left ventricle; PA, pulmonary artery; RA, right atrium; RV, right ventricle.

In TGA the right ventricle is connected to the systemic circulation (aorta) and the left ventricle to the pulmonary circulation (pulmonary artery) (Fig. 13.97). The neonate with TGA is cyanosed as the blood circulates in two parallel circuits, i.e. deoxygenated blood from the systemic veins passes to the systemic circulation and oxygenated blood from the pulmonary veins passes back to the lungs.

Clinical features

TGA will present within a few hours of birth as a cyanotic (blue) baby, as maternal/placental oxygenation is removed. Ninety per cent of patients present within a few days. However, in patients with a significant shunt (ASD and VSD or a large PDA) diagnosis may be delayed. Clinical features include dyspnoea, tachycardia and cool clammy skin.

FURTHER READING

Lindsey J, Hillis LD. Clinical update: atrial septal defects in adults. *Lancet* 2007; **369**: 1244–1246.

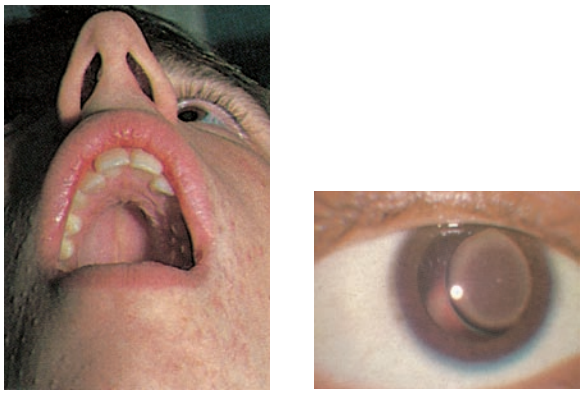
Investigations

2-D Echocardiography can identify the anatomical connections between the great vessels and the cardiac chambers. It can also detect an associated ASD, VSD or PDA.

Treatment

The initial management in a cyanotic neonate with TGA is an atrial septostomy (Rashkind procedure). Balloon dilatation of the foramen ovale is performed to allow mixing of atrial blood that maintains oxygen saturation of 50–80% prior to more definitive surgery. Prostaglandins (E_1) may be given initially to maintain a PDA.

The arterial switch operation is now performed in the first two weeks of life – the aorta is reconnected to the left ventricle; the pulmonary artery is connected to the right ventricle; the coronary arteries are connected to the aorta and any associated defects are repaired.



(a) (b)

Fig. 13.98 **Marfan's syndrome.** High arched palate and eye lens dislocation. From Forbes CD, Jackson WF. *Color Atlas and Text of Clinical Medicine*. St Louis: Mosby, 2003: 134, copyright Elsevier.

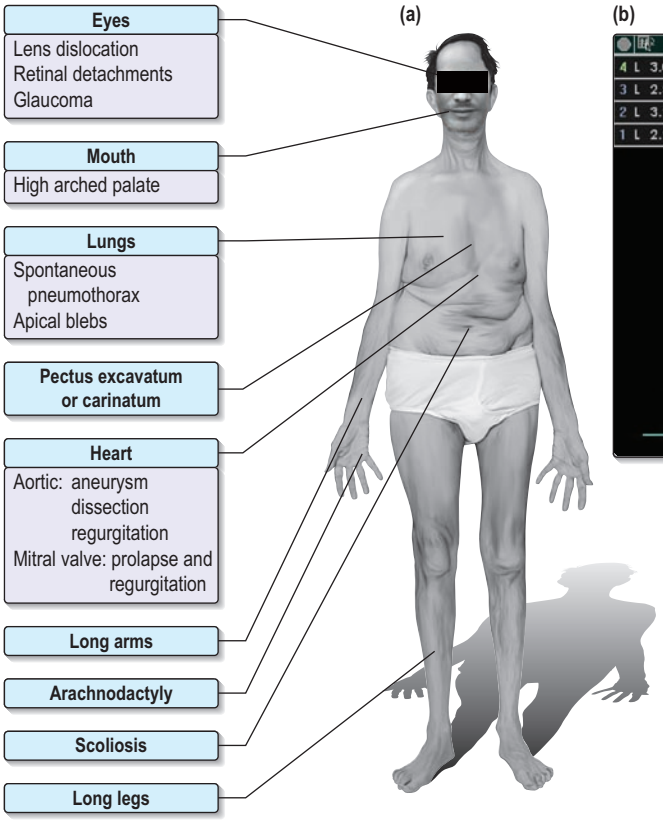
MARFAN'S SYNDROME

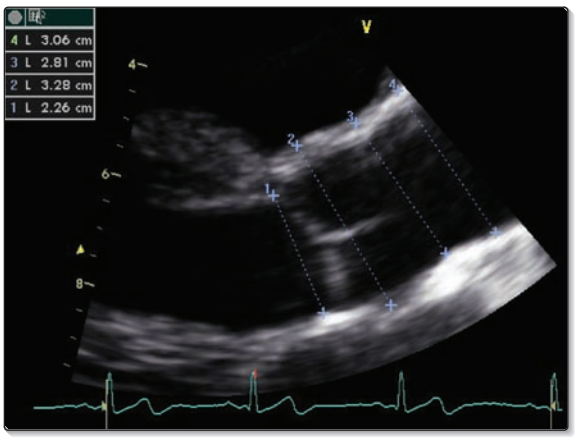
Clinical features

Marfan's syndrome (MFS) is one of the most common autosomal dominant inherited disorders of connective tissue, affecting the heart (aortic aneurysm and dissection, mitral valve prolapse), eye (dislocated lenses, retinal detachment) and skeleton (tall, thin body build with long arms, legs and fingers; scoliosis and pectus deformity) (Figs 13.98 and 13.99; Table 13.41).

Table 13.41 **Yearly risk (%) of complications based on aortic size**

	Aortic size (cm)			
	>3.5	>4	>5	>6
Rupture	0.0	0.3	1.7	3.6
Dissection	2.2	1.5	2.5	3.7
Death	5.9	4.6	4.8	10.8
Any of the above	7.2	5.3	6.5	14.1





(c)

	Aortic size			
Yearly risk	> 3.5cm	> 4cm	> 5cm	> 6cm
Rupture	0.0%	0.3%	1.7%	3.6%
Dissection	2.2%	1.5%	2.5%	3.7%
Death	5.9%	4.6%	4.8%	10.8%
Any of the above	7.2%	5.3%	6.5%	14.1%

Fig. 13.99 **Marfan's syndrome. (a)** Photograph of a 63-year-old man. **(b)** 2-D Longitudinal echocardiogram of the aortic root showing measurements of aortic size.

Clinically, two of three major systems must be affected, to avoid overdiagnosing the condition. Diagnosis may be confirmed by studying family linkage to the causative gene, or by demonstrating a mutation in the Marfan's syndrome gene (*MFS1*) for fibrillin (*FBN-1*) on chromosome 15q21.

MFS affects approximately 1 in 5000 population worldwide and 25% of patients are affected as a result of a new mutation. This group includes many of the more severely affected patients, with high cardiovascular risk. Other known associations with early death due to aortic aneurysm and dissection are: family history of early cardiac involvement; family history of dissection with an aortic root diameter of >5 cm; male sex; and extreme physical characteristics, including markedly excessive stature and widespread striae. Histological examination of aortas often shows widespread medial degeneration, described as 'cystic medial necrosis'.

Cardiac investigations

- **Chest X-ray** is often normal but may show signs of aortic aneurysm and unfolding, or of widened mediastinum. Pneumothorax affects 11% and scoliosis is present in 70% of patients.
- **ECG** may be misleadingly normal with an acute dissection. In conjunction with mitral valve prolapse, 40% of patients usually have arrhythmia, with premature ventricular and atrial arrhythmias.
- **Echocardiography** shows mitral valve prolapse, and mitral regurgitation in the majority of patients. High-quality serial echocardiogram measurements of aortic root diameter in the sinuses of Valsalva, at 90° to the direction of flow are the basis for medical and surgical management (Fig. 13.98b).
- **CT or CMR** can detect aortic dilatation and are useful in monitoring.

Management

- **Beta-blocker therapy** slows the rate of dilatation of the aortic root.
- **ACE receptor blockers**. In Marfan's there is upregulation of TNF-β which is specifically inhibited by ACE blockers. A recent small trial has shown no increase in aortic root diameter on this therapy.
- **Lifestyle alterations** are required because of ocular, cardiac or skeletal involvement. Sports that necessitate prolonged exertion at maximum cardiac output, such as cross-country running, are to be avoided. Sedentary occupations are usually best, as patients tend to suffer from easy fatiguability and hypermobile painful joints.
- **Monitoring** with yearly echocardiograms up to aortic root diameter of 4.5 cm, 6-monthly from 4.5 to 5 cm, and then referred directly to a surgeon who is experienced in aortic root replacement in Marfan's syndrome for elective surgery.

Pregnancy is generally well tolerated if no serious cardiac problems are present, but is preferably avoided if the aortic root diameter is over 4 cm, with aortic regurgitation. Echocardiography should be performed every 6–8 weeks throughout pregnancy and during the initial 6 months postpartum. Blood pressure should be regularly monitored and hypertension treated actively. Delivery should be by the least stressful method; ideally a vaginal delivery. Caesarean section should not be routinely performed. However, if the aortic root is over 4.5 cm, delivery at 39 weeks by induction or caesarean section should be considered. Beta-blocker therapy may be safely instituted or continued throughout pregnancy, to help prevent aortic dissection.

Careful medical and surgical management have increased the overall survival rate. On average, 13 years of life is added when surgical survival is compared to that reported in the natural history of MFS.

Genetic counselling

The condition is inherited in an autosomal dominant mode, with each child of one affected parent having a 50:50 chance of inheriting the condition. Males and females are equally often affected. In 25% of all cases, the condition arises as the result of a spontaneous mutation in gene 5 of one of the parents. Fibrillin-1 gene mutations can be identified in 80% of those affected, confirming diagnosis and aiding prognosis. The mutation can also be used to screen at-risk family members, including postnatal or prenatal offspring.

FURTHER READING

Judge DP, Dietz AC. Marfan's syndrome. *Lancet* 2005; **366**: 1965–1976.

PULMONARY HEART DISEASE

The normal values for mean pulmonary artery pressure (mPAP), mean capillary wedge pressure (mPCWP) and cardiac output (CO) are 12 ± 2 mmHg, 6 ± 2 mmHg and 5 L/min, respectively. The fall in pressure across the lung circulation is known as the transpulmonary gradient and reflects the difference between mPAP and mPCWP. The normal transpulmonary gradient is 6 ± 2 mmHg.

The *pulmonary vascular resistance* (PVR) is calculated by the formula:

$$\frac{\text{mPAP} - \text{mPCWP}}{\text{CO}}$$

It is normally about 1.5 mmHg/L/min (1.5 Wood units). Approximately 60% of the body's endothelial surface is in the lungs and the lungs normally offer a low resistance to blood flow. This is because the media of the precapillary pulmonary arterioles is thin as compared with their more muscular systemic counterparts that have to respond constantly to postural changes under the influence of gravity. The fact that the lung circulation normally offers a low resistance to flow explains the preferential passage of blood through the lungs in specific forms of congenital heart disease, which may eventually lead to remodelling of the lung circulation and pulmonary hypertension.

Pulmonary hypertension

Definition

Pulmonary hypertension is defined as an mPAP of greater than 25 mmHg at rest or of greater than 30 mmHg during exercise. It may present spontaneously with no apparent underlying disease association (and is then known as primary pulmonary hypertension, PPH) or it can occur in association with other disease processes as listed in Table 13.42.

Causes

The most common cause of pulmonary hypertension is chronic obstructive pulmonary disease (COPD). In general, the causes of pulmonary hypertension (Table 13.42) can be subdivided depending on whether their effects on the lung circulation are:

Ask the authors
www.studentconsult.com

Table 13.42 Causes of pulmonary hypertension

Pulmonary vascular disorders
Acute pulmonary thromboembolism (rarely tumour emboli)
Primary pulmonary hypertension
Multiple pulmonary artery stenoses
Pulmonary veno-occlusive disease
Chronic pulmonary thromboembolism
Parasitic infection, e.g. schistosomiasis

Diseases of the lung and parenchyma
COPD
Other chronic lung disorders (see Chapter 14)

Musculoskeletal disorders (causing chronic underventilation)
Kyphoscoliosis
Poliomyelitis
Myasthenia gravis

Disturbance of respiratory control
Obstructive sleep apnoea
Morbid obesity (Pickwickian syndrome)
Cerebrovascular disease

Cardiac disorders
Mitral stenosis
Left ventricular failure
Left atrial myxoma
Congenital heart disease with Eisenmenger's reaction

Miscellaneous
Appetite-suppressant drugs, e.g. dexfenfluramine
Type 1 glycogen storage diseases
Lipid storage diseases, e.g. Gaucher's disease
Rheumatic autoimmune disease, e.g. SLE
Hepatic cirrhosis
Sickle cell disease

- *Precapillary (i.e. in the pulmonary arteries and arterioles).* The most severe elevations in PAP and in PVR occur with these disorders. Examples include PPH, congenital heart disease with Eisenmenger's reaction and multiple pulmonary embolism. In PPH the PCWP is usually normal.
- *Capillary disorders* causing damage to the alveolar capillary mechanisms also cause pulmonary hypertension and these usually fall under the umbrella of parenchymal lung diseases, e.g. COPD and pulmonary fibrosis. Pulmonary arterial hypoxaemia is a potent cause of vasoconstriction and hence of pulmonary hypertension. Any diseases which lead to low pulmonary artery oxygen levels will give rise to pulmonary hypertension.
- *Postcapillary (passive)* processes involve disease in structures distal to the pulmonary capillary bed. They cause pulmonary hypertension through elevated venous pressures. Examples include pulmonary veno-occlusive disease, chronic left ventricular failure and mitral valve disease.

A variety of miscellaneous conditions can also cause pulmonary hypertension, and these include a history of appetite suppressant drug ingestion, and glycogen and lipid storage diseases.

Progressive elevation in PVR ultimately leads to right ventricular dilatation, right ventricular failure and death. When pulmonary hypertension occurs in association with other conditions such as lung disorders or chronic left ventricular failure, the primary underlying cause may have the major influence on mortality and the secondary effects of pulmonary hypertension are additive.

Mechanisms

These depend on the cause and involve hypoxic vasoconstriction, e.g. COPD, decreased surface area of the pulmonary vascular bed, e.g. lung fibrosis, and increased right ventricular volume/pressure, e.g. congenital heart disease.

Following the rise in pulmonary arterial pressure, damage to the pulmonary endothelium leads to excessive release of endothelium-derived vasoconstrictors such as endothelin (ET, see p. 577) which contribute to an increase in PVR. Increased platelet and leucocyte adhesion, elevated serotonin, plasminogen activator inhibitor-1 (PAI-1) and fibrinopeptide A can produce inappropriate intravascular thrombus. ET, angiotensin II and thromboxane A_2 can act as growth factors, giving rise ultimately to vasoconstriction, cell proliferation, fibrosis and smooth muscle hypertrophy. Such vascular remodelling increases the PVR.

The symptoms, signs and investigations of pulmonary hypertension are discussed below.

Management

Since pulmonary hypertension is a condition caused by a heterogeneous group of diseases, treatment is directed at the primary cause, e.g. chest physiotherapy, good nutrition and early antibiotics to treat infection in patients with cystic fibrosis and bronchiectasis.

Surgical or medical therapies to optimize cardiac and respiratory function, supplemental oxygen and anticoagulation are necessary together with newer specific therapies designed at modulating pulmonary vasomotor tone.

Patients who have symptoms resulting in marked limitation of physical activity or those who are symptomatic at rest have a poor median survival (32 and 6 months, respectively). Adverse factors include right ventricular dysfunction and reduced 6-minute walk distance.

Primary pulmonary hypertension

Primary pulmonary hypertension is a condition of unknown cause characterized by clinical, radiological and electrocardiographic evidence of pulmonary hypertension and increased PAP and PVR with normal PCWP. It is estimated that one to two individuals per million per year are diagnosed with PPH compared to five to six individuals per million per year with pulmonary hypertension secondary to rheumatic autoimmune disease. There is an unexplained predominance of females (approximately a 3:1 ratio) with women mainly presenting in the third decade and men in the fourth. Approximately 6–12% of cases are believed to have a familial origin with inheritance in an autosomal dominant fashion.

Some patients with familial PPH have mutations in the gene encoding bone morphogenic protein receptor 2 (*BMPR 2*) on chromosome 2q33. This gene encodes TGF-β, which acts as a growth suppressor and opposes cellular proliferation. Other gene mutations exist.

PPH has also been linked to certain drugs: aminorex fumarate, an amphetamine-like appetite suppressant, talc which is often inhaled with cocaine, fenfluramine and phenteramine in combination (a prescription weight-loss drug).

Plexogenic pulmonary arteriopathy, identical to that seen in PPH, has also been described in patients with rheumatic autoimmune disease (limited cutaneous scleroderma, p. 756). These may have a rapidly progressive and fatal disease associated with pulmonary hypertension.

Patients with *systemic lupus erythematosus* can also develop pulmonary hypertension secondary to vascular remodelling, with identical histopathological features to patients with PPH. Patients with *hepatic cirrhosis* and *Eisenmenger's reaction* also have identical pathology to patients with PPH.

Clinical presentation

Patients with pulmonary hypertension have an insidious onset and indeed often their condition comes to the attention of the physician late in the course of their illness or when symptoms of right ventricular failure develop. Physical examination may reveal findings consistent with pulmonary hypertension and right ventricular overload (Table 13.43). Clinical signs of right ventricular dysfunction may be present.

Chest X-ray may demonstrate enlargement of the pulmonary arteries and the major branches, with marked tapering (pruning) of peripheral arteries. The lung fields are usually lucent and there may be right atrial and right ventricular enlargement. *ECG* may show right ventricular hypertrophy and right atrial enlargement (P pulmonale). *Echocardiography* demonstrates enlarged right ventricular dimensions. *Pulmonary function tests* are usually normal.

The *diagnosis* can be confirmed by CT pulmonary angiography.

The differential diagnoses for PPH are an exclusion of secondary causes as shown in Table 13.42. In particular, mitral stenosis, congenital heart disease with Eisenmenger's reaction, connective tissue disease and sickle cell disease should be excluded.

Treatment

Warfarin and oxygen, followed by oral calcium-channel blockers, is initially used but only helps about 10% of patients. Oral endothelin receptor antagonists (bosentan, sitaxentam), prostanoid analogues (inhaled iloprost, treprostinil, beraprost) or intravenous epoprostenol are also used.

Sildenafil is an orally active phosphodiesterase inhibitor that increases levels of cyclic guanosine monophosphate, resulting in pulmonary arterial vasodilatation and inhibition of smooth muscle cell proliferation. In randomized controlled trials, sildenafil has reduced pulmonary artery pressure and improved exercise capacity, and WHO functional class.

Table 13.43 Clinical signs of pulmonary hypertension

Large '*a*' wave in jugular venous waveform
Right ventricular (parasternal) heave
Pulmonary ejection sound and flow murmur
Loud pulmonary component to second heart sound
Right ventricular fourth heart sound
Right ventricular failure (hepatomegaly, ascites, peripheral oedema)
Prominent '*v*' wave in jugular venous waveform
Right ventricular third heart sound
Tricuspid and pulmonary regurgitation murmurs

Prognosis

Several studies have reported a mean survival of 2–3 years from the time of diagnosis. The cause of death is usually right ventricular failure or sudden death. Increased right atrial pressure above 15 mmHg and cardiac index below 2 L/min/m^2 are haemodynamic predictors of poor prognosis. Heart and lung transplantation is used (p. 741).

Chronic cor pulmonale

Cor pulmonale is enlargement of the right ventricle because of increase in afterload that is due to diseases of the thorax, lung and pulmonary circulation; the presence of right ventricular failure is not necessary for the diagnosis of cor pulmonale.

Pathophysiology

The precise mechanism varies according to the cause of cor pulmonale, but chronic obstructive pulmonary disease (COPD), which is discussed here, is illustrative. Pulmonary vascular resistance is increased because of loss of pulmonary vascular tissue and because of pulmonary vasoconstriction caused by hypoxia and acidosis. The increased pulmonary vascular resistance leads to pulmonary hypertension, which initially occurs only during an acute respiratory infection. Eventually, the pulmonary hypertension becomes persistent and progressively more severe. The pulmonary vascular bed is gradually obliterated by muscular hypertrophy of the arterioles and thrombus formation. Right ventricular function is progressively compromised because of the increased pressure load. Hypoxia further impairs right ventricular function and, as it develops, left ventricular function is also depressed.

Clinical features

Chest pain, exertional dyspnoea, syncope and *fatigue* are common symptoms, and sudden death occurs. Other symptoms are due to the cause of the pulmonary hypertension.

On *physical examination* (see Table 13.43), there is a prominent '*a*' wave in the jugular venous pulse, a right ventricular (parasternal) heave and a loud pulmonary component to the second heart sound. Other findings include a right ventricular fourth heart sound, a systolic pulmonary ejection click, a mid-systolic ejection murmur and an early diastolic murmur due to pulmonary regurgitation (Graham Steell murmur). If tricuspid regurgitation develops, there is a pansystolic murmur and a large jugular '*cv*' venous wave (see p. 691).

Investigations

- **Chest X-ray** may show right ventricular enlargement and right atrial dilatation. The pulmonary artery is usually prominent and the enlarged proximal pulmonary arteries taper rapidly. Peripheral lung fields are oligaemic.
- **ECG** demonstrates right ventricular hypertrophy (right axis deviation, possibly a dominant R wave in lead V_1, and inverted T waves in right precordial leads) and a

right atrial abnormality (tall peaked P waves in lead II) (Fig. 13.100).

- **Echocardiography** will usually demonstrate right ventricular dilatation and/or hypertrophy. It is often possible to measure the peak pulmonary artery pressure indirectly with Doppler echocardiography. The echocardiogram may also reveal the cause of pulmonary hypertension, such as an intracardiac shunt.

Other investigations may also be required to evaluate the cause of pulmonary hypertension and to look for treatable conditions, such as left-to-right shunts, mitral stenosis or left atrial tumours. With direct measurement of pulmonary artery pressure and pulmonary wedge pressure, *cardiac catheterization* is necessary in some patients with severe pulmonary hypertension of unknown cause. *Pulmonary angiography* may be indicated if multiple pulmonary emboli are suspected, but it is dangerous.

If no other cause is found, then a diagnosis of primary pulmonary hypertension is made (see p. 782).

Treatment

Treatment is determined by the condition underlying pulmonary hypertension. Diuretic treatment is used for right ventricular failure, but care should be taken to avoid excessive fluid depletion as this will result in reduced output from the impaired right ventricle. Hypoxia is avoided by the use of oxygen therapy when safe and necessary. In those with COPD and some others, long-term oxygen therapy improves symptoms and prognosis (p. 839). In contrast to their enormous value in those with left ventricular impairment, angiotensin-converting enzyme inhibitors are seldom useful and may make matters worse.

Pulmonary embolism

Thrombus, usually formed in the systemic veins or rarely in the right heart (<10% of cases), may dislodge and embolize into the pulmonary arterial system. Post-mortem studies indicate that this is a very common condition (microemboli are found in up to 60% of autopsies) but it is not usually diagnosed this frequently in life. Ten per cent of clinical pulmonary emboli are fatal.

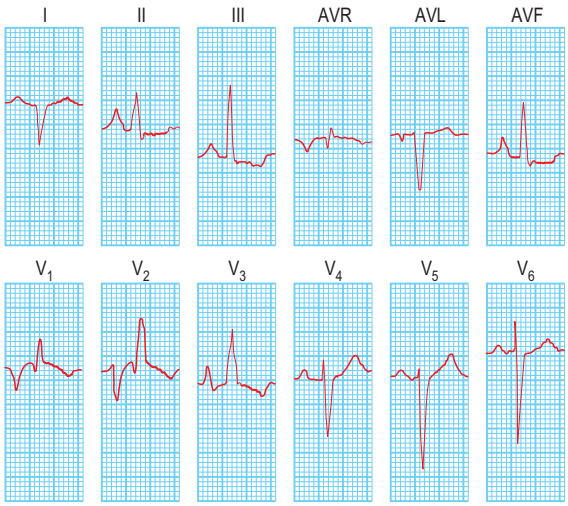


Fig. 13.100 **Pulmonary hypertension shown by a 12-lead ECG.** There is right axis deviation (+ 120°), right ventricular hypertrophy (dominant secondary R wave [R′] in V_1) and a combination of left and right atrial conduction abnormalities.

Most clots which cause clinically relevant pulmonary emboli actually come from the pelvic and abdominal veins, but femoral deep venous thrombosis, and even occasionally axillary thrombosis, can be the origin of the clot. Clot forms as a result of a combination of sluggish blood flow, local injury or compression of the vein and a hypercoagulable state. Emboli can also occur from tumour, fat (long bone fractures), amniotic fluid and foreign material during i.v. drug abuse. Risk factors are shown in Table 8.27 and discussed on page 442.

After pulmonary embolism, lung tissue is ventilated but not perfused – producing an intrapulmonary dead space and resulting in impaired gas exchange. After some hours the non-perfused lung no longer produces surfactant. Alveolar collapse occurs and exacerbates hypoxaemia. The primary haemodynamic consequence of pulmonary embolism is a reduction in the cross-sectional area of the pulmonary arterial bed which results in an elevation of pulmonary arterial pressure and a reduction in cardiac output. The zone of lung that is no longer perfused by the pulmonary artery may infarct, but often does not do so because oxygen continues to be supplied by the bronchial circulation and the airways.

Clinical features

Sudden onset of *unexplained dyspnoea* is the most common, and often the only symptom of pulmonary embolism. *Pleuritic chest pain* and *haemoptysis* are present only when infarction has occurred. Many pulmonary emboli occur silently, but there are three typical clinical presentations. A clinical deep venous thrombosis is not commonly observed, although detailed investigation of the lower limb and pelvic veins will reveal thrombosis in more than half of the cases.

Small/medium pulmonary embolism

In this situation an embolus has impacted in a terminal pulmonary vessel. Symptoms are pleuritic chest pain and breathlessness. Haemoptysis occurs in 30%, often 3 or more days after the initial event. On examination, the patient may be tachypnoeic with a localized pleural rub and often coarse crackles over the area involved. An exudative pleural effusion (occasionally blood-stained) can develop. The patient may have a fever, and cardiovascular examination is normal.

Massive pulmonary embolism

This is a much rarer condition where sudden collapse occurs because of an acute obstruction of the right ventricular outflow tract. The patient has severe central chest pain (cardiac ischaemia due to lack of coronary blood flow) and becomes shocked, pale and sweaty. Syncope may result if the cardiac output is transiently but dramatically reduced, and death may occur. On examination, the patient is tachypnoeic, has a tachycardia with hypotension and peripheral shutdown. The jugular venous pressure (JVP) is raised with a prominent '*a*' wave. There is a right ventricular heave, a gallop rhythm and a widely split second heart sound. There are usually no abnormal chest signs.

Multiple recurrent pulmonary emboli

This leads to increased breathlessness, often over weeks or months. It is accompanied by weakness, syncope on exertion and occasionally angina. The physical signs are due to the pulmonary hypertension that has developed from multi-

ple occlusions of the pulmonary vasculature. On examination, there are signs of right ventricular overload with a right ventricular heave and loud pulmonary second sound.

Diagnosis

The symptoms and signs of small and medium-sized pulmonary emboli are often subtle and non-specific, so the diagnosis is often delayed or even completely missed. Pulmonary embolism should be considered if patients present with symptoms of unexplained cough, chest pain, haemoptysis, new-onset atrial fibrillation (or other tachycardia), or signs of pulmonary hypertension if no other cause can be found.

The Revised Geneva Score (Table 13.44) gives the clinical probability of a pulmonary embolus.

Investigations

Small/medium pulmonary emboli

- **Chest X-ray** is often normal, but linear atelectasis or blunting of a costophrenic angle (due to a small effusion) is not uncommon. These features develop only after some time. A raised hemidiaphragm is present in some patients. More rarely, a wedge-shaped pulmonary infarct, the abrupt cut-off of a pulmonary artery or a translucency of an underperfused distal zone is seen. Previous infarcts may be seen as opaque linear scars.
- **ECG** is usually normal, except for sinus tachycardia, but sometimes atrial fibrillation or another tachyarrhythmia occurs. There may be evidence of right ventricular strain.
- **Blood tests**. Pulmonary infarction results in a polymorphonuclear leucocytosis, an elevated ESR and increased lactate dehydrogenase levels in the serum. Immediately prior to commencing anticoagulants a thrombophilia screen should be checked.

Table 13.44 **Revised Geneva score for the clinical prediction of a pulmonary embolism.** After Righini M, Le Gal G, Aujesky D, et al. *Lancet* 2008; **371**: 1343–1352, with permission from Elsevier

	Score
Risk factors	
Age >65 years	+1
Previous deep venous thrombosis or pulmonary embolism	+3
Surgery or fracture within 1 month	+2
Active malignancy	+2
Symptoms	
Unilateral leg pain	+3
Haemoptysis	+2
Clinical signs	
Heart rate (b.p.m.):	
75–94	+3
≥95	+5
Pain on leg deep vein palpation and unilateral oedema	+4
Clinical probability	
Low	0–3
Intermediate	4–10
High	≥11

- **Plasma D-dimer** (see p. 441) – if this is undetectable, it excludes a diagnosis of pulmonary embolism.
- **Radionuclide ventilation/perfusion scanning ($\dot{V}/\dot{Q}$ scan)** is a good and widely available diagnostic investigation. Pulmonary ^{99m}Tc scintigraphy demonstrates underperfused areas (Fig. 13.101) which, if not accompanied by a ventilation defect on a ventilation scintigram performed after inhalation of radioactive xenon gas (see p. 824), is highly suggestive of a pulmonary embolus. There are limitations to the test, however. For example, a matched defect may arise with a pulmonary embolus which causes an infarct, or from emphysematous bullae. This test is therefore conventionally reported as a probability (low, medium or high) of pulmonary embolus and should be interpreted in the context of the history, examination and other investigations.
- **Ultrasound scanning** can be performed for the detection of clots in pelvic or iliofemoral veins (see p. 810).
- **CT scans**. Contrast-enhanced multidetector CT angiograms (CTA) (Fig. 13.102), have a sensitivity of 83% and specificity of 96%, with a positive predictive value of 92%. These values will increase with the use of 64-multislice scanners.
- **MR imaging** gives similar results and is used if CT angiography is contraindicated.

Massive pulmonary emboli

- **Chest X-ray** may show pulmonary oligaemia, sometimes with dilatation of the pulmonary artery in the hila. Often there are no changes.
- **ECG** shows right atrial dilatation with tall peaked P waves in lead II. Right ventricular strain and dilatation give rise to right axis deviation, some degree of right bundle branch block, and T wave inversion in the right precordial leads (Fig. 13.103). The 'classic' ECG pattern with an S wave in lead I, and a Q wave and inverted T waves in lead III (S1, Q3, T3), is rare.

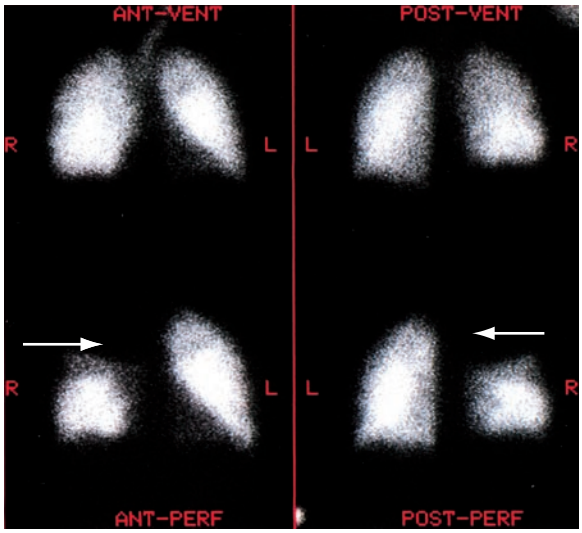


Fig. 13.101 **Ventilation (top) and perfusion (bottom) lung scans** which demonstrate absence of perfusion in the right upper lobe, i.e. probably pulmonary embolism (arrows).

- **Blood gases** show arterial hypoxaemia with a low arterial CO_2 level, i.e. type I respiratory failure pattern.
- **Echocardiography** shows a vigorously contracting left ventricle, and occasionally a dilated right ventricle and a clot in the right ventricular outflow tract.
- **Pulmonary angiography** has now been replaced by CT and MR angiography.

Multiple recurrent pulmonary emboli

- **Chest X-ray** may be normal. Enlarged pulmonary arterioles with oligaemic lung fields indicate advanced disease.
- **ECG** can be normal or show signs of pulmonary hypertension (Fig. 13.100).
- **Leg imaging** with ultrasound and venography may show thrombi.
- $\dot{V}/\dot{Q}$ scan may show evidence of pulmonary infarcts.
- **Multidetector CT scans** can detect small emboli.

Further tests looking for exercise-induced hypoxaemia and catheter studies to estimate pulmonary artery pressures are sometimes required.

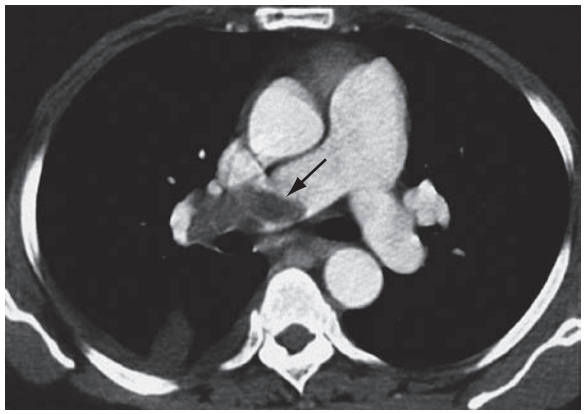


Fig. 13.102 **CT pulmonary angiographic image** at level of the main right and left pulmonary arteries showing a large thrombus (arrow) in the right pulmonary artery.

FURTHER READING

Righini M, Le Gal G, Aujesky G, et al. Diagnosis of pulmonary embolism by multidetector CT alone or combined with venous ultrasonography of the leg: a randomized non-inferiority trial. *Lancet* 2008; **371**: 1343–1352.

Tapson VF. Acute pulmonary embolism. *New England Journal of Medicine* 2008; **358**: 1037–1052.

Treatment

Acute management

- All patients should receive high-flow oxygen (60–100%) unless they have significant chronic lung disease. Patients with pulmonary infarcts require bed rest and analgesia.
- In severe cases, intravenous fluids and even inotropic agents to improve the pumping of the right heart are sometimes required, and very ill patients will require care on the intensive therapy unit (see p. 904).
- *Fibrinolytic therapy* such as streptokinase (250 000 units by i.v. infusion over 30 minutes, followed by streptokinase 100 000 units i.v. hourly for up to 12–72 hours according to manufacturer's instructions) has been shown in controlled trials to clear pulmonary emboli more rapidly and to confer a survival benefit in massive PE. It should be used more often.
- *Surgical embolectomy* is rarely necessary, but there may be no alternative when the haemodynamic circumstances are very severe.

Prevention of further emboli

A comparison of low-molecular-weight heparin (LMWH) with unfractionated heparin has shown no difference. As LMWHs simplify treatment (see p. 444) they are used exclusively if available, although they are more expensive. Oral anticoagulants are usually begun immediately and the heparin is tapered off as the oral anticoagulant becomes effective. Oral anticoagulants are continued for 6 weeks to 6 months, depending on the likelihood of recurrence of venous thrombosis or embolism. In some situations, such as after recurrent embolism, lifelong treatment is indicated.

Occasionally, physical methods are required to prevent further emboli. This is usually because recurrent emboli occur despite adequate anticoagulation, but it is also indicated in high-risk patients in whom anticoagulation is absolutely contraindicated. The most common method by which pulmonary embolism is treated in this situation is by insertion of a filter in the inferior vena cava via the femoral vein to above the level of the renal veins.

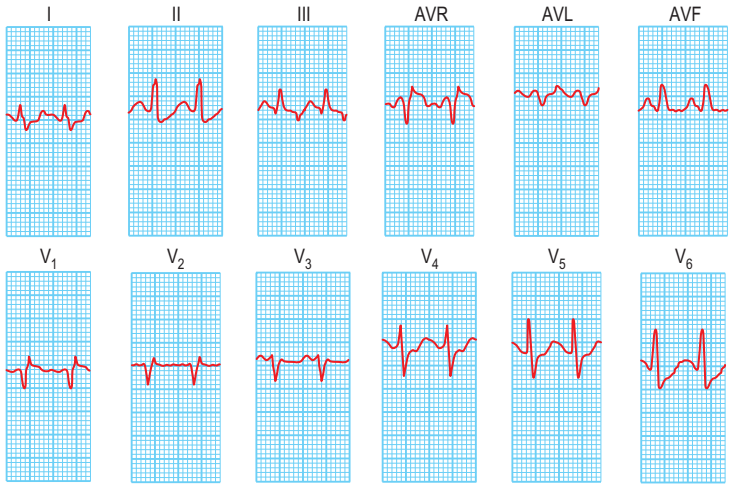


Fig. 13.103 **Acute pulmonary embolism shown by a 12-lead ECG.** There is an S wave in lead I, a Q wave in lead III and an inverted T wave in lead III (the S1, Q3, T3 pattern). There is sinus tachycardia (160 b.p.m.) and an incomplete right bundle branch block pattern (an R wave in AVR and V_1 and an S wave in V_6).

MYOCARDIAL AND ENDOCARDIAL DISEASE

ATRIAL MYXOMA

This is the most common primary cardiac tumour. It occurs at all ages and shows no sex preference. Although most myxomas are sporadic, some are familial or are part of a multiple system syndrome. Histologically they are benign. The majority of myxomas are solitary, usually develop in the left atrium and are polypoid, gelatinous structures attached by a pedicle to the atrial septum. The tumour may obstruct the mitral valve or may be a site of thrombi that then embolize. It is also associated with constitutional symptoms: the patient may present with dyspnoea, syncope or a mild fever. The physical signs are a loud first heart sound, a tumour 'plop' (a loud third heart sound produced as the pedunculated tumour comes to an abrupt halt), a mid-diastolic murmur and signs due to embolization. A raised ESR is usually present.

The diagnosis is easily made by echocardiography because the tumour is demonstrated as a dense space-occupying lesion (Fig. 13.104). Surgical removal usually results in a complete cure.

Myxomas may also occur in the right atrium or in the ventricles. Other primary cardiac tumours include rhabdomyomas and sarcomas.

MYOCARDIAL DISEASE

Myocardial disease that is not due to ischaemic, valvular or hypertensive heart disease, or a known infiltrative, metabolic/toxic or neuromuscular disorder may be caused by:

- an acute or chronic inflammatory pathology (myocarditis)
- idiopathic myocardial disease (cardiomyopathy).

Myocarditis

Acute inflammation of the myocardium has many causes (Table 13.45). Establishment of a definitive aetiology with isolation of viruses or bacteria is difficult in routine clinical practice.

In western societies, the commonest causes of infective myocarditis are Coxsackie or adenoviral infection. Myocarditis in association with HIV infection is seen at post-mortem in up to 20% of cases but causes clinical problems in less than 10% of cases. Chagas' disease, due to *Trypanosoma cruzi*, which is endemic in South America, is one of the commonest causes of myocarditis world-wide. Additionally toxins (including prescribed drugs), physical agents, hypersensitivity reactions and autoimmune conditions may also cause myocardial inflammation.

Pathology

In the acute phase myocarditic hearts are flabby with focal haemorrhages; in chronic cases they are enlarged and hypertrophied. Histologically an inflammatory infiltrate is present – lymphocytes predominating in viral causes; polymorphonuclear cells in bacterial causes; eosinophils in allergic and hypersensitivity causes (Fig. 13.105).

Clinical features

Myocarditis may be an acute or chronic process; its clinical presentations range from an asymptomatic state associated with limited and focal inflammation to *fatigue*, *palpitations*, *chest pain*, *dyspnoea* and *fulminant congestive cardiac*

Table 13.45 Causes of myocarditis

Idiopathic

Infective

- **Viral:** Coxsackievirus, adenovirus, CMV, echovirus, influenza, polio, hepatitis, HIV
- **Parasitic:** *Trypanosoma cruzi, Toxoplasma gondii* (a cause of myocarditis in the newborn or immunocompromised)
- **Bacterial:** Streptococcus (most commonly rheumatic carditis), diphtheria (toxin-mediated heart block common)
- **Spirochaetal:** Lyme disease (heart block common), leptospirosis
- **Fungal**
- **Rickettsial**

Toxic

- **Drugs:** Causing hypersensitivity reactions, e.g. methyldopa, penicillin, sulphonamides, antituberculous, modafinil
- **Radiation:** May cause myocarditis but pericarditis more common

Autoimmune

An autoimmune form with autoactivated T cells and organ-specific antibodies may occur

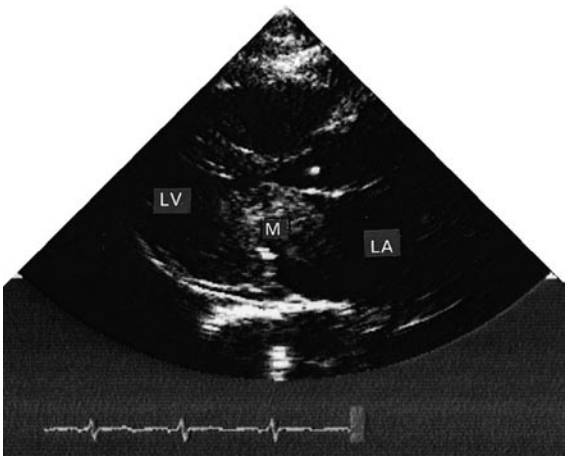
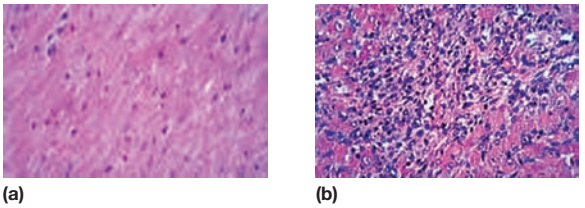


Fig. 13.104 **Atrial myxoma shown by a 2-D echocardiogram** (long-axis view). The myxoma is an echo-dense mass obstructing the mitral valve orifice. It was removed surgically. LV, left ventricle; LA, left atrium; M, mass.



(a) (b)

Fig. 13.105 **Biopsies showing (a) normal myocardium and (b) myocarditis,** with increased interstitial inflammatory cells.

failure due to diffuse myocardial involvement. An episode of viral myocarditis, perhaps unrecognized and forgotten, may be the initial event that eventually culminates in an 'idiopathic' dilated cardiomyopathy. Physical examination includes soft heart sounds, a prominent third sound and often a tachycardia. A pericardial friction rub may be heard.

Investigations

- **Chest X-ray** may show some cardiac enlargement, depending on the stage and virulence of the disease.
- **ECG** demonstrates ST- and T wave abnormalities and arrhythmias. Heart block may be seen with diphtheritic myocarditis, Lyme disease and Chagas' disease (see below).
- **Cardiac enzymes** are elevated.
- **Viral antibody titres** may be increased. However, since enteroviral infection is common in the general population, the diagnosis depends on the demonstration of acutely rising titres.
- **Endomyocardial biopsy** may show acute inflammation but false negatives are common by conventional criteria. Biopsy is of limited value outside specialized units.
- **Viral RNA** can be measured from biopsy material using polymerase chain reaction (PCR). Specific diagnosis requires demonstration of active viral replication within myocardial tissue.

Treatment

The underlying cause must be identified, treated, eliminated or avoided. Bed rest is recommended in the acute phase of the illness and athletic activities should be avoided for 6 months. Heart failure should be treated conventionally with the use of diuretics, ACE inhibitors/AII receptor antagonists, beta-blockers, spironolactone ± digoxin. *Antibiotics* should be administered immediately where appropriate. *NSAIDs* are contraindicated in the acute phase of the illness but may be used in the late phase. The use of *corticosteroids* is controversial and no studies have demonstrated an improvement in left ventricular ejection fraction or survival following their use. The administration of high-dose intravenous *immunoglobulin* on the other hand appears to be associated with a more rapid resolution of the left ventricular dysfunction and improved survival. Novel and effective antiviral, immunosuppressive (e.g. gamma-interferon) and *immunomodulating* (e.g. IL-10) agents may become available in the future to treat viral myocarditis.

Giant cell myocarditis

This is a severe form of myocarditis characterized by the presence of multinucleated giant cells within the myocardium. The cause is unknown but it may be associated with sarcoidosis, thymomas and autoimmune disease. It has a rapidly progressive course and a poor prognosis. Immunosuppression is recommended.

Chagas' disease

Chagas' disease is caused by the protozoan *Trypanosoma cruzi* and is endemic in South America where upwards of 20 million people are infected. Acutely, features of myocarditis are present with fever and congestive heart failure. Chronically, there is progression to a dilated cardiomyopathy with a propensity towards heart block and ventricular arrhythmias. Treatment is discussed on page 158. Amiodarone is helpful for the control of ventricular arrhythmias; heart failure is treated in the usual way (p. 737).

CARDIOMYOPATHY

Cardiomyopathies are a group of diseases of the myocardium that affect the mechanical or electrical function of the heart. They are frequently genetic and may produce inappropriate ventricular hypertrophy or dilatation and can be primarily a cardiac disorder or part of a multi-system disease (Table 13.46). Abnormal myocardial function produces systolic or diastolic heart failure; abnormal electrical conduction results in cardiac arrhythmias and sudden cardiac death.

Hypertrophic cardiomyopathy (HCM)

HCM includes a group of inherited conditions that produce hypertrophy of the myocardium in the absence of an alternate cause (e.g. aortic stenosis or hypertension). It is the most common cause of sudden cardiac death in young people and affects 1 in 500 of the population. The majority of cases are familial autosomal dominant, due to mutations in the genes encoding sarcomeric proteins (Fig. 13.106). The most common causes of HCM are mutations of the β-myosin heavy chain and myosin-binding protein C. Other mutations include troponin T and I, regulatory and essential myosin light chains, titin, α tropomyosin, α actin, α myosin heavy chain and muscle LIM protein (although over 400 mutations have been identified.) There are non-sarcomeric protein mutations in genes that control cardiac metabolism that result in glycogen storage diseases (Danon's, Pompe's and Fabry's disease) that are indistinguishable from HCM.

Clinical features

HCM is characterized by variable myocardial hypertrophy frequently involving the interventricular septum and

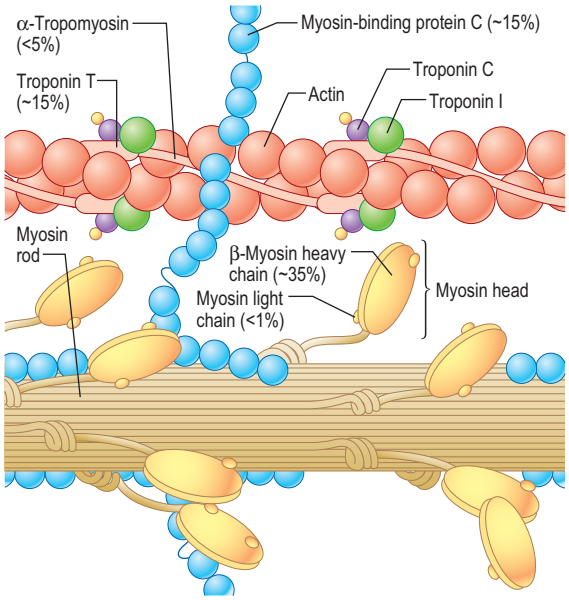


Fig. 13.106 **Sarcomeric proteins implicated in hypertrophic cardiomyopathy.** From Spirito P, Seidman CE, McKenna WJ, Maron BJ. The management of hypertrophic cardiomyopathy. *New England Journal of Medicine* 1997; **336**: 775–785. Copyright © Massachusetts Medical Society, all rights reserved.

Table 13.46 Cardiomyopathies

(a) Primary cardiomyopathies

Genetic	Mixed (genetic/acquired)	Acquired
Hypertrophic cardiomyopathy	Dilated cardiomyopathy	Inflammatory (myocarditis)
Arrhythmogenic right ventricular cardiomyopathy	Restrictive (non-hypertrophic and non-dilated)	Stress produced (Tako-tsubo)
Left ventricular non-compaction		Peripartum
		Tachycardia induced
		Infants of mothers with Type 1 diabetes
Conduction defects (see p. 726)		
Mitochondrial myopathies (see p. 37)		
*Ion channel disorders		
LQTS		
Brugada		
SQTS		
CPVT		
Asian SUNDS		

LQTS, long QT syndrome; SQTS, short QT syndrome; CPVT, catecholaminergic polymorphic ventricular tachycardia; Asian SUNDS, Asian sudden unexpected nocturnal death syndrome.
*Excluded from the European Society of Cardiology classification.

(b) Secondary cardiomyopathies

Infiltrative*	e.g. amyloidosis, Gaucher's disease†, Hurler's disease†, Hunter's disease†
Storage‡	e.g. hereditary haemochromatosis, Fabry's disease†, glycogen storage disease (type II, Pompe)†, Niemann–Pick disease†
Toxicity	e.g. drugs (e.g. cocaine), alcohol, heavy metals (e.g. cobalt), chemical agents
Endomyocardial	e.g. endomyocardial fibrosis, Loeffler's endocarditis
Inflammatory (granulomatous)	e.g. sarcoidosis, post-infective (e.g. Chagas' disease)
Endocrine	e.g. diabetes mellitus†, hyper- or hypo-thyroidism, hyperparathyroidism, phaeochromocytoma, acromegaly
Cardiofacial	e.g. Noonan syndrome†, lentiginosis†
Neuromuscular/neurological	e.g. Friedrich's ataxia†, Duchenne–Becker muscular dystrophy†, myotonic dystrophy†, neurofibromatosis†
Nutritional deficiencies	e.g. beriberi (thiamin), pellagra, scurvy, selenium, carnitine, kwashiorkor
Autoimmune/collagen	e.g. systemic lupus erythematosus, dermatomyositis, systemic sclerosis
Electrolyte imbalance	
Consequences of cancer therapy	e.g. anthracyclines (e.g. doxorubicin), daunorubicin, cyclophosphamide, radiation

*Accumulation of abnormal substances between myocytes (i.e. extracellular).
†Genetic (familial) origin.
‡Accumulation of abnormal substances within myocytes (i.e. intracellular).

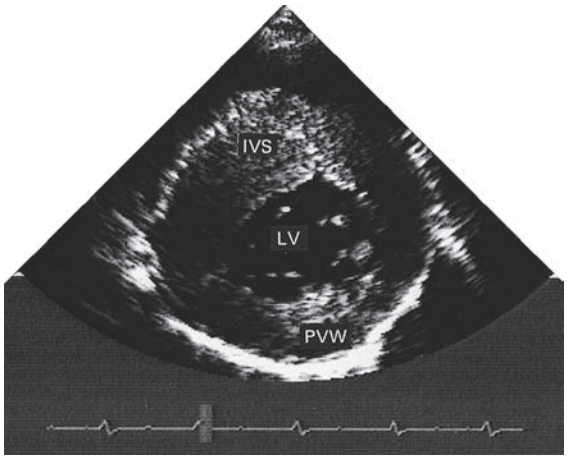
disorganization ('disarray') of cardiac myocytes and myofibrils. Twenty-five per cent of patients have dynamic left ventricular outflow tract obstruction due to the combined effects of hypertrophy, systolic anterior motion (SAM) of the anterior mitral valve leaflet and rapid ventricular ejection. The salient clinical and morphological features of the disease vary according to the underlying genetic mutation. For example, marked hypertrophy is common with β myosin heavy chain mutations whereas mutations in troponin T may be associated with mild hypertrophy but a high risk of sudden death. The hypertrophy may not manifest before completion of the adolescent growth spurt, making the diagnosis in children difficult. HCM due to myosin-binding protein may not manifest until the sixth decade of life or later.

Symptoms

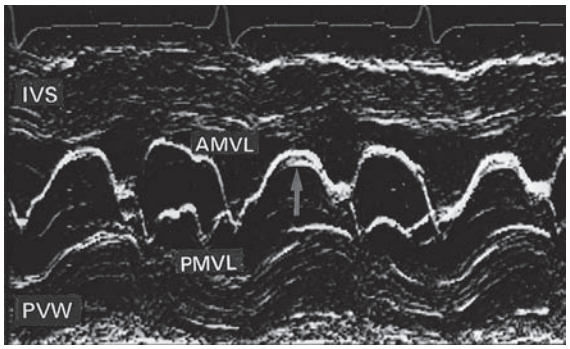
- many are asymptomatic and are detected through family screening of an affected individual or following a routine ECG examination
- chest pain, dyspnoea, syncope or pre-syncope (typically with exertion), cardiac arrhythmias and sudden death are seen
- sudden death occurs at any age but the highest rates (up to 6% per annum) occur in adolescents or young adults. Risk factors for sudden death are discussed below
- dyspnoea occurs due to impaired relaxation of the heart muscle or the left ventricular outflow tract obstruction that occurs in some patients. The systolic cavity remains small until the late stages of disease when progressive dilatation may occur. If a patient develops atrial fibrillation there is often a rapid deterioration in clinical status due to the loss of atrial contraction and the tachycardia – resulting in elevated left atrial pressure and acute pulmonary oedema.

Signs

- double apical pulsation (forceful atrial contraction producing a fourth heart sound)



(a)



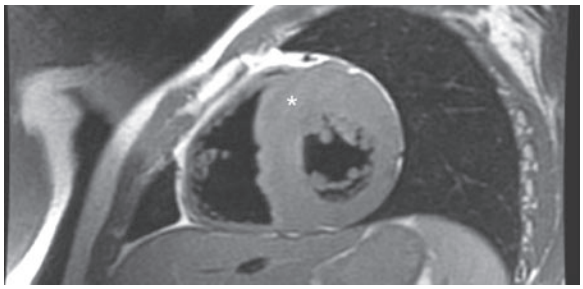
(b)

Fig. 13.107 **Hypertrophic cardiomyopathy. (a)** A 2-D echocardiogram (short-axis view). **(b)** M-mode recording. The grossly thickened interventricular septum is shown, resulting in a small left ventricular cavity. This condition is associated with an abnormal anterior motion of the mitral valve during systole (arrow). IVS, interventricular septum; LV, left ventricle; PVW, posterior ventricular wall; AMVL, PMVL, anterior and posterior mitral valve leaflets.

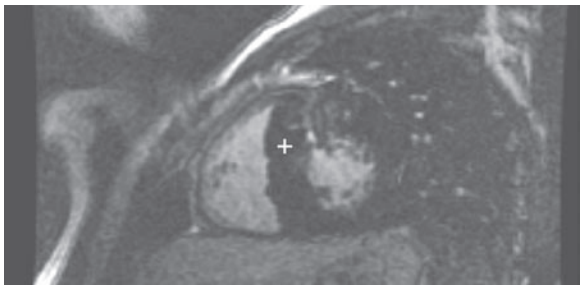
- jerky carotid pulse because of rapid ejection and sudden obstruction to left ventricular outflow during systole
- ejection systolic murmur due to left ventricular outflow obstruction late in systole – it can be increased by manoeuvres that decrease after-load, e.g. standing or Valsalva, and decreased by manoeuvres that increase after-load and venous return, e.g. squatting
- pan-systolic murmur due to mitral regurgitation (secondary to SAM)
- fourth heart sound (if not in AF).

Investigations

- **ECG** abnormalities of HCM include left ventricular hypertrophy (see Fig. 13.80), ST and T wave changes, and abnormal Q waves especially in the infero-lateral leads.
- **Echocardiography** is usually diagnostic and in classical HCM there is asymmetric left ventricular hypertrophy (involving the septum more than the posterior wall), systolic anterior motion of the mitral valve, and a vigorously contracting ventricle (Fig. 13.107). However, any pattern of hypertrophy may be seen, including concentric and apical hypertrophy.
- **Cardiac MR** can detect both the hypertrophy but also abnormal myocardial fibrosis (Fig. 13.108).



(a)



(b)

Fig. 13.108 **Cardiac MR in a patient with hypertrophic cardiomyopathy. (a)** Turbo spin-echo demonstrates marked thickening of the basal to mid-anterior wall and septum (*). **(b)** Inversion recovery image post gadolinium demonstrates regional fibrosis with enhancement (+).

- **Genetic analysis**, where available, may confirm the diagnosis and provide prognostic information for the patient and relatives.

Treatment

The management of HCM includes treatment of symptoms and the prevention of sudden cardiac death in the patient and relatives.

Risk factors for sudden death:

- massive left ventricular hypertrophy (>30 mm on echocardiography)
- family history of sudden cardiac death (<50 years old)
- non-sustained ventricular tachycardia on 24-hour Holter monitoring
- prior unexplained syncope
- abnormal blood pressure response on exercise (flat or hypotensive response).

The presence of these cardiac risk factors is associated with an increased risk of sudden death, and patients with two or more should be assessed for implantable cardioverter–defibrillator (ICD). In patients in whom the risk is less, amiodarone is an appropriate alternative.

Chest pain and dyspnoea are treated with beta-blockers and verapamil either alone or in combination. An alternative agent is disopyramide if patients have left ventricular outflow tract obstruction. In some patients with significant left outflow obstruction and symptoms, dual-chamber pacing is necessary. Alcohol (non-surgical) ablation of the septum has been investigated and appears to give good results in reduction of outflow tract obstruction and subsequent improvement in exercise capacity. This procedure carries risks, including the development of complete heart block and myocardial infarction. Occasionally surgical resection of septal myocardium may be indicated. Vasodilators should be avoided because

they may aggravate left ventricular outflow obstruction or cause refractory hypotension.

Arrhythmogenic right ventricular cardiomyopathy (ARVC)

ARVC is an uncommon (1 in 5000 population) inherited condition that predominantly affects the right ventricle with fatty or fibro-fatty replacement of myocytes, leading to segmental or global dilatation (Fig. 13.109). Left ventricular involvement has been reported in up to 75% of cases. The fibro-fatty replacement leads to ventricular arrhythmia and risk of sudden death in its early stages, and right ventricular or biventricular failure in its later stages.

Autosomal dominant ARVC has been mapped to eight chromosomal loci within mutations in four genes – cardiac ryanodine receptor RyR2 (also responsible for familial catecholaminergic polymorphic ventricular tachycardia, CPVT), desmoplakin, plakophillin-2 and mutations altering the regulatory sequences of the transforming growth factor-β gene.

There are two recessive forms – *Naxos disease* (associated with palmoplantar keratoderma and woolly hair) that is due to a mutation in junctional plakoglobin, and *Carvajal syndrome*, due to a mutation in desmoplakin.

Clinical features

Most patients are asymptomatic. Symptomatic ventricular arrhythmias, syncope or sudden death occur. Occasionally presentation is with symptoms and signs of right heart failure, although this is more common in the later stages of the disease. Some patients may be detected through family screening, although frequently the morphological appearances of the right ventricle is normal despite significant cardiac arrhythmias.

Investigations

- **ECG** is usually normal but may demonstrate T wave inversion in the precordial leads related to the right ventricle (V_1–V_3). Small-amplitude potentials occurring at the end of the QRS complex (epsilon waves) may be present (Fig. 13.110) and incomplete or complete RBBB is sometimes seen. Signal averaged ECG may indicate the presence of late potentials, the delayed depolarization of individual muscle cells. Twenty-four-hour Holter monitoring may demonstrate frequent extrasystoles of right ventricular origin or runs of non-sustained ventricular tachycardia.
- **Echocardiography** is frequently normal but with more advanced cases may demonstrate right ventricular dilatation and aneurysm formation, and there may be left ventricular dilatation.
- **Cardiac MR** can more accurately assess the right ventricle and in some cases can demonstrate fibro-fatty infiltration (Fig. 13.111).
- **Genetic testing** may soon be the diagnostic gold standard.

Treatment

Beta-blockers are first-line treatment for patients with non-life-threatening arrhythmias. Amiodarone or sotalol are used for symptomatic arrhythmias but for refractory or life-threatening arrhythmias an ICD is required. Occasionally cardiac transplantation is indicated either for intractable arrhythmia or cardiac failure.

Left ventricular non-compaction (LVNC)

LVNC is a recently recognized condition associated with a sponge-like appearance of the left ventricle. The condition

Fig. 13.109 **Gross pathological specimen demonstrating thinning and fibrofatty replacement of right ventricular free wall.** From Basso C, Thiene G, Corrado D, et al. *Circulation* 1996; **94**: 983–991, with permission.

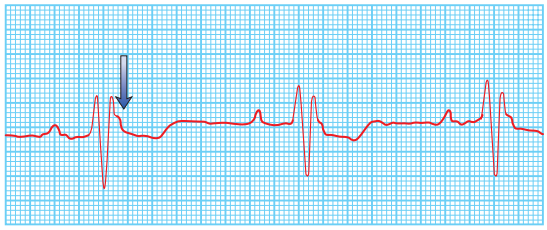


Fig. 13.110 **Electrocardiogram from an adult with arrhythmogenic right ventricular cardiomyopathy (ARVC)** demonstrating RBBB and precordial T wave insertion with epsilon waves visible at the terminal of the QRS complex (arrow).

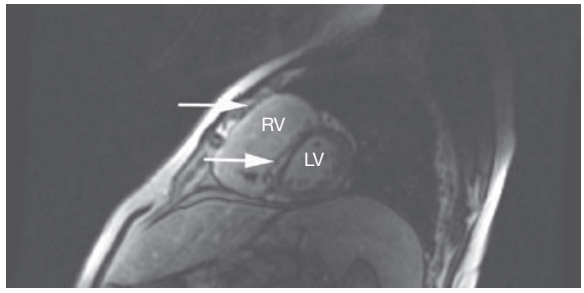


Fig. 13.111 **Arrhythmogenic right ventricular cardiomyopathy (ARVC).** Inversion recovery MR image post gadolinium demonstrates marked hyperenhancement (arrows) in the right ventricular free wall and infero-septum consistent with fibro-fatty replacement in right and left ventricles.

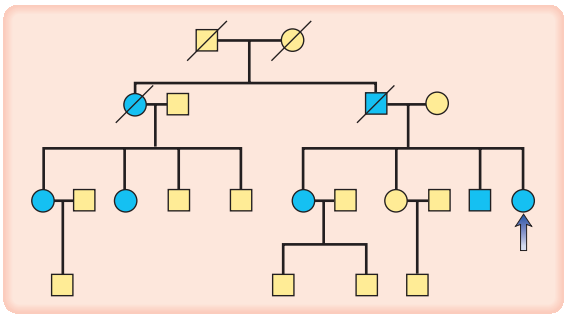


Fig. 13.112 **Pedigree of a family with dilated cardiomyopathy.** Blue symbols are affected family members. The arrow indicates the index case.

predominantly affects the apical portion of the left ventricle with persistent deep sinusoids in communication with the cavity due to arrested embryological differentiation. LVNC may be associated with congenital heart abnormalities. The condition is diagnosed by echocardiography, cardiac MR or left ventricular angiography. The natural history of the condition is unresolved but includes congestive cardiac failure, thromboembolism, cardiac arrhythmias and sudden death. Familial and spontaneous cases have been described.

Conduction system disease

Lenegre disease (p. 719) is a progressive disease of the cardiac conduction system (His–Purkinje system) that causes broad QRS duration, long pauses and bradycardia which presents with syncope. Sick-sinus syndrome is phenotypically similar.

Ion channelopathies

Features of the ion channelopathies are long QT syndrome (LQTS) (p. 727), Brugada syndrome (p. 727), catecholaminergic polymorphic ventricular tachycardia (CPVT), short QT syndrome (SQTS) and idiopathic ventricular fibrillation (VF).

Dilated cardiomyopathy (DCM)

DCM has a prevalence of 1 in 2500 and is characterized by dilatation of the ventricular chambers and systolic dysfunction with preserved wall thickness.

Familial DCM is predominantly autosomal dominant and can be associated with over 20 abnormal loci and genes (Fig. 13.112). Many of these are genes encoding cytoskeletal or associated myocyte proteins (dystrophin in X-linked cardiomyopathy; actin, desmin, troponin T, beta myosin heavy chain, sarcoglycans, vinculin and lamin a/c in autosomal dominant DCM) (Fig. 13.113). Many of these have prominent associated features such as skeletal myopathy or conduction system disease and therefore differ from the majority of cases of DCM.

Sporadic DCM can be caused by multiple conditions (Table 13.46):

- myocarditis – Coxsackie, adenoviruses, erythroviruses, HIV, bacteria, fungae, mycobacteria, parasitic (Chagas' disease)
- toxins – alcohol, chemotherapy, metals (cobalt, lead, mercury, arsenic)
- autoimmune

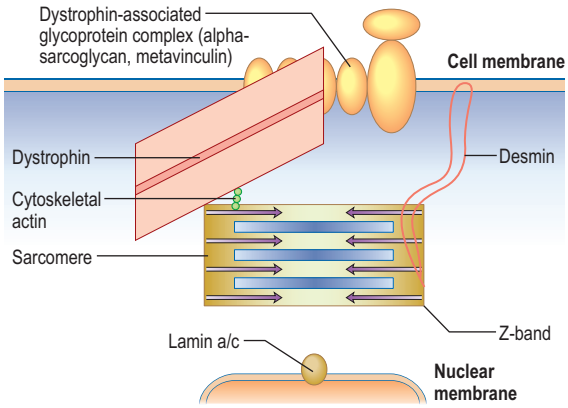


Fig. 13.113 **Schematic representation of myocyte proteins implicated in dilated cardiomyopathy (DCM).** See also Figure 13.4.

- endocrine
- neuromuscular.

Clinical features

DCM can present with heart failure, cardiac arrhythmias, conduction defects, thromboembolism or sudden death. Increasingly, evaluation of relatives of DCM patients is allowing identification of early asymptomatic disease, prior to the onset of these complications. Clinical evaluation should include a family history and construction of a pedigree where appropriate.

Investigations

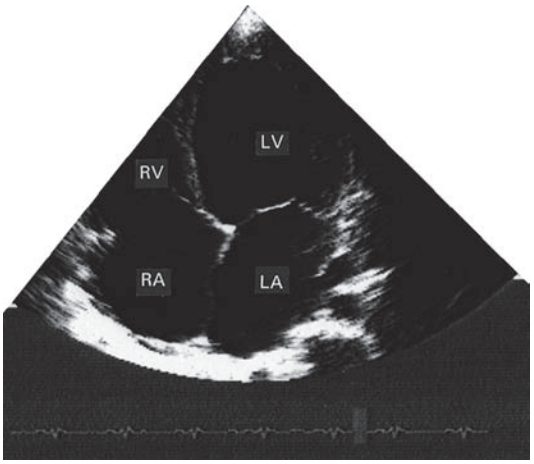
- **Chest X-ray** demonstrates generalized cardiac enlargement.
- **ECG** may demonstrate diffuse non-specific ST segment and T wave changes. Sinus tachycardia, conduction abnormalities and arrhythmias (i.e. atrial fibrillation, ventricular premature contractions or ventricular tachycardia) are also seen.
- **Echocardiogram** reveals dilatation of the left and/or right ventricle with poor global contraction function (Fig. 13.114).
- **Cardiac MR** may demonstrate other aetiologies of left ventricular dysfunction (e.g. previous myocardial infarction) or demonstrate abnormal myocardial fibrosis (Fig. 13.115). Cardiac MR is also useful for identifying myocardial thrombus (Fig. 13.116).
- **Coronary angiography** should be performed to exclude coronary artery disease in all individuals at risk (generally patients >40 years or younger if symptoms or risk factors are present).
- **Biopsy** is generally not indicated outside specialist care.

Treatment

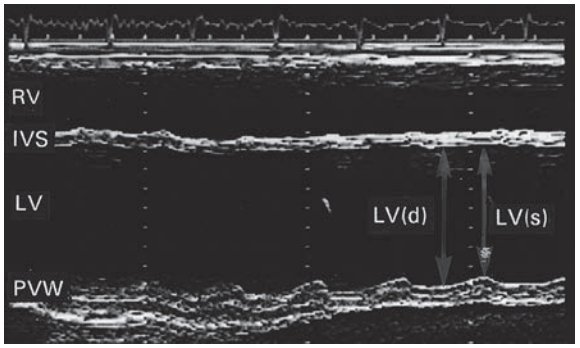
Treatment consists of the conventional management of heart failure (see p. 737) with the option of cardiac resynchronization therapy and ICDs in patients with NYHA III/IV grading. Cardiac transplantation is appropriate for certain patients.

Primary restrictive non-hypertrophic cardiomyopathy

This is a rare condition in which there is normal or decreased volume of both ventricles with bi-atrial enlargement, normal



(a)



(b)

Fig. 13.114 **Dilated cardiomyopathy. (a)** 2-D (apical four-chamber view) and **(b)** M-mode echocardiograms. The heart has a 'globular' appearance with all four chambers dilated. The extremely impaired left ventricular function can be appreciated from the M-mode recording. Compare the systolic shortening fraction with that of Figures 13.24 and 13.28. LA, left atrium; RA, right atrium; LV, left ventricle; LV(d) and LV(s), diastolic and systolic left ventricular dimensions; IVS, interventricular septum; PVW, posterior ventricular wall.

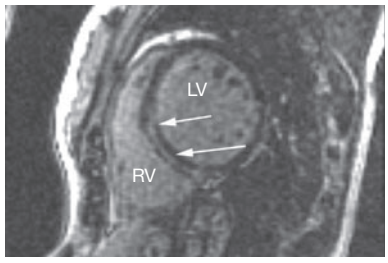


Fig. 13.115 **Dilated cardiomyopathy.** Inversion recovery MR post gadolinium demonstrates septal mid-wall hyperenhancement (arrows) consistent with fibrosis in a patient with dilated cardiomyopathy.

wall thickness, normal cardiac valves and impaired ventricular filling with restrictive physiology but near normal systolic function. The restrictive physiology produces symptoms and signs of heart failure. Conditions associated with this form of cardiomyopathy include amyloidosis (commonest), sarcoidosis, Loeffler's endocarditis and endomyocardial fibrosis; in the latter two conditions there is myocardial and endocardial fibrosis associated with eosinophilia. The idiopathic form of restrictive cardiomyopathy may be familial.

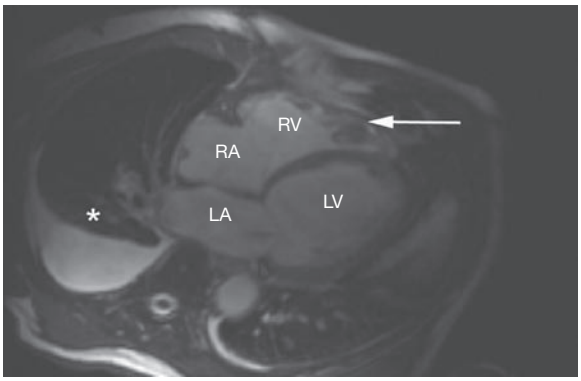


Fig. 13.116 **Dilated cardiomyopathy.** MR showing thrombus in myocardium. Horizontal long-axis CMR demonstrates biventricular dilatation with reduced systolic function and a mobile thrombus (arrow) in the right ventricular apex in a patient with dilated cardiomyopathy. There is also a right pleural effusion (*).

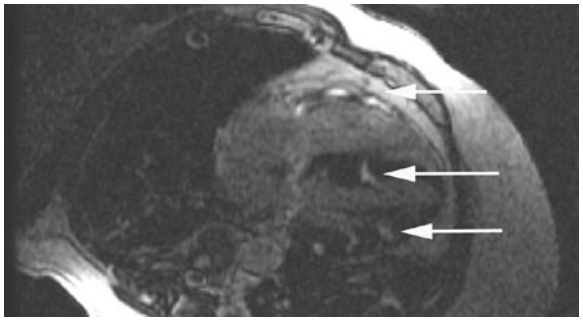


Fig. 13.117 **Sarcoidosis.** Inversion recovery MR post gadolinium demonstrates marked hyperenhancement (arrows) in the right ventricular free wall, septum and lateral wall consistent with myocardial sarcoidosis.

Clinical features

Patients with restrictive cardiomyopathy may present with dyspnoea, fatigue and embolic symptoms. On clinical examination there will be elevated jugular venous pressure with diastolic collapse (Friedreich's sign) and elevation of venous pressure with inspiration (Kussmaul's sign), hepatic enlargement, ascites and dependent oedema. Third and fourth heart sounds may be present.

Investigations

- **Chest X-ray** may show pulmonary venous congestion. The cardiac silhouette can be normal or show cardiomegaly and/or atrial enlargement.
- **ECG** may demonstrate low-voltage QRS and ST segment and T wave abnormalities.
- **Echocardiography** shows symmetrical myocardial thickening and often a normal systolic ejection fraction but impaired ventricular filling.
- **Cardiac MR** may demonstrate abnormal myocardial fibrosis in amyloidosis or sarcoidosis (Fig. 13.117).
- **Cardiac catheterization** and haemodynamic studies may help distinguish between restrictive cardiomyopathy and constrictive pericarditis, although volume loading may be required.
- **Endomyocardial biopsy** in contrast with other cardiomyopathies is often useful in this condition and

may permit a specific diagnosis, such as amyloidosis, to be made.

Treatment

There is no specific treatment. Cardiac failure and embolic manifestations should be treated. Cardiac transplantation is necessary in some severe cases, especially the idiopathic variety. In primary amyloidosis combination therapy with melphalan plus prednisolone with or without colchicine may improve survival. However, patients with cardiac amyloidosis have a worse prognosis than those with other forms of the disease, and the disease often recurs after transplantation. Liver transplantation may be effective in familial amyloidosis (due to production of mutant pre-albumin) and may lead to reversal of the cardiac abnormalities.

Acquired cardiomyopathies

Myocarditis (see p. 787)

Stress (Tako-tsubo/octopus pot) cardiomyopathy

This is a recently described acute and reversible cardiomyopathy that occurs in the absence of coronary artery disease and is usually triggered by profound psychological stress. It is more common in middle–old aged women and classically affects the left ventricular apex causing apical ballooning (Fig. 13.118).

Peri-partum cardiomyopathy

This rare condition affects women in the last trimester of pregnancy or within 5 months of delivery. It presents as a dilated cardiomyopathy, is more common in obese, multiparous women over 30 years old and is associated with pre-eclampsia. Nearly half of patients will recover to normal function within 6 months but in some patients it can causes progressive heart failure and sudden death.

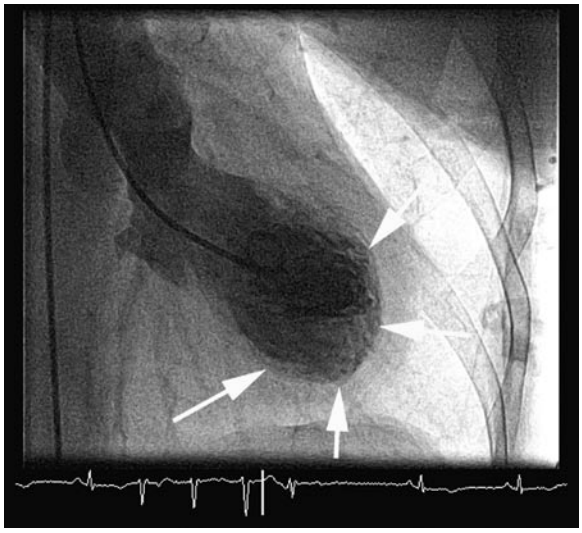


Fig. 13.118 **Tako-tsubo cardiomyopathy.** Left ventricular angiography in a patient admitted with chest pain following emotional stress demonstrates apical ballooning (arrows) consistent with Tako-tsubo cardiomyopathy. Diagnostic coronary angiography showed normal coronary arteries.

Tachycardia cardiomyopathy

Prolonged periods of supraventricular or ventricular tachycardia will lead to dilated cardiomyopathy. Cardioversion and ablation may be necessary to restore sinus rhythm and allow for recovery of cardiac function.

FURTHER READING

Lewis GD, Shah R, Shahzad K, et al. Sildenafil improves exercise capacity and quality of life in patients with systolic heart failure and secondary pulmonary hypertension. *Circulation* 2007; **116**: 1555–1562.

Maron BJ, Tobin JA, Thiene G, et al. Definition and classification of cardiomyopathies. *Circulation* 2006; **113**: 1807–1816.

PERICARDIAL DISEASE

The pericardium acts as a protective covering for the heart. It consists of two separate layers, the inner visceral pericardium and the outer parietal pericardium. The visceral pericardium reflects back upon itself at the level of the great vessels to join the parietal pericardium, thus forming a sac. The pericardial sac contains up to 50 mL of pericardial fluid in the normal heart, although this is a potential space for fluid to collect. The pericardium serves to lubricate the surface of the heart, prevents deformation and dislocation of the heart and acts as a barrier to the spread of infection.

Presentations of pericardial disease include:

- acute pericarditis
- pericardial effusion and cardiac tamponade
- constrictive pericarditis.

Acute pericarditis

This refers to inflammation of the pericardium. Classically, fibrinous material is deposited into the pericardial space and pericardial effusion often occurs. Acute pericarditis has numerous aetiologies (Table 13.47). Most commonly in the UK, it is due to viral infection and myocardial infarction, although in many cases the cause is unknown.

Viral pericarditis. The most common viral causes are Coxsackie B virus and echovirus. Viral pericarditis is usually painful but has a short time course and rarely long-term effects. Increasingly, HIV is implicated in the aetiology of pericarditis, both directly and via immunosuppression, which predisposes the subject to infective causes.

Post-myocardial infarction pericarditis occurs in about 20% of patients in the first few days following MI. It occurs more commonly with anterior MI and ST elevation MI with high serum cardiac enzymes, but its incidence is reduced to 5–6% with thrombolysis. It may be difficult to differentiate this pain from recurrent angina when it occurs early (day 1–2 post-infarct) but a good history of the pain and serial ECG monitoring is helpful. Pericarditis may also occur later on in the recovery phase after infarction. This usually occurs as a feature of *Dressler's syndrome*, an autoimmune response to cardiac damage occurring 2–10 weeks post-infarct. Autoimmune reaction to myocardial damage is the main aetiology, and antimyocardial antibodies can often be found. Recurrences are common. Differential diagnosis includes a new myocardial infarction or unstable angina.

Uraemic pericarditis is due to irritation of the pericardium by accumulating toxins. It can occur in 6–10% of patients with advanced renal failure if dialysis is delayed. It is an indication for urgent dialysis as it continues to be associated with significant morbidity and mortality.

Table 13.47 Aetiology of pericarditis

I. Infectious pericarditis
- Viral (Coxsackievirus, echovirus, mumps, herpes, HIV)
- Bacterial (*Staphylococcus*, *Streptococcus*, *Pneumococcus*, *Meningococcus*, *Haemophilus influenzae*, *Mycoplasmosis*, *Borreliosis*, *Chlamydia*)
- Tuberculous
- Fungal (*Histoplasmosis*, *Coccidioidomycosis*, *Candida*)

II. Post-myocardial infarction pericarditis
- Acute myocardial infarction (early)
- Dressler's syndrome (late)

III. Malignant pericarditis
- Primary tumours of the heart (mesothelioma)
- Metastatic pericarditis (breast and lung carcinoma, lymphoma, leukaemia, melanoma)

IV. Uraemic pericarditis

V. Myxoedematous pericarditis

VI. Chylopericardium

VII. Autoimmune pericarditis
- Collagen–vascular (rheumatoid arthritis, rheumatic fever, systemic lupus erythematosus, scleroderma)
- Drug-induced (procainamide, hydralazine, isoniazid, doxorubicin, cyclophosphamides)

VIII. Post-radiation pericarditis

IX. Post-surgical pericarditis
- Post-pericardiotomy syndrome

X. Post-traumatic pericarditis

XI. Familial and idiopathic pericarditis

Bacterial pericarditis may rarely occur with septicaemia or pneumonia, or it may stem from an early postoperative infection after thoracic surgery or trauma, or may complicate endocarditis.

Staphylococcus aureus is a frequent cause of purulent pericarditis in HIV patients. This form of pericarditis, especially staphylococcal, is fulminant and often fatal.

Other *endemic infectious pericarditis* includes mycoplasmosis and Lyme pericarditis which are often effusive and require pericardial drainage. The diagnosis is based on serological tests of pericardial fluid and identification of organisms in pericardial or myocardial biopsies.

Tuberculous pericarditis usually presents with chronic low-grade fever, particularly in the evening, associated with features of acute pericarditis, dyspnoea, malaise, night sweats and weight loss. Pericardial aspiration is often required to make the diagnosis. Constrictive pericarditis is a frequent outcome.

Fungal pericarditis is a common complication of endemic fungal infections, such as histoplasmosis and coccidioidomycosis but may be also caused by *Candida albicans*, especially in immunocompromised patients, drug addicts or after cardiac surgery.

Malignant pericarditis. Carcinoma of the bronchus, carcinoma of the breast and *Hodgkin's lymphoma* are the most common causes of malignant pericarditis. Leukaemia and malignant melanoma are also associated with pericarditis. A substantial pericardial effusion is very typical and is due to the obstruction of the lymphatic drainage from the heart. The effusion is often haemorrhagic. Radiation and therapy for thoracic tumours may cause radiation injury to the pericardium resulting in serous or haemorrhagic pericardial effusion and pericardial fibrosis. Absence of neoplastic cells in the pericardial fluid in these conditions often helps diagnosis.

Clinical features

Pericardial inflammation produces sharp central *chest pain* exacerbated by movement, respiration and lying down. It is typically relieved by sitting forward. It may be referred to the neck or shoulders. The main differential diagnoses are angina and pleurisy.

The classical clinical sign is a pericardial friction rub occurring in three phases corresponding to atrial systole, ventricular systole and ventricular diastole. It may also be heard as a biphasic 'to and fro' rub. The rub is heard best with the diaphragm of the stethoscope at the lower left sternal edge at the end of expiration with the patient leaning forward. There is usually a *fever, leucocytosis* or *lymphocytosis* when pericarditis is due to viral or bacterial infection, rheumatic fever or myocardial infarction. Features of a pericardial effusion may also be present (p. 796). Large *pericardial effusion* can compress adjacent bronchi and lung tissue and may cause *dyspnoea*.

Investigations

ECG is diagnostic. There is concave-upwards (saddle-shaped) ST elevation (Fig. 13.119). These changes evolve over time, with resolution of the ST elevation, T wave flattening/inversion and finally T wave normalization. The early ECG changes must be differentiated from ST elevation found in myocardial infarction.

Sinus tachycardia may result from fever or haemodynamic embarrassment, and rhythm and conduction abnormalities may be present if myocardium is involved. Cardiac enzymes should be assayed as they may be elevated if there is associated myocarditis (see p. 787).

Chest X-ray, echocardiograms and radionucleotide scans are of little value in uncomplicated acute pericarditis.

Treatment

If a cause is found, this should be treated. Bed rest and oral NSAIDs (high-dose aspirin, indometacin or ibuprofen) are effective in most patients. In the few days following a myocardial infarction, NSAID use is associated with a higher rate of myocardial rupture and these drugs should not be used. Corticosteroids have been used when the disease does not subside rapidly, but they are associated with side-effects. About 20% of cases of acute pericarditis go on to develop idiopathic relapsing pericarditis. The first-line treatment is again oral NSAIDs. In resistant cases, oral corticosteroids are again used to provide symptomatic relief. However, symptoms commonly recur on dose reduction or withdrawal, and prolonged steroid use is associated with its own side-effects. Numerous other treatments have been studied including azathioprine, colchicine, intravenous corticosteroids and pericardiectomy. Current data tend to favour the use of colchicine to prevent recurrent attacks of pericarditis.

Ask the authors
www.studentconsult.com

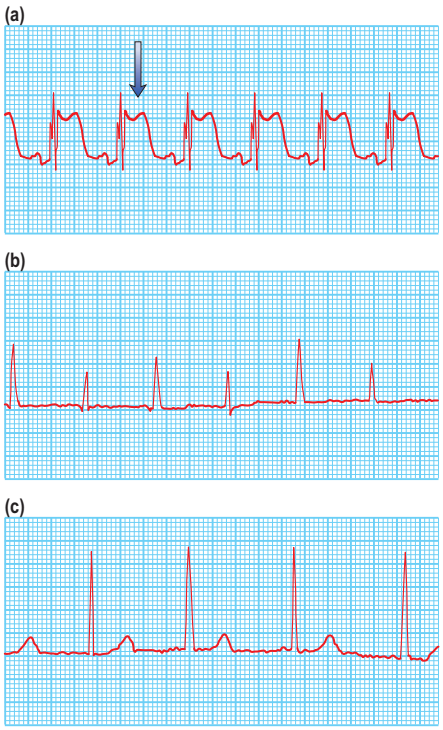


Fig. 13.119 **ECGs associated with pericarditis. (a) Acute pericarditis.** Note the raised ST segment, concave upwards (arrow). **(b) Chronic phase of pericarditis** associated with a pericardial effusion. Note the T wave flattening and inversion, and the alternation of the QRS amplitude (QRS alternans). **(c)** The same patient after evacuation of the pericardial fluid. Note that the QRS voltage has increased and the T waves have returned to normal.

If pericarditis persists for 6–12 months following the acute episode, it is considered chronic. If the pericardium thickens and restricts ventricular filling, constrictive pericarditis is said to have developed.

Pericardial effusion and cardiac tamponade

A pericardial effusion is a collection of fluid within the potential space of the serous pericardial sac (Fig. 13.120), commonly accompanying an episode of acute pericarditis. When a large volume collects in this space, ventricular filling is compromised leading to embarrassment of the circulation. This is known as cardiac tamponade.

Clinical features

Symptoms of a pericardial effusion commonly reflect the underlying pericarditis. On examination:

- Heart sounds are soft and distant.
- Apex beat is commonly obscured.
- A friction rub may be evident due to pericarditis in the early stages, but this becomes quieter as fluid accumulates and pushes the layers of the pericardium apart.
- Rarely, the effusion may compress the base of the left lung, producing an area of dullness to percussion below the angle of the left scapula (Ewart's sign).
- As the effusion worsens, signs of cardiac tamponade may become evident:
  - raised jugular venous pressure with sharp rise and '*y*' descent (Friedreich's sign)
  - Kussmaul's sign (rise in JVP/increased neck vein distension during inspiration)
  - pulsus paradoxus
  - reduced cardiac output.

Fig. 13.120 **Chest X-ray showing a pericardial effusion;** the heart appears globular.

Investigations

- **ECG** reveals low-voltage QRS complexes.
- **Chest X-ray** (Fig. 13.120) shows large globular or pear-shaped heart with sharp outlines. Typically, the pulmonary veins are not distended.
- **Echocardiography** (Fig. 13.29, p. 704) is the most useful technique for demonstrating the effusion and looking for evidence of tamponade.
- **MRI** should be considered if haemopericardium (blood in the pericardial space) or loculated pericardial effusions are suspected.
- **Pericardiocentesis** (Fig. 13.39) is the removal of pericardial fluid with aseptic technique under echocardiographic guidance. It is indicated when a tuberculous, malignant or purulent effusion is suspected.
- **Pericardial biopsy** may be needed if tuberculosis is suspected and pericardiocentesis is not diagnostic.

Other tests include looking for underlying causes, e.g. blood cultures, autoantibody screen.

Treatment

An underlying cause should be sought and treated if possible. Most pericardial effusions resolve spontaneously. However, when the effusion collects rapidly, tamponade may result. Pericardiocentesis is then indicated to relieve the pressure – a drain may be left in temporarily to allow sufficient release of fluid.

Pericardial effusions may reaccumulate, most commonly due to malignancy (in the UK). This may require pericardial fenestration, i.e. creation of a window in the pericardium to allow the slow release of fluid into the surrounding tissues. This procedure may either be performed transcutaneously under local anaesthetic or using a conventional surgical approach.

Constrictive pericarditis

Certain causes of pericarditis such as tuberculosis, haemopericardium, bacterial infection and rheumatic heart disease result in the pericardium becoming thick, fibrous and calcified. This may also develop late after open heart surgery, and fibrosis also occurs with the use of dopamine agonists, e.g. cabergoline, perigolide. In many cases these pericardial changes do not cause any symptoms. If, however, the pericardium becomes so inelastic as to interfere with diastolic filling of the heart, constrictive pericarditis is said to have developed. As these changes are chronic, allowing the body time to compensate, this condition is not as immediately life-threatening as cardiac tamponade, in which the circulation is more acutely embarrassed.

Constrictive pericarditis should be distinguished from restrictive cardiomyopathy (see p. 792). The two conditions are very similar in their presentation, but the former is fully treatable, whereas most cases of the latter are not. In the later stages of constrictive pericarditis the subepicardial layers of myocardium may undergo fibrosis, atrophy and calcification.

Clinical features

The symptoms and signs of constrictive pericarditis occur due to:

- reduced ventricular filling (similar to cardiac tamponade, i.e. Kussmaul's sign, Friedreich's sign, pulsus paradoxus)
- systemic venous congestion (ascites, dependent oedema, hepatomegaly and raised JVP)
- pulmonary venous congestion (dyspnoea, cough, orthopnoea, PND) less commonly
- reduced cardiac output (fatigue, hypotension, reflex tachycardia)
- rapid ventricular filling ('pericardial knock' heard in early diastole at the lower left sternal border)
- atrial dilatation (30% of cases have atrial fibrillation).

Investigations

- **Chest X-ray** (Fig. 13.121): shows a relatively small heart in view of the symptoms of heart failure. Pericardial calcification is present in up to 50%. A lateral chest film may be useful in detecting calcification that is missed on an AP film. However, a calcified pericardium is not necessarily a constricted one
- **ECG** reveals low-voltage QRS complexes with generalized T wave flattening or inversion.
- **Echocardiography** shows thickened calcified pericardium, and small ventricular cavities with normal wall thickness. Doppler studies may be useful.
- **CT and CMR** are used to assess pericardial anatomy and thickness (3 mm or greater) (see Fig. 13.33).
- **Endomyocardial biopsy** may be helpful in distinguishing constrictive pericarditis from restrictive cardiomyopathy in difficult cases.
- **Cardiac catheterization**. End-diastolic pressures in the left and right ventricles measured during this procedure are usually equal, owing to pericardial constriction.

Restrictive cardiomyopathy is a close mimic of constrictive pericarditis and all the above tests may not help to distinguish the two conditions.

Treatment

The treatment for chronic constrictive pericarditis is complete resection of the pericardium. This is a risky procedure with a high complication rate due to the presence of myocardial atrophy in many cases at the time of surgery. Thus early pericardiectomy is suggested in non-tuberculous cases, before severe constriction and myocardial atrophy have developed.

In cases of tuberculous constriction, the presence of pericardial calcification implies chronic disease. Current evidence tends to favour early pericardiectomy with antituberculous drug cover in these cases. If there is no calcification, a course of antituberculous therapy should be attempted first. If the patient's haemodynamic state remains static or deteriorates after 4–6 weeks of therapy, pericardiectomy is recommended.

FURTHER READING

Yusuf S, Cairns JA, Camm AJ, Fallen EL, Gersh BJ (eds). Pericardial disease. In: *Evidence-Based Cardiology,* 3rd edn. London: BMJ Books, 2003: 407–422.

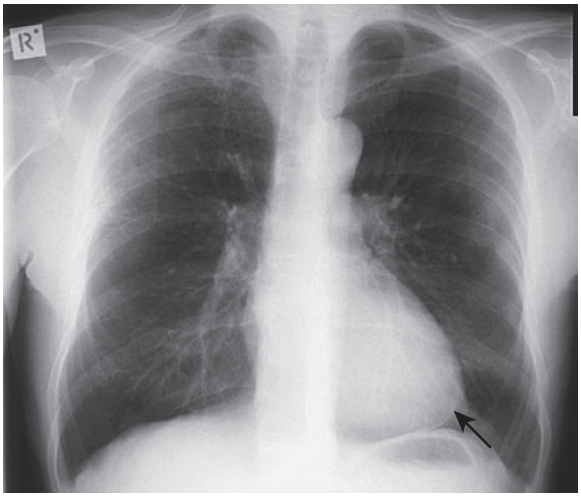


Fig. 13.121 **Chest X-ray showing a pericardial calcification** (arrow).

HEART DISEASE IN THE ELDERLY

The increasing age of the population is associated with an increased burden of cardiovascular disease in the elderly. In general, the management of elderly patients with cardiovascular disease is the same as for younger patients. However, elderly patients may present with atypical symptoms:

- falls – bradycardia, atrioventricular block, vasovagal syncope
- confusion – myocardial infarction, infective endocarditis
- breathlessness – pneumonia, left ventricular failure, diastolic heart failure, ischaemic heart disease.

The drug treatment of elderly patients may require modification as drug absorption, and renal and hepatic clearance are reduced. These patients are frequently on polypharmacy with multiple drugs that may interact with each other or influence co-morbid conditions, e.g. NSAIDs may cause fluid retention

and exacerbate heart failure. Elderly patients are at increased risk both from interventions (percutaneous coronary angiography, coronary artery bypass grafting) and cardiovascular disease – the decision to intervene should be considered on an individual basis.

SYSTEMIC HYPERTENSION

Definitions of hypertension

Elevated arterial blood pressure is a major cause of premature vascular disease leading to cerebrovascular events, ischaemic heart disease and peripheral vascular disease. Blood pressure is a characteristic of each individual, like height and weight, with marked interindividual variation, and has a continuous (bell-shaped) distribution. The levels of blood pressure observed depend on the characteristics of the population studied – in particular, the age and ethnic background. Blood pressure in industrialized countries rises with age, certainly up to the seventh decade. This rise is more marked for systolic pressure and is more pronounced in men. Hypertension is very common in the developed world. Depending on the diagnostic criteria, hypertension is present in 20–30% of the adult population. Hypertension rates are much higher in black Africans (40–45% of adults).

The definition of an abnormal blood pressure is indicated in Table 13.48.

The risk of mortality or morbidity rises progressively with increasing systolic and diastolic pressures, with each measure having an independent prognostic value; for example, isolated systolic hypertension is associated with a two- to three-fold increase in cardiac mortality.

A prehypertension category has been added to reflect the continium between normal and abnormal blood pressure.

Table 13.48 Classification of blood pressure levels of the British Hypertension Society

Category	Systolic blood pressure (mmHg)	Diastolic blood pressure (mmHg)
Blood pressure		
Optimal	<120 and	<80
Normal	120–129 and/or	<85
High normal*	130–139 and/or	85–89
Hypertension		
Grade 1 (mild)	140–159 and/or	90–99
Grade 2 (moderate)	160–179 and/or	100–109
Grade 3 (severe)	≥180	≥109
Isolated systolic hypertension		
Grade 1	140–149	<90
Grade 2	≥160	<90

*Equivalent to pre-hypertension.

The European Societies of Hypertension and Cardiology Guidelines 2007 are based on clinical blood pressure and not values for ambulatory blood pressure measurement. Threshold blood pressure levels for the diagnosis of hypertension using self/home monitoring are greater than 135/85 mmHg. For ambulatory monitoring, 24-hour values are greater than 125/80 mmHg. If systolic blood pressure and diastolic blood pressure fall into different categories, the higher value should be taken for classification.

All adults should have blood pressure measured routinely at least every 5 years until the age of 80 years. Seated blood pressure when measured after 5 minutes' resting with appropriate cuff size and arm supported is usually sufficient, but standing blood pressure should be measured in diabetic and elderly subjects to exclude orthostatic hypotension. The cuff should be deflated at 2 mm/s and the blood pressure measured to the nearest 2 mmHg. Two consistent blood pressure measurements are needed to estimate blood pressure, and more are recommended if there is variation in the pressure. When assessing the cardiovascular risk, the average blood pressure at separate visits is more accurate than measurements taken at a single visit.

Causes

The majority (80–90%) of patients with hypertension have primary elevation of blood pressure, i.e. essential hypertension of unknown cause.

Essential hypertension

Essential hypertension has a multifactorial aetiology.

Genetic factors

Blood pressure tends to run in families and children of hypertensive parents tend to have higher blood pressure than age-matched children of parents with normal blood pressure. This familial concordance of blood pressure may be explained, at least in part, by shared environmental influences. However, there still remains a large, still largely unidentified genetic component.

Fetal factors

Low birth weight is associated with subsequent high blood pressure. This relationship may be due to fetal adaptation to intrauterine undernutrition with long-term changes in blood vessel structure or in the function of crucial hormonal systems.

Environmental factors

Amongst the several environmental factors that have been proposed, the following seem to be the most significant:

Obesity. Fat people have higher blood pressures than thin people. There is a risk, however, of overestimation if the blood pressure is measured with a small cuff. Adjust the bladder size to the arm circumference. Sleep disordered breathing (see p. 842) often seen with obesity may be an additional risk factor.

Alcohol intake. Most studies have shown a close relationship between the consumption of alcohol and blood pressure level. However, subjects who consume small amounts of alcohol seem to have lower blood pressure level than those who consume no alcohol.

Sodium intake. A high sodium intake has been suggested to be a major determinant of blood pressure differences between and within populations around the world. Populations with higher sodium intakes have higher average blood pressures than those with lower sodium intake. Migration from a rural to an urban environment is associated with an increase in blood pressure that is in part related to the amount of salt in the diet. Studies of the restriction of salt intake have shown a beneficial effect on blood pressure in

hypertensives. There is some evidence that a high potassium diet can protect against the effects of a high sodium intake.

Stress. Whilst acute pain or stress can raise blood pressure, the relationship between chronic stress and blood pressure is uncertain.

Humoral mechanisms

The autonomic nervous system, as well as the renin–angiotensin, natriuretic peptide and kallikrein–kinin system, plays a role in the physiological regulation of short-term changes in blood pressure and have been implicated in the pathogenesis of essential hypertension. A low renin, salt-sensitive, essential hypertension in which patients have renal sodium and water retention has been described. However, there is no convincing evidence that the above systems are directly involved in the maintenance of hypertension.

Insulin resistance

An association between diabetes and hypertension has long been recognized and a syndrome has been described of hyperinsulinaemia, glucose intolerance, reduced levels of HDL cholesterol, hypertriglyceridaemia and central obesity (all of which are related to insulin resistance) in association with hypertension. This association (also called the 'metabolic syndrome', p. 231) is a major risk factor for cardiovascular disease.

Secondary hypertension

Secondary hypertension is where blood pressure elevation is the result of a specific and potentially treatable cause. Secondary forms of hypertension include the following:

Renal diseases

These account for over 80% of the cases of secondary hypertension. The common causes are:

- diabetic nephropathy
- chronic glomerulonephritis
- adult polycystic disease
- chronic tubulointerstitial nephritis
- renovascular disease.

Hypertension can itself cause or worsen renal disease. The mechanism of this blood pressure elevation is primarily due to sodium and water retention, although there can be inappropriate elevation of plasma renin levels.

Endocrine causes

These include:

- Conn's syndrome
- adrenal hyperplasia
- phaeochromocytoma
- Cushing's syndrome
- acromegaly.

Congenital cardiovascular causes

The major cause is coarctation of the aorta.

Drugs

Many drugs have been shown to cause or aggravate hypertension, or interfere with the response to some antihypertensive agents: NSAIDs, oral contraceptives, steroids, carbenoxolone, liquorice, sympathomimetics and vasopressin. Patients taking monoamine oxidase inhibitors who consume tyramine-containing foods may develop paroxysms of severe hypertension.

Pregnancy

Cardiac output rises in pregnancy but, owing to a relatively greater fall in peripheral resistance, blood pressure in pregnant women is usually lower than in those who are not pregnant. Hypertension is noted in 8–10% of pregnancies; when detected in the first half of pregnancy or persisting after delivery, it is usually due to pre-existing essential hypertension. Hypertension presenting in the second half of pregnancy – or 'pregnancy-induced hypertension' – usually resolves after delivery. When the blood pressure increases to >160/110 mmHg treatment is warranted for the protection of the mother. Pre-eclampsia is a syndrome consisting of pregnancy-induced hypertension with proteinuria. The primary pathology is unknown, but is likely to involve a disturbance of the uteroplacental circulation and result in intrauterine growth restriction. Hypertension in pregnancy, together with pulmonary embolus, are the most common causes of maternal death, with a rate of 10 per million pregnancies. Furthermore, the critical condition of eclampsia, which is associated with severe hypertension, may ultimately lead to convulsions, cerebral and pulmonary oedema, jaundice, clotting abnormalities and fetal death.

Pathophysiology

The pathogenesis of essential hypertension remains unclear. In some young hypertensive patients, there is an early increase in cardiac output, in association with increased pulse rate and circulating catecholamines. This could result in changes in baroreceptor sensitivity, which would then operate at a higher blood pressure level.

In *chronic hypertension*, the cardiac output is normal and it is an increased peripheral resistance that maintains the elevated blood pressure. The resistance vessels (the small arteries and arterioles) show structural changes in hypertension with an increase in wall thickness and a reduction in the vessel lumen diameter. There is also some evidence for rarefaction (decreased density) of these vessels. These mechanisms would result in an increased overall peripheral vascular resistance.

Hypertension also causes changes in the large arteries. There is thickening of the media, an increase in collagen and the secondary deposition of calcium. These changes result in a loss of arterial compliance, which in turn leads to a more pronounced arterial pressure wave.

Pulse wave velocity is a measure of arterial stiffness and is inversely related to distensibility. With each systolic contraction a pulse wave travels down the arterial wall before the flow of blood. Thus, the more rigid the arterial wall, the faster the wave travels. It can be measured but is not in routine use. Atheroma develops in the large arteries owing to the interaction of these mechanical stresses and low growth factors (see p. 744). Endothelial dysfunction with alternations in agents such as nitric oxide and endothelins appear to be involved.

Left ventricular hypertrophy, which results from increased peripheral vascular resistance and increased left ventricular load, is a significant prognostic indicator of future cardiovascular events.

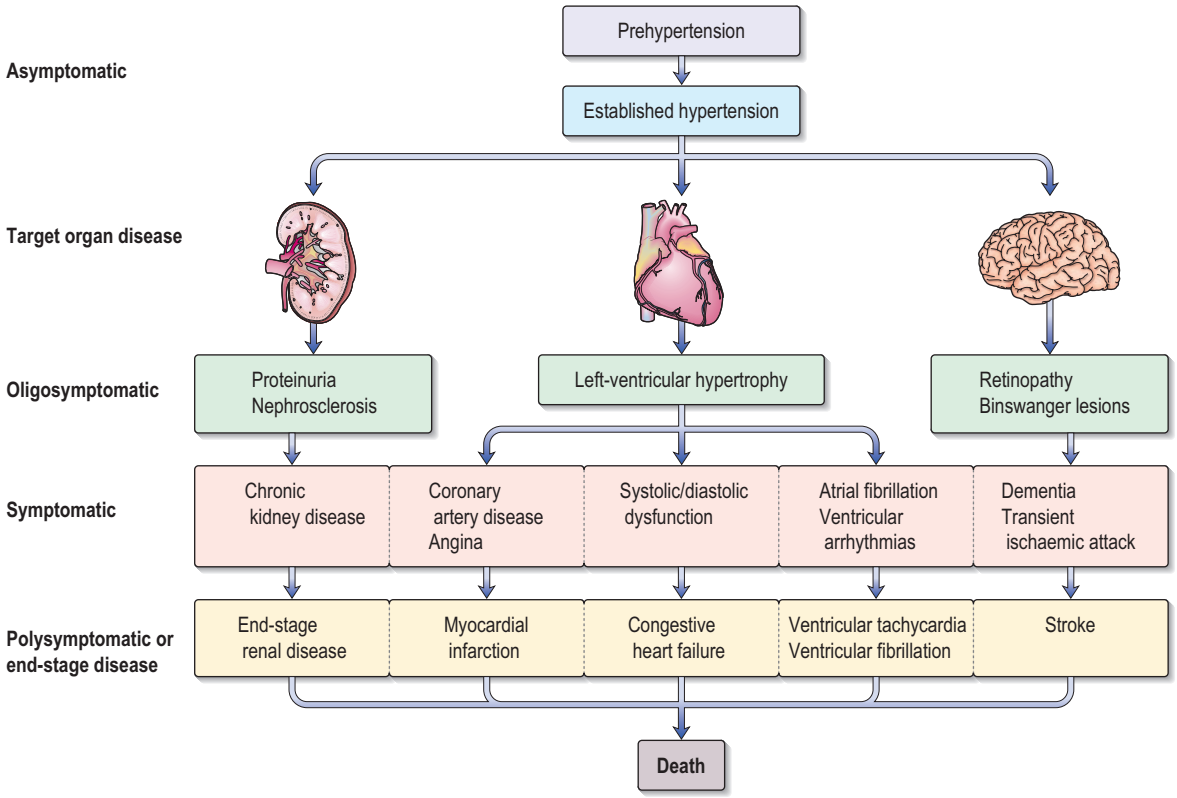


Fig. 13.122 **Range of hypertensive cardiovascular disease from prehypertension to target-organ damage and end-stage disease.** From Messerli FH, Williams B, Ritz E. Essential hypertension. *Lancet* 2007; **370**: 591–603, with permission from Elsevier.

Changes in the *renal vasculature* eventually lead to a reduced renal perfusion, reduced glomerular filtration rate and, finally, a reduction in sodium and water excretion. The decreased renal perfusion may lead to activation of the renin–angiotensin system (renin converts angiotensinogen to angiotensin I, which is in turn converted to angiotensin II by angiotensin-converting enzyme) with increased secretion of aldosterone and further sodium and water retention.

Complications

Cerebrovascular disease and coronary artery disease are the most common causes of death, although hypertensive patients are also prone to renal failure and peripheral vascular disease (Fig. 13.122).

Hypertensives have a sixfold increase in stroke (both haemorrhagic and atherothrombotic). There is a threefold increase in cardiac death (due either to coronary events or to cardiac failure). Furthermore, peripheral arterial disease is twice as common.

Malignant hypertension

Malignant or accelerated hypertension occurs when blood pressure rises rapidly and is considered with severe hypertension (diastolic blood pressure >120 mmHg) (see p. 805). The characteristic histological change is fibrinoid necrosis of the vessel wall and, unless treated, it may lead to death from progressive renal failure, heart failure, aortic dissection or stroke. The changes in the renal circulation result in rapidly progressive renal failure, proteinuria and haematuria. There is also a high risk of cerebral oedema and haemorrhage with resultant hypertensive encephalopathy. In the retina there may be flame-shaped haemorrhages, cotton wool spots, hard exudates and papilloedema (Fig. 13.123). Without effective treatment there is a 1-year survival of less than 20%.

Fig. 13.123 **Fundus showing hypertensive changes:** Grade 4 retinopathy with papilloedema, haemorrhages and exudates.

Assessment

Management should be considered in three stages: assessment, non-pharmacological treatment and drug treatment. During the assessment period, secondary causes of hypertension should be excluded, target-organ damage from the

blood pressure should be evaluated and any concomitant conditions (e.g. dyslipidaemia or diabetes) that may add to the cardiovascular burden should be identified.

History

The patient with mild hypertension is usually asymptomatic. Attacks of sweating, headaches and palpitations point towards the diagnosis of phaeochromocytoma. Higher levels of blood pressure may be associated with headaches, epistaxis or nocturia. Breathlessness may be present owing to left ventricular hypertrophy or cardiac failure, whilst angina or symptoms of peripheral arterial vascular disease suggest the diagnosis of atheromatous renal artery stenosis. This is usually a local manifestation of more generalized atherosclerosis, and patients are often elderly with coexistent vascular disease (Fig. 13.124). Fibromuscular disease of the renal arteries encompasses a group of conditions in which fibrous or muscular proliferation results in morphologically simple or complex stenoses and tends to occur in younger patients. Malignant hypertension may present with severe headaches, visual disturbances, fits, transient loss of consciousness or symptoms of heart failure.

Examination

Elevated blood pressure is usually the only abnormal sign. Signs of an underlying cause should be sought, such as renal artery bruits in renovascular hypertension, or radiofemoral delay in coarctation of the aorta. The cardiac examination may also reveal features of left ventricular hypertrophy and a loud aortic second sound. If cardiac failure develops, there may be a sinus tachycardia and a third heart sound.

Fundoscopy is an essential part of the examination of any hypertensive patient (Fig. 13.123). The abnormalities are graded according to the Keith–Wagener classification:

- *Grade 1* – tortuosity of the retinal arteries with increased reflectiveness (silver wiring).
- *Grade 2* – grade 1 plus the appearance of arteriovenous nipping produced when thickened retinal arteries pass over the retinal veins.
- *Grade 3* – grade 2 plus flame-shaped haemorrhages and soft ('cotton wool') exudates actually due to small infarcts.
- *Grade 4* – grade 3 plus papilloedema (blurring of the margins of the optic disc).

Grades 3 and 4 are diagnostic of malignant hypertension.

Ambulatory blood pressure monitoring

Indirect automatic blood pressure measurements can be made over a 24-hour period using a measuring device worn by the patient. The clinical role of such devices remains uncertain, although they are used to confirm the diagnosis in those patients with 'white-coat' hypertension, i.e. blood pressure is completely normal at all stages except during a clinical consultation (Fig. 13.125a). These patients do not have any evidence of target-organ damage, and unnecessary treatment can be avoided. These devices may also be used

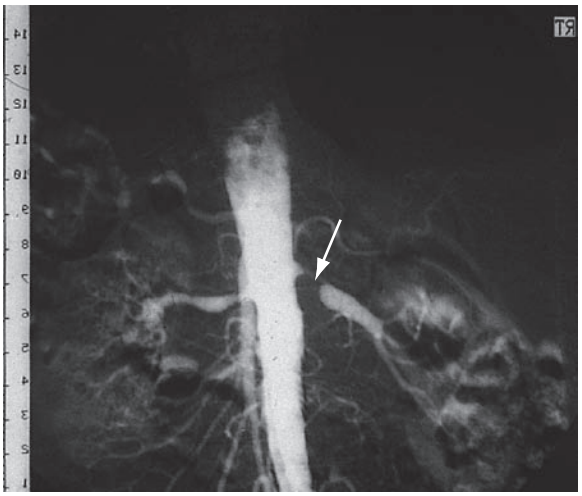


Fig. 13.124 **Digital subtraction angiography, showing typical unilateral atheromatous renal artery stenosis with post-stenotic dilatation** (arrow).

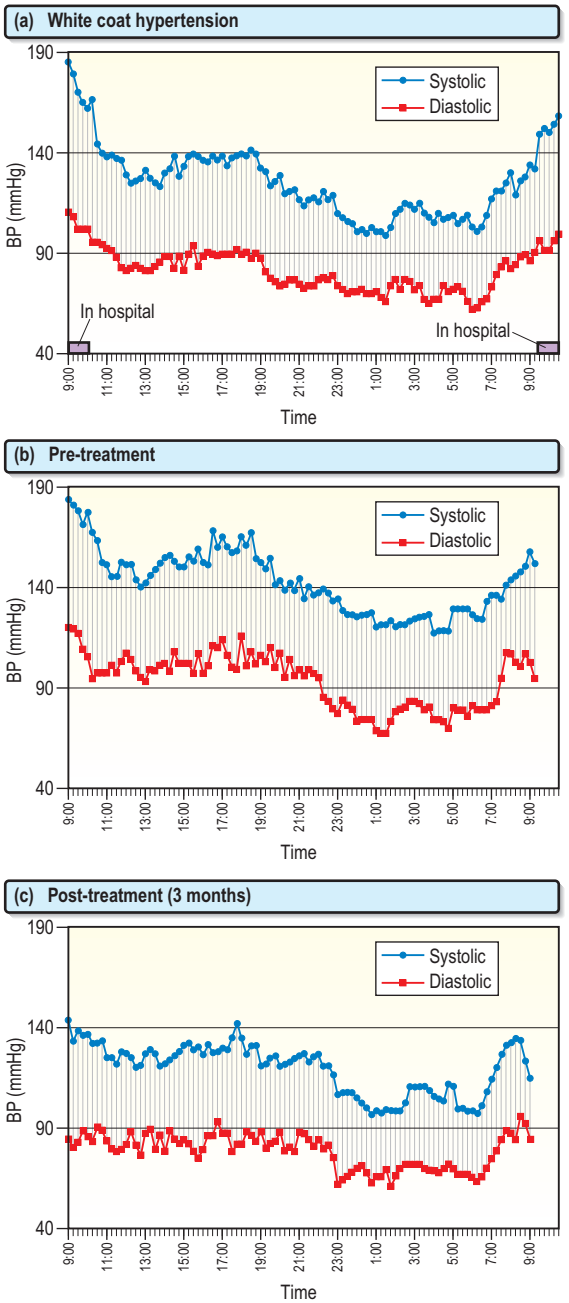


Fig. 13.125 **24-Hour ambulatory blood pressure monitoring**, showing: **(a)** white-coat hypertension; **(b)** pre-treatment; **(c)** after 3 months' treatment.

to monitor the response of patients to drug treatment and, in particular, can be used to determine the adequacy of 24-hour control with once-daily medication (Fig. 13.125b,c).

Ambulatory blood pressure recordings seem to be better predictors of cardiovascular risk than clinic measurements. Analysis of the diurnal variation in blood pressure suggests that those hypertensives with loss of the usual nocturnal fall in blood pressure ('non-dippers') have a worse prognosis than those who retain this pattern.

Investigations

Routine investigation of the hypertensive patient should include:

- ECG
- urine stix test for protein and blood
- fasting blood for lipids (total and HDL cholesterol) and glucose
- serum urea, creatinine and electrolytes.

If the urea or creatinine is elevated, more specific renal investigations are indicated – creatinine clearance, renal ultrasound (in case of polycystic kidney disease, or parenchymal renal artery disease) and a renal isotope scan or renal angiography if renovascular disease (either atheromatous or fibromuscular dysplasia) is suspected. A low serum potassium may indicate an endocrine disorder (either primary hyperaldosteronism or glucocorticoid excess), and aldosterone, cortisol and renin measurements must then be made, preferably prior to initiating pharmacological therapy. Clinical suspicion of phaeochromocytoma should be investigated further with measurement of urinary metanephrines and plasma or urinary catecholamines.

If the ECG shows evidence of coronary artery disease the coronary vascular status should be assessed. If left ventricular hypertrophy is suspected *echocardiography* (or MRI) should be undertaken. A *chest X-ray* is indicated if cardiac involvement or aortic coarctation is likely.

Treatment

Unless the patient has severe or malignant hypertension, there should be a period of assessment with repeated blood pressure measurements, combined with advice and non-pharmacological measures prior to the initiation of drug therapy. The guidelines of the British Hypertension Society (BHS) suggest the following.

Use of non-pharmacological therapy in all hypertensive and borderline hypertensive people:

- weight reduction – BMI should be <25 kg/m^2
- low-fat and saturated fat diet
- low-sodium diet – <6 g sodium chloride per day
- limited alcohol consumption – ≤ 21 units/week for men and ≤ 14 units/week for women
- dynamic exercise – at least 30 minutes' brisk walk per day
- increased fruit and vegetable consumption
- reduce cardiovascular risk by stopping smoking and increasing oily fish consumption.

Pharmacological therapy (Table 13.49) should be based on the following:

- The initiation of antihypertensive therapy in subjects with sustained systolic blood pressure (BP) ≥ 160 mmHg, or sustained diastolic BP ≥ 100 mmHg.
- Decide on treatment in subjects with sustained systolic blood pressure between 140 and 159 mmHg, or sustained diastolic BP between 90 and 99 mmHg, according to the presence or absence of target organ damage or a 10-year cardiovascular disease risk $>20\%$.
- In patients with diabetes mellitus, the initiation of antihypertensive drug therapy if systolic BP is sustained ≥ 140 mmHg, or diastolic BP is sustained ≥ 90 mmHg.
- *Target blood pressure*. For most patients a target of ≈ 140 mmHg systolic blood pressure and ≈ 85 mmHg diastolic blood pressure is recommended. For patients with diabetes, renal impairment or established cardiovascular disease a lower target of $\approx 130/80$ mmHg is recommended.
- When using ambulatory blood pressure readings, mean daytime pressures are preferred and this value would be expected to be approximately 10/5 mmHg lower than the clinic blood pressure equivalent for both thresholds and targets. Similar adjustments are recommended for averages of home blood pressure readings.

Table 13.49 Advantages and disadvantages of drugs used in hypertension with respect to associated conditions

	Diuretics	Beta-blockers	ACE inhibitor/ angiotensin II receptor antagonist	Calcium-channel blockers	Alpha-blockers
Diabetes	Care*	Care*	Yes	Yes	Yes
Gout	No	Yes	Yes	Yes	Yes
Dyslipidaemia	Care†	Care†	Yes	Yes	Yes
Ischaemic heart disease	Yes	Yes	Yes	Yes	Yes
Heart failure	Yes	Yes	Yes	Care‡	Yes
Asthma	Yes	No	Yes	Yes	Yes
Peripheral vascular disease	Yes	Care	Care§	Yes	Yes
Renal artery stenosis	Yes	Care	No	Yes	Yes
Pregnancy	Caution	Not in late pregnancy	No	No	Caution

*Diuretics may aggravate diabetes: beta-blockers worsen glucose intolerance and mask symptoms of hypoglycaemia.
†Both diuretics and beta-blockers disturb the lipid profile.
‡Verapamil and diltiazem may exacerbate heart failure, although amlodipine appears to be safe.
§Patients with peripheral vascular disease may also have renal artery stenosis; therefore, ACE inhibitors should be used cautiously.

- The main determinant of outcome following treatment is the level of blood pressure reduction that is achieved rather than the specific drug used to lower blood pressure.
- Most hypertensive patients will require a combination of antihypertensive drugs to achieve the recommended targets.
- In most hypertensive patients therapy with statins and aspirin is added to reduce the overall cardiovascular risk burden. Glycaemic control should be optimized in diabetics (HbA_{1c} <7%).

The decision to commence specific drug therapy should usually be made only after a careful period of assessment, of up to 6 months, with repeated measurements of blood pressure (Fig. 13.126). The aim of drug treatment to reduce the risk of complications of hypertension should be carefully explained to the patient and a plan for the patient's treatment (drug dose titration, change of drug and combination of drugs) should be agreed with the patient. All of the drugs used to treat hypertension have side-effects and, since the benefits of drug treatment are not immediate, compliance may be a major problem.

Several classes of drugs are available to treat hypertension. The usual are: (a) ACE inhibitors or angiotensin receptor antagonists; (b) beta-blockers; (c) calcium-channel blockers; or (d) diuretics. It is recommended that drugs are chosen according to the scheme laid out in Figure 13.127.

The rationale for *step 1* in this scheme is that young Caucasians are more likely to have high renin hypertension, and older patients and black patients usually have low renin hypertension. If a drug within each pair is not tolerated, the alternative drug type can be used (e.g. if an ACE inhibitor is not tolerated, an angiotensin receptor antagonist). If a drug is not effective, a drug from the other group should be selected. Thus, if a calcium-channel blocker is not helpful, an ACE inhibitor/angiotensin receptor antagonist should be tried. Almost all patients will need more than one drug to effectively lower blood pressure.

Step 2 involves combining one drug from each group.

In *step 3* an ACE inhibitor (or angiotensin receptor antagonist) is combined with a calcium-channel blocker and diuretic. If triple therapy is not sufficient to achieve target blood pressure readings, an alpha-blocker, beta-blocker or spironolactone, or another agent may be used. It is not advised to combine a diuretic with a beta-blocker since both aggravate diabetes.

Diuretics

Thiazide diuretics such as bendroflumethiazide (2.5–5 mg daily) and cyclopenthiazide (0.25–0.5 mg daily) are well-

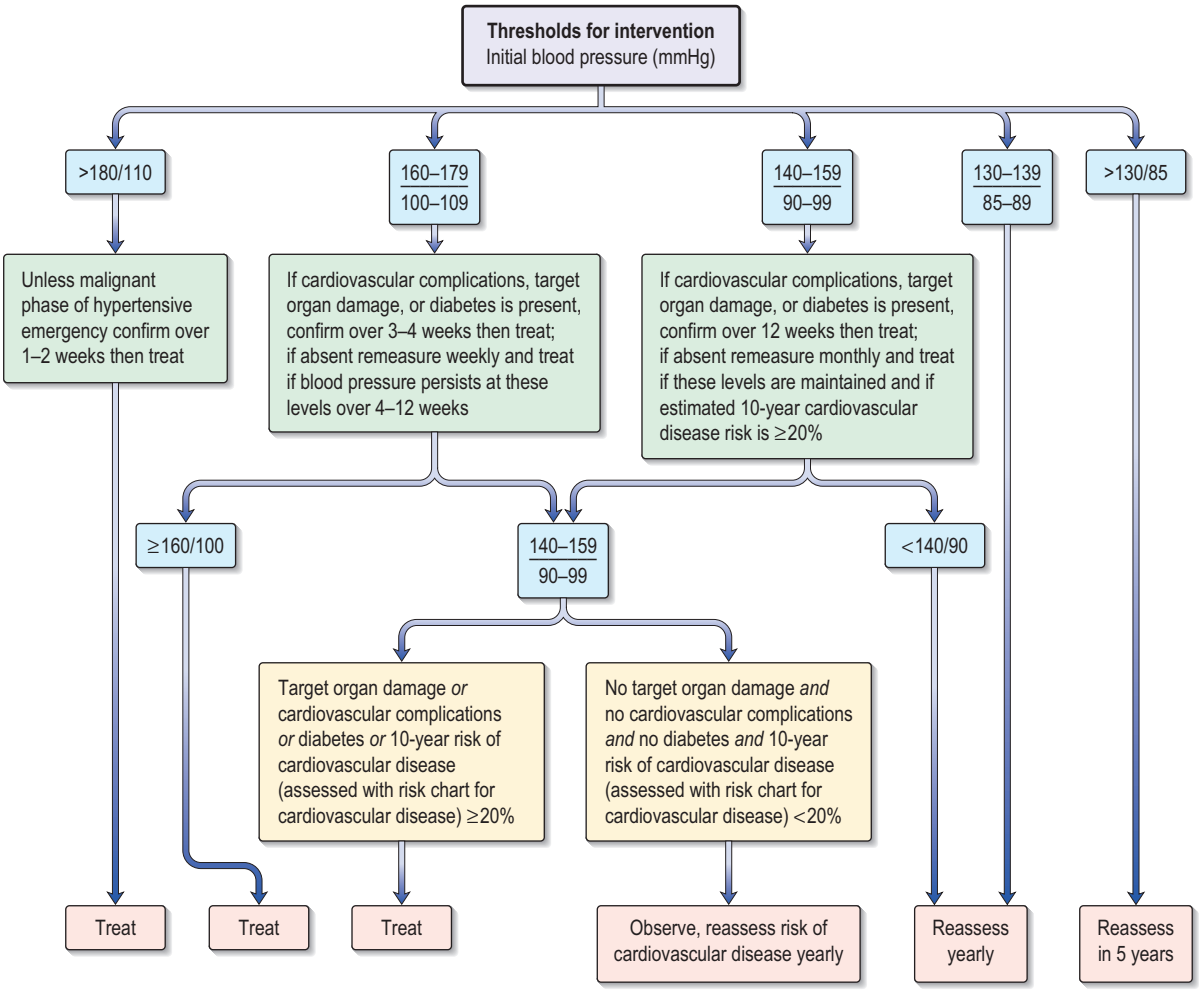


Fig. 13.126 **When to initiate treatment**. From Williams B, Poulter MR, Brown MJ, et al. British Hypertension Society guidelines for hypertension management 2004 (BHS-IV): summary. *British Medical Journal* 2004; **328**: 634–640, with permission from the BMJ Publishing Group.

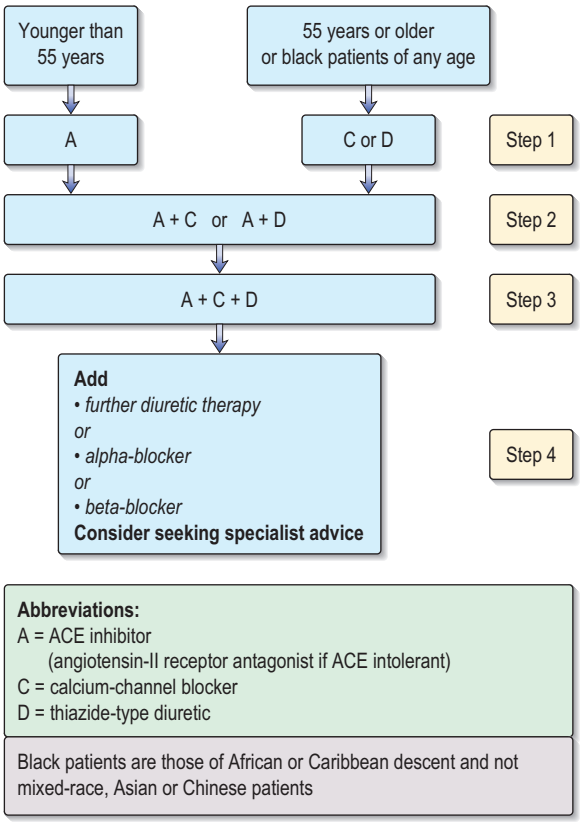


Fig. 13.127 **Choosing drugs for patients newly diagnosed with hypertension.** National Institute for Health and Clinical Excellence. 'Choosing drugs for patients newly diagnosed with hypertension' in CG 34 Hypertension: management of hypertension in adults in primary care (Quick Reference Guide). London: NICE, 2006. Available from www.nice.org.uk/CG034. Reproduced with permission.

established agents which have been shown to reduce the risk of stroke in patients with hypertension. The lower doses seems to be equally effective as higher doses in the reduction of blood pressure and most have a duration of up to 24 hours. The major concern with these agents is their adverse metabolic effects, particularly increased serum cholesterol, impaired glucose tolerance, hyperuricaemia (which may precipitate gout) and hypokalaemia. These tend to occur with higher doses of thiazide diuretics.

Loop diuretics such as furosemide (40 mg daily) do have a hypotensive effect, but are not routinely used in the treatment of essential hypertension. Potassium-sparing diuretics such as amiloride (5–10 mg daily) or spironolactone (50–200 mg daily) are not effective agents when used alone, with the exception of spironolactone in the treatment of hypertension and hypokalaemia associated with primary hyperaldosteronism.

Beta-adrenoceptor blockers

Beta-blockers are no longer a preferred initial therapy for hypertension. However, beta-blockers are used in younger people, particularly those with an intolerance or contraindication to ACE inhibitors and angiotensin-II receptor antagonists; women of child-bearing potential; or patients with evidence of increased sympathetic drive. In these circumstances, if therapy is initiated with a beta-blocker and a second drug is required, add a calcium-channel blocker rather than a thiazide-type diuretic to reduce the patient's risk of developing diabetes. Beta-blockers exert their effects by attenuating the effects of the sympathetic nervous and the renin–angiotensin systems. The major side-effects of this class of agents are bradycardia, bronchospasm, cold extremities, fatigue, bad dreams and hallucinations. These agents are especially useful in the treatment of patients with both hypertension and angina. The drugs include atenolol (50 mg daily), bisoprolol (10–20 mg daily), metoprolol (100–200 mg in divided doses daily) and propranolol (160–320 mg in divided doses daily).

Atenolol has been shown to reduce brachial arterial pressure but not aortic pressure, which is more significant in causing strokes and heart attacks. Atenolol is now not a preferred drug for hypertension.

Angiotensin-converting enzyme (ACE) inhibitors

These drugs block the conversion of angiotensin I to angiotensin II, which is a potent vasoconstrictor. They also block the degradation of bradykinin, a potent vasodilator. There is evidence that black African patients respond less well to ACE inhibitors unless combined with diuretics. They are particularly useful in diabetics with nephropathy, where they have been shown to slow disease progression, and in those patients with symptomatic or asymptomatic left ventricular dysfunction, where they have been shown to improve survival.

The major potential side-effects are profound hypotension following the first dose, which is usually seen in sodium-depleted patients or in those on treatment with large doses of diuretics, and deterioration of renal function in those with severe bilateral renovascular disease (in whom the production of angiotensin II is playing a major role in maintaining renal perfusion by causing efferent arteriolar constriction at the glomerulus). They also cause mild dry cough in a number of patients, especially if prescribed at high doses, due to their effect on bradykinin.

These are several ACE inhibitors available and there are no significant differences between them in terms of blood pressure effect; those with the longest duration of action may be taken once-daily, which is clearly a benefit in terms of compliance. The drugs include enalapril (10–20 mg daily), captopril (50–150 mg daily in divided doses), ramipril (2.5–10 mg daily), lisinopril (10–20 mg daily) and trandolapril (1–4 mg daily).

Angiotensin II receptor antagonists

This group of agents selectively block the receptors for angiotensin II. They share many of the actions of ACE inhibitors but, since they do not have any effect on bradykinin, do not cause a cough. They are currently used for patients who cannot tolerate ACE inhibitors because of persistent cough. Angioneurotic oedema and renal dysfunction are encountered less with these drugs than with ACE inhibitors. The agents include losartan (50–100 mg daily), candesartan (up to 32 mg daily), valsartan (80–160 mg daily), irbesartan (75–300 mg daily) and telmisartan (20–80 mg/daily).

Calcium-channel blockers

These agents effectively reduce blood pressure by causing arteriolar dilatation, and some also reduce the force of cardiac contraction. Like the beta-blockers, they are especially useful in patients with concomitant ischaemic heart disease. The major side-effects are particularly seen with the

FURTHER READING

National Institute of Clinical Excellence. Hypertension: management of hypertension in adults in primary care. Clinical Guideline 34, 2006. www.nice.org.uk.

O'Brien E, Asmar R, Beilin L, et al. European Society of Hypertension recommendations for conventional, ambulatory and home blood pressure measurement. *Journal of Hypertension* 2003; **21**: 821–848.

Williams B, et al.; British Hypertension Society. Guidelines for management of hypertension: report of the fourth working party of the British Hypertension Society, 2004-BHS IV. *Journal of Human Hypertension* 2004; **18**: 139–185.

short-acting agents and include headache, sweating, swelling of the ankles, palpitations and flushing. Many of these side-effects can be lessened by the co-administration of a beta-blocker. The short-acting agents, such as nifedipine (10–20 mg three times daily) are being replaced by once-daily agents that are very well tolerated and include amlodipine (5–10 mg daily), felodipine (5–20 mg daily) and long-acting nifedipine (20–90 mg daily).

Alpha-blockers

These agents cause postsynaptic α_1-receptor blockade with resulting vasodilatation and blood pressure reduction. Earlier short-acting agents caused serious first-dose hypotension, but the newer longer-acting agents are far better tolerated. These include doxazosin (1–4 mg daily). Labetalol is an agent that has combined alpha- and beta-blocking properties, but is not commonly used, except in pregnancy-induced hypertension.

Renin inhibitors

Aliskerin is the first orally active renin inhibitor which directly inhibits plasma renin activity: it reduces the negative feedback by which angiotensin II inhibits renin release. It has been used in combination with ACE inhibitors and angiotensin receptor blockers with a significant reduction in blood pressure. Side-effects are few but hypokalaemia occurs.

Other vasodilators

These include hydralazine (up to 100 mg daily) and minoxidil (up to 50 mg daily). Both are extremely potent vasodilators that are reserved for patients resistant to other forms of treatment. Hydralazine can be associated with tachycardia, fluid retention and a systemic lupus erythematosus-like syndrome. Minoxidil can cause severe oedema, excessive hair growth and coarse facial features. If these agents are used, it is usually in combination with a beta-blocker.

Centrally acting drugs

Reserpine is used in a low dose of 0.05 mg/day, which provides almost all its antihypertensive action with fewer side-effects than higher doses. It has a slow onset of action (measured in weeks). Methyldopa is still widely used despite central and potentially serious hepatic and blood side-effects. It acts on central α_2-receptors, usually without slowing the heart. Clonidine and moxonidine provide all the benefits of methyldopa with none of the rare (but serious) autoimmune reactions.

Management of severe or malignant hypertension

Patients with severe hypertension (diastolic pressure >140 mmHg), malignant hypertension (grades 3 or 4 retinopathy), hypertensive encephalopathy or with severe hypertensive complications, such as cardiac failure, should be admitted to hospital for immediate initiation of treatment. However, it is unwise to reduce the blood pressure too rapidly since this may lead to cerebral, renal, retinal or myocardial infarction, and the blood pressure response to therapy must be carefully monitored, preferably in a high-dependency unit. In most cases, the aim is to reduce the diastolic blood pressure to 100–110 mmHg over 24–48 hours. This is usually achieved with oral medication, e.g. atenolol or amlodipine. The blood pressure can then be normalized over the next 2–3 days.

When rapid control of blood pressure is required (e.g. in an aortic dissection), the agent of choice is intravenous sodium nitroprusside. Alternatively, an infusion of labetalol can be used. The infusion dosage must be titrated against the blood pressure response. Fenoldopam, a selective peripheral dopamine receptor agonist, is as effective as nitroprusside.

Management of hypertension in pregnancy

Many antihypertensive agents are contraindicated in pregnancy. Mild hypertension can be treated with methyldopa, which has been established as being safe in pregnancy, or labetalol. Pre-eclamptic hypertension can be treated with the same agents, or nifedipine, although the only method for reversal of overt pre-eclampsia is delivery. More severe hypertension or eclampsia requires treatment with intravenous hydralazine and may even require termination of the pregnancy.

Prognosis

The prognosis from hypertension depends on a number of features:

- level of blood pressure
- presence of target-organ changes (retinal, renal, cardiac or vascular)
- coexisting risk factors for cardiovascular disease, such as hyperlipidaemia, diabetes, smoking, obesity, male sex
- age at presentation.

Several studies have confirmed that the treatment of hypertension, even mild hypertension, will reduce the risk not only of stroke but of coronary artery disease as well.

PERIPHERAL VASCULAR DISEASE

PERIPHERAL ARTERIAL DISEASE

Peripheral vascular disease (PVD) is commonly caused by atherosclerosis and usually affects the aorto-iliac or infra-inguinal arteries. It is present in 7% of middle-aged men and 4.5% of middle-aged women, but these patients are more likely to die of myocardial infarction or stroke than lose their leg.

Limb ischaemia may be classified as chronic or acute.

Chronic lower limb ischaemia

Symptoms

On exercise, patients complain of a severe cramp, usually in the calf, which resolves when they stop walking. They may be unable to continue walking with the pain and often the symptoms are worse walking uphill but never occur at rest. This is called intermittent claudication. Patients may experience similar pain in buttocks and thighs associated with male impotence, the 'Leriche syndrome'. Claudication can occur in both legs but is often worse in one leg.

Rest pain is defined as a severe unremitting pain in the foot, which stops a patient from sleeping. It is partially relieved by dangling the foot over the edge of the bed or standing on a cold floor. Patients with severe PVD or critical lower limb ischaemia may have ulceration or necrosis of the tissue (gangrene).

Signs

The lower limbs are cold with dry skin and lack of hair. Pulses may be diminished or absent. Ulceration may occur in association with dark discoloration of the toes or gangrene.

Risk factors

Common risk factors are:

- smoking
- diabetes
- hypercholesterolaemia
- hypertension.

Premature atherosclerosis in patients aged <45 years may be associated with thrombophilia and hyperhomocysteinaemia.

Differential diagnosis

Symptoms may be confused with those of:

- spinal canal claudication (but all pulses are present)
- osteoarthritis hip/knee (knee pain often at rest)
- peripheral neuropathy (associated with numbness and tingling)
- popliteal artery entrapment (young patients who may have normal pulses)
- venous claudication (bursting pain on walking with a previous history of a DVT)
- fibromuscular dysplasia
- 'Buerger's' disease (young males, heavy smokers).

Investigations

An estimation of the anatomical level of disease may be possible with the examination of pulses. The severity of disease is indicated by *ankle/brachial pressure index* (ABPI). This is a measurement of the cuff pressure at which blood flow is detectable by Doppler in the posterior tibial or anterior tibial arteries compared to the brachial artery (ankle/brachial pressure). Intermittent claudication is associated with an ABPI of 0.4–0.9. Values of <0.4 are associated with critical limb ischaemia. The sensitivity of the test may be improved by a fall in ABPI after exercise. If the arteries are heavily calcified and incompressible, i.e. in renal or diabetic disease, the ABPI will be falsely elevated. In these patients toe pressure values are more sensitive.

Diagnostic angiograms are now performed less commonly as *Doppler* and *duplex imaging* give an accurate anatomical assessment of the level and degree of disease. Angiograms are performed via a percutaneous arterial catheter and allow therapeutic interventions to be performed (Fig. 13.128).

Magnetic resonance and *CT angiography* are not routinely performed.

Management

Medical

All patients with peripheral vascular disease need aggressive risk factor management. Patients are encouraged to stop smoking and need smoking cessation advice. Patients with diabetes mellitus need regular chiropody care and diabetic management. Hypercholesterolaemia should be treated as this reduces disease progression. It has been shown by the Heart Protection Study that even the reduction of a normal cholesterol level reduces mortality from cardiovascular disease. Low-dose aspirin reduces the risk of myocardial infarction and stroke in patients with peripheral vascular disease. There are as yet no proven oral medications that are of benefit in patients with claudication. Supervised exercise programmes significantly improve walking distance and quality of life.

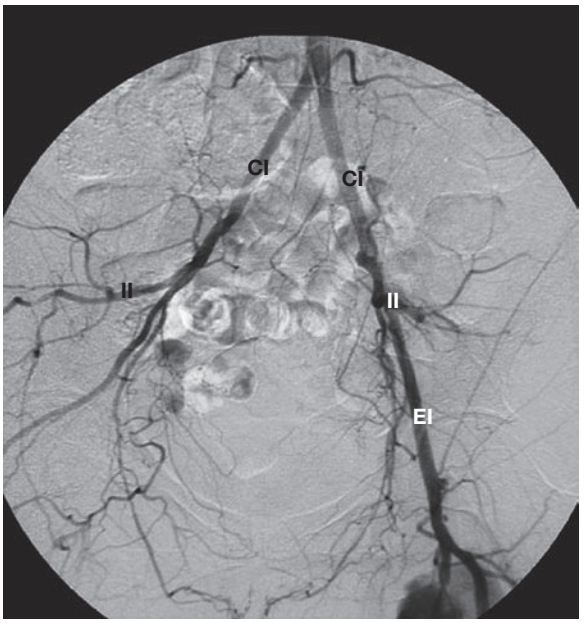


Fig. 13.128 **Angiogram showing occlusion of the right external iliac artery.** CI, common iliac; II, internal iliac; EI, external iliac.

Surgical/radiological

These are only considered in patients who have had their risk factors addressed and who feel that their lifestyle is disabled by their symptoms. For instance a person who can only walk 50 yards before claudicating but who also has severe breathing problems may never need lower limb intervention.

Percutaneous transluminal angioplasty is the first option and is carried out via a catheter inserted into the femoral artery. The long-term patency rates decrease as the angioplasty becomes more distal. The long-term results of angioplasty appear to be similar to those of a continued exercise programme. Arterial stents may be deployed in recurrent iliac disease, and drug-eluting stents allowing long-term patency are being used, e.g. paclitaxel.

Bypass procedures may be performed using Dacron, polytetrafluroethylene (PTFE) or autologous veins. Bypasses to distal vessels have poorer long-term patencies. Prosthetic grafts have equal patencies in above-knee bypasses but are inferior to veins below the knee.

In severe ischaemia with unreconstructable arterial disease an *amputation* may be necessary. An amputation may lead to loss of independence, with only 70% of below-knee and 30% of above-knee amputees achieving full mobility.

Acute lower limb ischaemia

Symptoms

Patients complain of the five Ps. They complain of pain, that the leg looks white (pallor), paraesthesia, paralysis and that it feels perishingly cold. The pain is unbearable and normally requires opioids for relief.

Signs

The limb is cold with mottling or marbling of the skin. Pulses are diminished or absent. The sensation and movement of the leg are reduced in severe ischaemia. Patients may develop a compartment syndrome with pain in the calf on compression.

Causes

Acute limb ischaemia (ALI) may occur because of embolic or thrombotic disease. ***Embolic disease*** is commonly due to cardiac thrombus and cardiac arrhythmias. Rheumatic fever is now an uncommon cause and the frequency of cardiac embolic ALI is also on the decline. Emboli may also occur secondary to aneurysm thrombus or thrombus on atherosclerotic plaques. Emboli from atrial myxomas are rare.

Acute limb ischaemia is now often due to ***thrombotic disease***. Acute thrombus may form on a chronic atherosclerotic stenosis in a patient who has previously reported symptoms of claudication. Thrombus may also form in normal vessels in patients who are hypercoagulable because of malignancy or thrombophilia defects. Prosthetic or venous grafts may also thrombose either de novo or secondary to a developing stenosis either in the graft or in the native vessels. Popliteal aneurysms may thrombose or embolize distally. ***Acute upper limb ischaemia*** may be caused by similar processes or occur secondary to external compression with a cervical rib/band.

Investigation and management

Investigations are similar to those described for chronic lower limb disease.

Medical

Management is dependent on the degree of ischaemia. Patients showing improvement may be treatable with heparin and appropriate treatment of the underlying cause. Patients with emboli following myocardial infarction or atrial fibrillation need long-term warfarin.

Surgical/radiological

Patients with mild to moderate ischaemic symptoms who have occluded a graft may need graft thrombolysis. Intra-arterial thrombolysis may reveal an underlying stenosis within a graft or native vessel that could be treated with angioplasty. Patients with an embolus may benefit from its surgical removal (embolectomy). A bypass graft may be required after occlusion of a popliteal aneurysm or acute-on-chronic lower limb arterial disease. When an ischaemic limb is revascularized, the sudden improvement in blood flow can cause reperfusion injury with release of toxic metabolites into the circulation. In muscle compartments the consequent oedema may lead to a 'compartment syndrome', which requires fasciotomies (release of the fascia to prevent muscle damage). An amputation may be warranted in unreconstructable or severe ischaemia. In patients dying from other causes, acute limb ischaemia may occur and intervention may then be inappropriate.

Aneurysmal disease

Aneurysms are classified as true and false. An aneurysm is defined if there is a permanent dilatation of the artery to twice the normal diameter. In true aneurysms the arterial wall forms the wall of the aneurysm. The arteries most frequently involved are the abdominal aorta, iliac, popliteal, femoral artery and thoracic aorta (in decreasing frequency).

In false aneurysms (pseudoaneurysms) the surrounding tissues form the wall of the aneurysm. False aneurysms can occur following femoral artery puncture. A haematoma is formed because of inadequate compression of the entry site and continued bleeding into the surrounding compressed soft tissue forms the wall of this aneurysm.

Abdominal aortic aneurysm

Abdominal aortic aneurysms (AAA) occur most commonly below the renal arteries (infrarenal). The incidence increases with age, being present in 5% of the population >60 years. They occur five times more frequently in men and in one in four male children of an affected individual. Aneurysms may occur secondary to atherosclerosis, infection (syphilis, *Escherichia coli, Salmonella*) and trauma, or may be genetic (Marfan's syndrome, Ehlers–Danlos syndrome).

Symptoms

Most aneurysms are asymptomatic and are found on routine abdominal examination, plain X-ray or during urological investigations. Rapid expansion or rupture of an AAA may cause *severe pain* (epigastric pain radiating to the back). A ruptured AAA causes *hypotension*, *tachycardia*, *profound anaemia* and *sudden death*. The symptoms of rupture may mimic renal colic, diverticulitis and severe lower abdominal or testicular pain. Gradual erosion of the vertebral bodies may cause non-specific back pain. The aneurysm may embolize distally. Inflammatory aneurysms can obstruct adjacent structures, e.g. ureter, duodenum and vena cava. Rarely patients with aneurysms can present with severe haematemesis secondary to an aortoduodenal fistula.

Signs

The aorta is retroperitoneal and in overweight patients there may be no overt signs. An aneurysm is suspected if a pulsatile, expansile abdominal mass is felt. The presence of an AAA should alert a clinician to the possibility of popliteal aneurysms. Patients may present with 'trash feet', dusky discoloration of the digits secondary to emboli from the aortic thrombus.

Investigations

An AAA is first assessed by *ultrasound*.

CT scan is more accurate and relates the anatomical relationship to the renal and visceral vessels.

Management

Like any operation, the management of an asymptomatic aneurysm depends on the balance of operative risk and conservative management. The UK Small Aneurysm Trial showed that patients with infrarenal AAA did best with an operation if the aneurysm was:

- ≥5.5 cm diameter
- expanding >1 cm/year
- symptomatic.

Medical

Patients with aneurysmal disease need careful control of hypertension, to stop smoking and to have lipid-lowering

medication. Patients with AAA <5.5 cm are followed up by regular ultrasound surveillance.

Repair of abdominal aortic aneurysm

Standard therapy is open surgical repair with insertion of a Dacron or Gore-Tex graft.

Endovascular stent

Endovascular stent insertion (via the femoral or iliac arteries) is a non-surgical approach to AAA repair. The EndoVascular Aneurysm Repair studies EVAR (stent versus open surgical repair) and EVAR 2 (stent versus medical therapy in patients unsuitable for open repair) investigated the role of endovascular stents in patients with AAA ≥5.5 cm on CT. In EVAR the 30-day mortality rate was 1.7% with stenting versus 4.7% with surgery (p = 0.009) but the long-term mortality rate was similar in both groups at 4 years. In EVAR 2 the 30-day mortality rate with stenting was 9%. Long-term mortality rate was similar in both stent and medical therapy groups.

A meta-analysis of three randomized control trials demonstrated a 30-day mortality rate of 2% for stent–graft repair versus 5% for open surgical repair; with reductions in ITU and in-hospital stay with stent–graft repair.

Laparoscopic surgery

An alternative to open-surgical repair or endovascular stenting is laparoscopic repair that is performed with hand-assisted laparoscopic surgery (HALS, requiring a midline mini-laparotomy) or by total laparoscopic surgery (TLS). In non-randomized controlled trials both methods were associated with reduced length of stay, although the operating times were longer.

Prognosis

After repair, patients with an AAA should return to normal activity within a few months.

Thoraco-abdominal aneurysm (TAA)

The ascending, arch or descending thoracic aorta may become aneurysmal. Ascending TAAs occur most commonly in patients with Marfan's syndrome or hypertension. Descending or arch TAAs occur secondary to atherosclerosis and are now rarely due to syphilis.

Symptoms

Most aneurysms are asymptomatic and are found on routine chest X-ray or cardiological investigation (Fig. 13.129). Rapid expansion may cause severe pain (chest pain radiating to the upper back) and rupture is associated with hypotension, tachycardia and death. Chest symptoms from expansion may include stridor (compressed bronchial tree), haemoptysis (aortobronchial fistula) and hoarseness (compression of the recurrent laryngeal nerve). Aorto-oesophageal fistula uncommonly causes haematemesis.

Investigations

- **CT scan** is used for assessment of a TAA.
- **Aortography** may be used to assess the position of the key branches in relation to the aneurysm.
- **Transoesophageal echocardiography** can be helpful by identifying an aortic dissection.

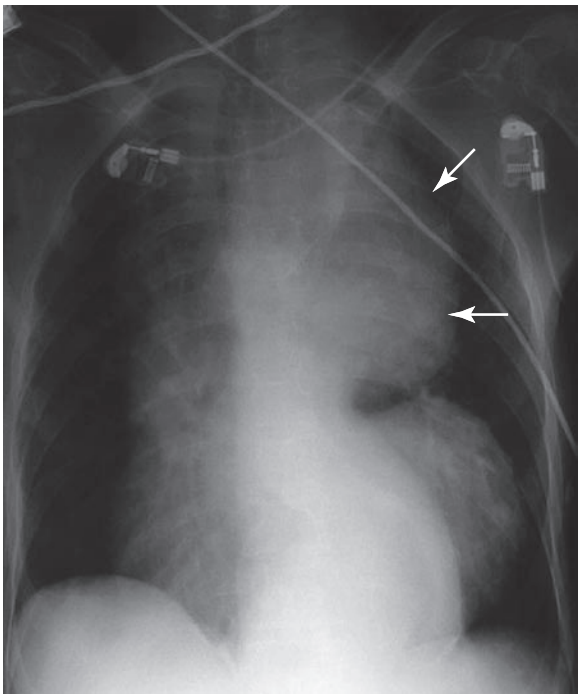


Fig. 13.129 **Plain PA chest X-ray demonstrating a markedly enlarged mediastinum due to an aneurysm of the aorta** (arrows).

Management

If the aneurysm is >6 cm then operative repair or stenting may be appropriate, but these can be technically difficult and carry a high risk of mortality and paraplegia. EVAR is at present the procedure of choice for isolated descending thoracic aneurysms.

Aortic dissection

Aortic dissection usually begins with a tear in the intima. Blood penetrates the diseased medial layer and then cleaves the intimal laminal plain leading to dissection. Thoracic aortic dissection may be classified into:

- *Type A*: involving the aortic arch and aortic valve proximal to the left subclavian artery origin.
- *Type B*: involving the descending thoracic aorta distal to the left subclavian artery origin.

Symptoms

Severe and central *chest pain* often radiates to the back and down the arms, mimicking myocardial infarction.

Signs

Patients may be shocked and may have neurological symptoms secondary to loss of blood supply to the spinal cord. They may develop renal failure, acute lower limb ischaemia or visceral ischaemia. Peripheral pulses may be absent.

Investigations

The mediastinum may be widened on *chest X-ray*, and *CT scan* or *transoesophageal echocardiography* will confirm the diagnosis (Fig. 13.98b).

Management

At least 50% of patients are hypertensive and they may require urgent antihypertensive medication. Type A dissec-

tions should undergo surgery (arch replacement) if fit enough, as medical management carries a high mortality. Type B dissections should be managed medically unless they develop complications.

Raynaud's phenomenon or Raynaud's disease

Raynaud's phenomenon consists of spasm of the digital arteries, usually precipitated by cold and relieved by heat. If there is no underlying cause, it is known as Raynaud's disease. This affects 5% of the population, mostly women. The disorder is usually bilateral with fingers affected more commonly than toes.

Symptoms

Vasoconstriction causes *skin pallor* followed by *cyanosis* due to sluggish blood flow, then redness secondary to hyperaemia. The duration of the attacks is variable but they can sometimes last for hours. Numbness, a burning sensation and severe pain occur as the fingers warm up. In chronic, severe disease tissue *infarction* and *digital loss* can occur.

Diagnosis

Primary Raynaud's disease needs to be differentiated from secondary treatable causes leading to Raynaud's phenomenon. These are the rheumatic autoimmune disorders such as systemic sclerosis (see p. 545). It can be associated with atherosclerosis or occupations that involve the use of vibrating tools. Ergot-containing drugs and beta-blockers, and smoking can aggravate symptoms.

Management

Patients should avoid cold provocation by wearing gloves and warm clothes, and stop smoking. Vasodilators can be prescribed but are often unacceptable as cerebral vasodilatation causes severe headaches. Sympathectomy or prostacyclin infusion can be helpful in severe disease.

Takayasu's disease

This is rare, except in Japan. It is known as the pulseless disease or aortic arch syndrome. It is of unknown aetiology and occurs in females. There is a vasculitis involving the aortic arch as well as other major arteries. There is also a systemic illness, with pain and tenderness over the affected arteries. Absent peripheral pulses and hypertension are common. Corticosteroids help the constitutional symptoms. Eventually heart failure and strokes may occur but most patients survive for at least 5 years. Treatment may require a surgical bypass to improve perfusion of the affected areas.

Thromboangiitis obliterans (Buerger's disease)

This disease, involving the small vessels of the lower limbs, occurs in young men who smoke. It is thought by some workers to be indistinguishable from atheromatous disease. However, pathologically there is inflammation of the arteries and sometimes veins that may indicate a separate disease entity. Clinically it presents with severe claudication and rest pain leading to gangrene. A thrombophlebitis is sometimes present. Treatment is as for all peripheral vascular disease, but patients must stop smoking.

Cardiovascular syphilis

This gives rise to:

- uncomplicated aortitis
- aortic aneurysms, usually in the ascending part
- aortic valvulitis with regurgitation
- stenosis of the coronary ostia.

The diagnosis is confirmed by serology. Treatment is with penicillin. Aneurysms and valvular disease are treated as necessary by the usual methods.

PERIPHERAL VENOUS DISEASE

Varicose veins

Varicose veins are a common problem, sometimes giving rise to pain. They are treated by injection or surgery.

Venous thrombosis

Thrombosis can occur in any vein, but the veins of the leg and the pelvis are the most common sites.

Superficial thrombophlebitis

This commonly involves the saphenous veins and is often associated with varicosities. Occasionally the axillary vein is involved, usually as a result of trauma. There is local superficial inflammation of the vein wall, with secondary thrombosis.

The clinical picture is of a painful, tender, cord-like structure with associated redness and swelling.

The condition usually responds to symptomatic treatment with rest, elevation of the limb and analgesics (e.g. non-steroidal anti-inflammatory drugs). Anticoagulants are not necessary, as embolism does not occur from superficial thrombophlebitis.

Deep vein thrombosis

A thrombus forms in the vein, and any inflammation of the vein wall is secondary to this.

Thrombosis commonly occurs after periods of immobilization, but it can occur in normal individuals for no obvious reasons. The precipitating factors are discussed on page 442.

A deep vein thrombosis in the legs occurs in 50% of patients after prostatectomy (without prophylactic heparin) or following a cerebral vascular event. In addition, 10% of patients with a myocardial infarct have a clinically detected deep vein thrombosis.

Thrombosis can occur in any vein of the leg or pelvis, but is particularly found in veins of the calf. It is often undetected; autopsy figures give an incidence of over 60% in hospitalized patients. Axillary vein thrombosis occasionally occurs, sometimes related to trauma, but usually for no obvious reason.

Clinical features

The individual may be asymptomatic, presenting with clinical features of pulmonary embolism (see p. 784).

A major presenting feature is *pain* in the calf, often with swelling, redness and *engorged superficial veins*. The

affected calf is often warmer and there may be ankle oedema. *Homan's sign* (pain in the calf on dorsiflexion of the foot) is often present, but is not diagnostic and occurs with all lesions of the calf.

Thrombosis in the iliofemoral region can present with severe pain, but there are often few physical signs apart from occasional swelling of the thigh and/or ankle oedema.

Complete occlusion, particularly of a large vein, can lead to a *cyanotic discoloration* of the limb and severe oedema, which can very rarely lead to venous gangrene.

Pulmonary embolism can occur with any deep vein thrombosis but is more frequent from an iliofemoral thrombosis and is rare with thrombosis confined to veins below the knee. In 20–30% of patients, spread of thrombosis can occur proximally without clinical evidence, so careful monitoring of the leg, usually by ultrasound, is required.

Investigations

Clinical diagnosis is unreliable but combined with D-dimer level it has a sensitivity of 80%. Confirmation of an iliofemoral thrombosis can usually be made with *B mode venous compression, ultrasonography* or *Doppler ultrasound* with a sensitivity and specificity over 90%.

Below-knee thromboses can be detected reliably only by *venography* with non-invasive techniques, ultrasound, fibrinogen scanning and impedance plethysmography, having a sensitivity of only 70%. A venogram is performed by injecting a vein in the foot with contrast, which will detect virtually all thrombi that are present.

Treatment

The main aim of therapy is to prevent pulmonary embolism, and all patients with thrombi above the knee must be anticoagulated. Anticoagulation of below-knee thrombi is now recommended for 6 weeks as 30% of patients will have an extension of the clot proximally. Bed rest is advised until the patient is fully anticoagulated. The patient should then be mobilized, with an elastic stocking giving graduated pressure over the leg.

Low-molecular-weight heparins (LMWH) (see p. 444) have replaced unfractionated heparin as they are more effective, they do not require monitoring and there is less risk of bleeding. DVTs are being treated at home with low-molecular-weight heparin. Warfarin is started immediately and the heparin stopped when the INR is in the target range. The duration of *warfarin* treatment is debatable – 3 months is the period usually recommended, but 4 weeks is long enough if a definite risk factor (e.g. bed rest) has been present. Recurrent DVTs need permanent anticoagulants. The target INR should be 2.5. Anticoagulants do not lyse the thrombus that is already present. Unfractionated heparin should only be used if LMWH is unavailable.

Thrombolytic therapy (see p. 443) is occasionally used for patients with a large iliofemoral thrombosis.

Prognosis

Destruction of the deep vein valves produces a clinically painful, swollen limb that is made worse by standing and is accompanied by oedema and sometimes venous eczema. It occurs in approximately half of the patients with clinically symptomatic deep vein thrombosis, and it means that elastic support stockings are then required for life.

Prevention

Subcutaneous low-molecular-weight heparin (see p. 444) should be given to patients with cardiac failure, a myocardial infarct or surgery to the leg or pelvis.

Early ambulation is indicated, as most thromboses occur within the first 72 hours following surgery. Leg exercises should be encouraged and patients should not sit in a chair with their legs immobilized on a stool. An elastic support stocking should be given to patients at high risk (e.g. those with a history of thrombosis or with obesity).

FURTHER READING

Greenhalgh RM, Brown LC, Kwong GP, et al.; EVAR trial participants. Comparison of endovascular aneurysm repair with open repair in patients with abdominal aortic aneurysm (EVAR trial 1), 30-day operative mortality results: randomised controlled trial. *Lancet* 2004; **364**: 843–848.

Johnston SL, Lock RJ, Gompels MM. Takayasu arteritis: a review. *Journal of Clinical Pathology* 2002; **55**: 481–486.

Kyrle PA, Eichinger S. Deep venous thrombosis. *Lancet* 2005; **365**: 1163–1174.

National Institute of Clinical Excellence. Stent-graft placement in abdominal aortic aneurysm. Interventional procedure guidance 163. www.nice.org.uk/IPG163distributionlist.

National Institute of Clinical Excellence. Laparoscopic repair of abdominal aortic aneurysm. Interventional procedure guidance 229. www.nice.org.uk/IPG229distributionlist.

Sakalihasan N, Limet R, Defawe OD. Abdominal aortic aneurysm. *Lancet* 2005; **365**: 1577–1589.

SIGNIFICANT WEBSITES

http://www.doh.gov.uk/nsf/coronarych4.htm
UK National Service Framework for Coronary Heart Disease (2000)

http://www.americanheart.org/
American Heart Association

http://www.erc.edu/
European Resuscitation Council

http://www.resus.org.uk
UK Resuscitation Council

http://homepages.enterprise.net/djenkins/ecghome.html
ECG tracings library

http://www.achd-library.com/
Nevil Thomas Adult Congenital Heart Library

14 Respiratory disease

Structure of the respiratory system 811
Physiology of the respiratory system 814
Defence mechanisms of the respiratory tract 818
Symptoms 819
Examination of the respiratory system 820
Investigation of respiratory disease 821
Smoking and air pollution 829
Smoking 829
Air pollution and epidemiology 830
Diseases of the upper respiratory tract 831
Diseases of the lower respiratory tract 835
Asthma 846
Pneumonia 857
Types of pneumonia 859
General management of pneumonia 862
Complications of pneumonia 863
Tuberculosis 863
Diffuse diseases of the lung parenchyma 868
Granulomatous lung disease 868
Granulomatous lung disease with vasculitis 870
Idiopathic interstitial pneumonias (IIP) 872
Other types of diffuse lung disease 874
Pulmonary infiltration with eosinophilia 874
Extrinsic allergic alveolitis 876
Occupational lung disease 878
Lung cysts 880
Tumours of the respiratory tract 880
Malignant tumours 880
Screening for lung cancer 884
Benign tumours 884
Disorders of the chest wall and pleura 884
Disorders of the diaphragm 886
Mediastinal lesions 887

The main role of the respiratory system is to extract oxygen from the external environment and dispose of waste gases, principally carbon dioxide. This requires the lungs to function as efficient bellows, bringing in fresh air and delivering it to the alveoli, and expelling used air at an appropriate rate. Gas exchange is achieved by exposing thin-walled capillaries to the alveolar gas and matching ventilation to blood flow through the pulmonary capillary bed. In doing this, the lungs expose a large area of tissue, which can be damaged by dusts, gases and infective agents. Host defence is therefore a key priority for the lung and is achieved by a combination of structural and immunological defences.

Ask the authors
www.studentconsult.com

STRUCTURE OF THE RESPIRATORY SYSTEM

The nose, pharynx and larynx

See Chapter 20 (pp. 1080 and 1082).

The trachea, bronchi and bronchioles

The trachea is 10–12 cm in length. It lies slightly to the right of the midline and divides at the carina into right and left main bronchi. The carina lies under the junction of the manubrium sterni and the second right costal cartilage. The right main bronchus is more vertical than the left and, hence, inhaled material is more likely to end up in the right lung.

The right main bronchus divides into the upper lobe bronchus and the intermediate bronchus, which further subdivides into the middle and lower lobe bronchi. On the left the main bronchus divides into upper and lower lobe bronchi only. Each lobar bronchus further divides into segmental and subsegmental bronchi. There are about 25 divisions in all between the trachea and the alveoli.

The first seven divisions are bronchi that have:

- walls consisting of cartilage and smooth muscle
- epithelial lining with cilia and goblet cells
- submucosal mucus-secreting glands
- endocrine cells – Kulchitsky or APUD (amine precursor and uptake decarboxylation) containing 5-hydroxytryptamine.

The next 16–18 divisions are bronchioles that have:

- no cartilage and a muscular layer that progressively becomes thinner
- a single layer of ciliated cells but very few goblet cells
- granulated Clara cells that produce a surfactant-like substance.

The ciliated epithelium is a key defence mechanism. Each cell bears approximately 200 cilia beating at 1000 beats per minute in organized waves of contraction. Each cilium consists of nine peripheral parts and two inner longitudinal fibrils in a cytoplasmic matrix (Fig. 14.1). Nexin links join the

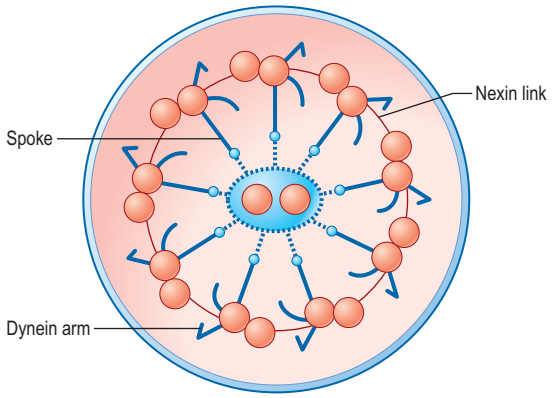


Fig. 14.1 **Cross-section of a cilium.** Nine outer microtubular doublets and two central single microtubules are linked by spokes, nexin links and dynein arms.

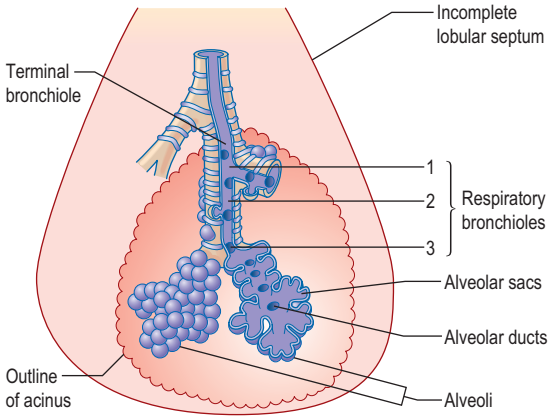


Fig. 14.2 **Branches of a terminal bronchiole ending in the alveolar sacs.**

peripheral pairs. Dynein arms consisting of ATPase protein project towards the adjacent pairs. Bending of the cilia results from a sliding movement between adjacent fibrils powered by an ATP-dependent shearing force developed by the dynein arms. Absence of dynein arms leads to immotile cilia. Mucus, which contains macrophages, cell debris, inhaled particles and bacteria, is moved by the cilia towards the larynx at about 1.5 cm/min (the 'mucociliary escalator', see below).

The bronchioles finally divide within the acinus into smaller respiratory bronchioles that have alveoli arising from the surface (Fig. 14.2). Each respiratory bronchiole supplies approximately 200 alveoli via alveolar ducts. The term 'small airways' refers to bronchioles of less than 2 mm; the average lung contains about 30 000 of these.

The alveoli

There are approximately 300 million alveoli in each lung. Their total surface area is 40–80 m^2. The epithelial lining consists mainly of *type I pneumocytes* (Fig. 14.3). These cells have an extremely thin layer of cytoplasm, which only offers a thin barrier to gas exchange. *Type I cells* are connected to each other by tight junctions that limit the movements of fluid in and out of the alveoli. Alveoli are not completely airtight – many have holes in the alveolar wall allowing communication between alveoli of adjoining lobules (pores of Kohn).

Type II pneumocytes are slightly more numerous than type I cells but cover less of the epithelial lining. They are

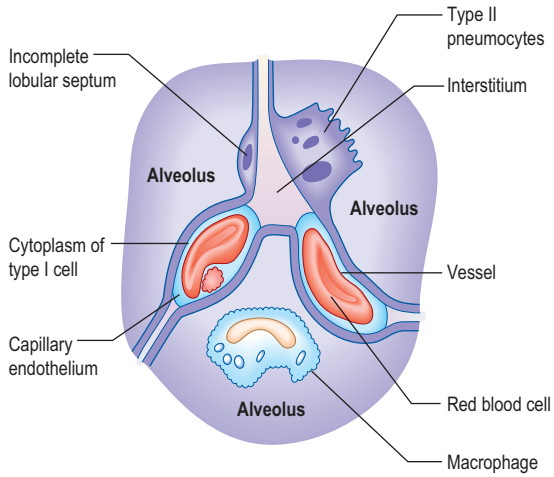


Fig. 14.3 **The structure of alveoli, showing the pneumocytes and capillaries.**

found generally in the borders of the alveolus and contain distinctive lamellar vacuoles, which are the source of surfactant. Type I pneumocytes are derived from type II cells. Large alveolar macrophages are present within the alveoli and assist in defending the lung.

The lungs

The lungs are separated into lobes by invaginations of the pleura, which are often incomplete. The right lung has three lobes, whereas the left lung has two. The positions of the oblique fissures and the right horizontal fissure are shown in Figure 14.4. The upper lobe lies mainly in front of the lower lobe and therefore physical signs on the right side in the front of the chest are due to lesions of the upper lobe or the middle lobe. Because of the contrast in density between healthy and diseased lung, plain radiography enables accurate localization of disease processes, especially if postero-anterior (PA) and lateral views are taken (Fig. 14.5).

Each lobe is further subdivided into bronchopulmonary segments by fibrous septa that extend inwards from the pleural surface. Each segment receives its own segmental bronchus.

The bronchopulmonary segment is further divided into individual lobules approximately 1 cm in diameter and generally pyramidal in shape, the apex lying towards the bronchiole that supplies it. Within each lobule a terminal bronchus supplies an acinus, and within this structure further divisions of the bronchioles eventually give rise to the alveoli.

The pleura

The pleura is a layer of connective tissue covered by a simple squamous epithelium. The visceral pleura covers the surface of the lung, lines the interlobar fissures, and is continuous at the hilum with the parietal pleura, which lines the inside of the hemithorax. At the hilum the visceral pleura continues alongside the branching bronchial tree for some distance before reflecting back to join the parietal pleura. In health, the pleurae are in apposition apart from a small quantity of lubricating fluid.

The diaphragm

The diaphragm is covered by parietal pleura above and peritoneum below. Its muscle fibres arise from the lower ribs and

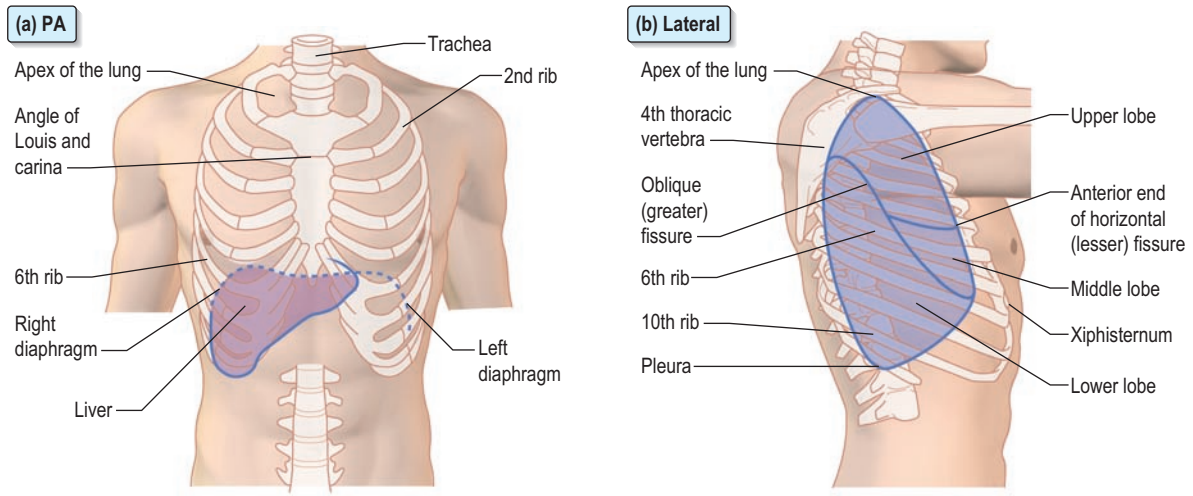


Fig. 14.4 **Surface anatomy of the chest. (a)** PA. **(b)** Lateral.

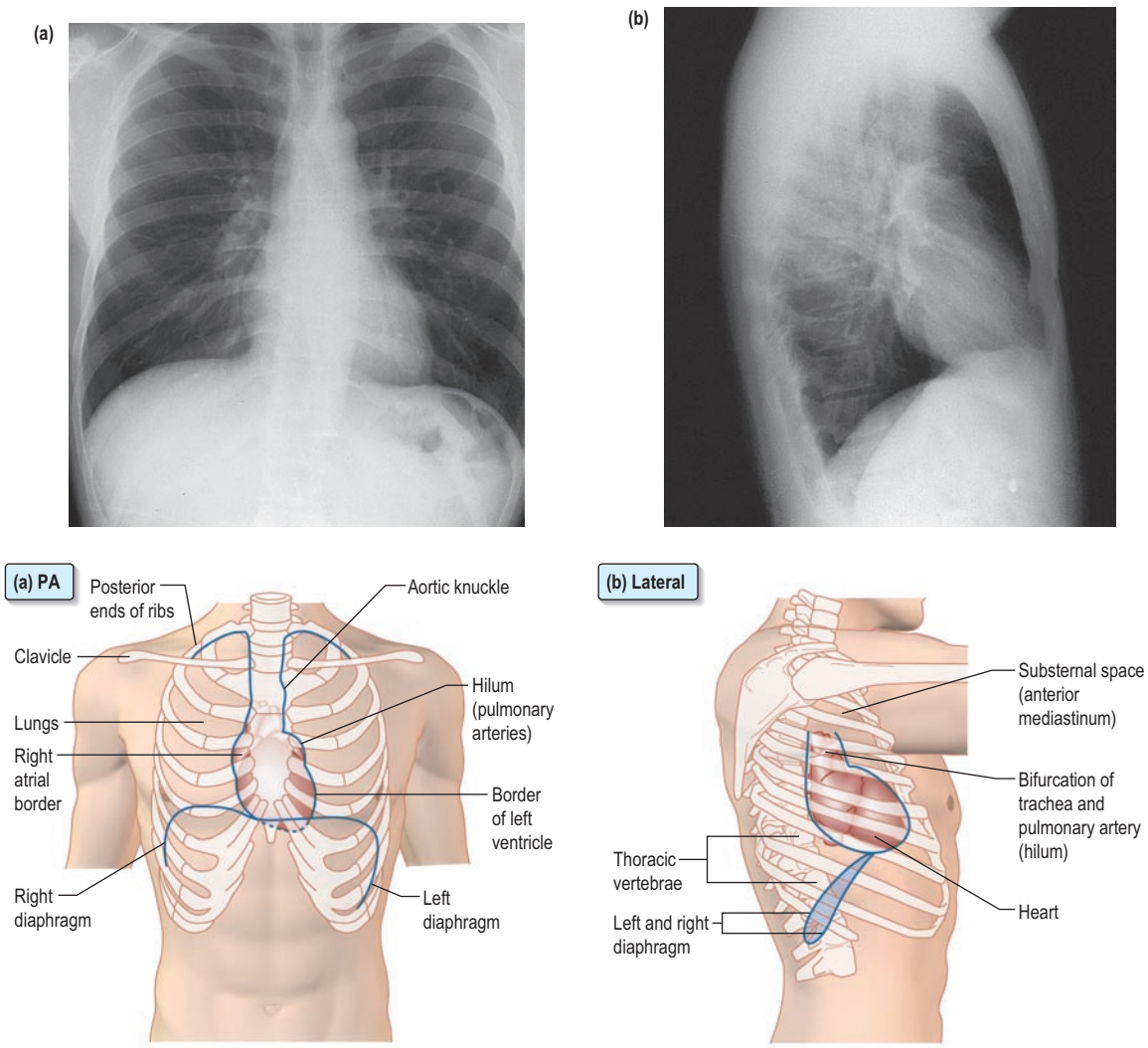


Fig. 14.5 **Chest X-rays. (a)** PA. **(b)** Lateral.

insert into the central tendon. Motor and sensory nerve fibres go separately to each half of the diaphragm via the phrenic nerves. Fifty per cent of the muscle fibres are of the slow-twitch type with a low glycolytic capacity; they are relatively resistant to fatigue.

Pulmonary vasculature and lymphatics

The lung has a dual blood supply, receiving deoxygenated blood from the right ventricle via the pulmonary artery and oxygenated blood via the bronchial circulation.

The pulmonary artery divides to accompany the bronchi. The arterioles accompanying the respiratory bronchioles are thin-walled and contain little smooth muscle. The pulmonary venules drain laterally to the periphery of the lobules, pass centrally in the interlobular and intersegmental septa, and eventually join to form the four main pulmonary veins.

The bronchial circulation arises from the descending aorta. These bronchial arteries supply tissues down to the level of the respiratory bronchiole. The bronchial veins drain into the pulmonary vein, forming part of the normal physiological shunt.

Lymphatic channels lie in the interstitial space between the alveolar cells and the capillary endothelium of the pulmonary arterioles.

The tracheobronchial lymph nodes are arranged in five main groups: pulmonary, bronchopulmonary, subcarinal, superior tracheobronchial and paratracheal. For all practical purposes these form a continuous network of nodes from the lung substance up to the trachea.

Nerve supply to the lung

The innervation of the lung remains incompletely understood. Parasympathetic and sympathetic fibres (from the vagus and sympathetic chain respectively) accompany the pulmonary arteries and the airways. Airway smooth muscle is innervated by vagal afferents, postganglionic muscarinic vagal efferents and vagally derived non-adrenergic non-cholinergic (NANC) fibres. Neurotransmitters (peptides and purines) are also involved. These include substance P, neurokinins A and B, calcitonin gene-related peptide, vasoactive intestinal polypeptide and a range of adenine and guanine phosphates. Three muscarinic receptor subtypes have been identified: M_1 receptors on parasympathetic ganglia, a smaller number of M_2 receptors on muscarinic nerve terminals, and M_3 receptors on airway smooth muscle. The parietal pleura is innervated from intercostal and phrenic nerves but the visceral pleura has no innervation.

FURTHER READING

Gibson J, Geddes D, Costabel U et al. *Respiratory Medicine*, 3rd edn, Vols 1 and 2. London: WB Saunders, 2002.

PHYSIOLOGY OF THE RESPIRATORY SYSTEM

The nose

The major functions of nasal breathing are:

- to heat and moisten the air
- to remove particulate matter.

About 10 000 L of air are inhaled daily. The relatively low flow rates and turbulence of inspired air are ideal for particle deposition, and few particles greater than 10 microns pass through the nose. Particles deposited on the nasal mucosa are removed within 15 minutes, compared with 60–120 days for particles that reach the alveoli. Nasal secretion contains IgA antibodies, lysozyme and interferons. In addition, the cilia of the nasal epithelium move the mucous gel layer rapidly back to the oropharynx where it is swallowed. Bacteria have little chance of settling in the nose. Mucociliary protection is less effective against viral infections because viruses bind to receptors on epithelial cells. The majority of rhinoviruses bind to an adhesion molecule, intercellular adhesion molecule 1 (ICAM-1), which is shared by neutrophils and eosinophils. Many noxious gases, such as SO_2, are almost completely removed by nasal breathing.

Breathing

Lung ventilation can be considered in two parts:

- the mechanical process of inspiration and expiration
- the control of respiration to a level appropriate for the metabolic needs.

Mechanical process

The lungs have an inherent elastic property that causes them to tend to collapse away from the thoracic wall, generating a negative pressure within the pleural space. The strength of this retractive force relates to the volume of the lung: at higher lung volumes the lung is stretched more, and a greater negative intrapleural pressure is generated. *Lung compliance* is a measure of the relationship between this retractive force and lung volume. At the end of a quiet expiration, the retractive force exerted by the lungs is balanced by the tendency of the thoracic wall to spring outwards. At this point, respiratory muscles are resting. The volume of air remaining in the lung after a quiet expiration is called the *functional residual capacity (FRC).*

Inspiration from FRC is an active process: a negative pressure is created by descent of the diaphragm and movement of the ribs upwards and outwards through contraction of the intercostal muscles. During tidal breathing in healthy individuals, inspiration is almost entirely due to contraction of the diaphragm. More vigorous inspiration may require the use of accessory muscles of ventilation, such as the sternomastoid and scalene muscles. Respiratory muscles are similar to other skeletal muscles but are less prone to fatigue. However, inspiratory muscle fatigue contributes to respiratory failure in patients with severe chronic airflow limitation and in those with primary neurological and muscle disorders.

At rest or during low-level exercise, expiration is passive and results from the natural tendency of the lung to collapse.

Forced expiration involves activation of accessory muscles, chiefly those of the abdominal wall, which help to push up the diaphragm.

The control of respiration

Coordinated respiratory movements result from rhythmical discharges arising in an anatomically ill-defined group of interconnected neurones in the reticular substance of the brainstem, known as the respiratory centre. Motor discharges from the respiratory centre travel via the phrenic and intercostal nerves to the respiratory musculature.

In healthy individuals, the main driver for respiration is the arterial pH, which is closely related to the pressure of carbon dioxide in arterial blood. Oxygen levels in arterial blood are usually above the level which triggers respiratory drive. In a typical normal adult at rest:

- The pulmonary blood flow of 5 L/min carries 11 mmol/min (250 mL/min) of oxygen from the lungs to the tissues.
- Ventilation at about 6 L/min carries 9 mmol/min (200 mL/min) of carbon dioxide out of the body.
- The normal pressure of oxygen in arterial blood (P_aO_2) is between 11 and 13 kPa.
- The normal pressure of carbon dioxide in arterial blood (P_aCO_2) is 4.8–6.0 kPa.

Ventilation is controlled by a combination of neurogenic and chemical factors (Fig. 14.6).

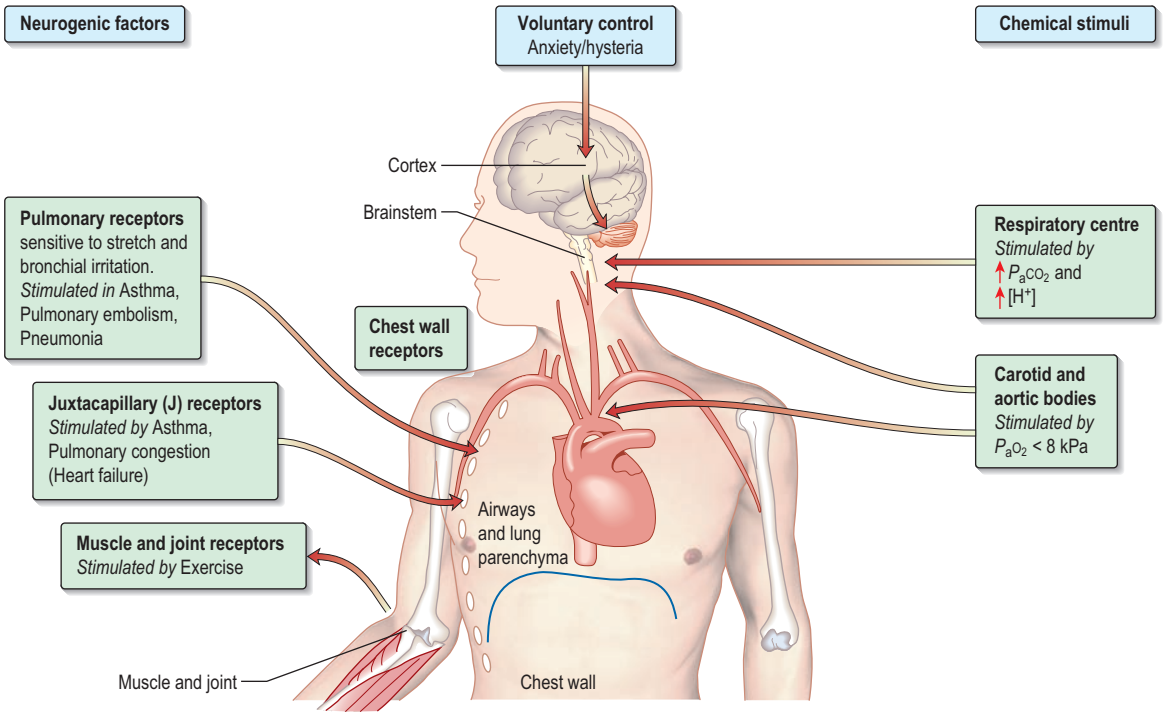


Fig. 14.6 **Chemical and neurogenic factors in the control of ventilation.** The strongest stimulant to ventilation is a rise in P_aCO_2 which increases [H^+] in CSF. Sensitivity to this may be lost in COPD. In these patients hypoxaemia is the chief stimulus to respiratory drive; oxygen treatment may therefore reduce respiratory drive and lead to a further rise in P_aCO_2. An increase in [H^+] due to metabolic acidosis as in diabetic ketoacidosis will increase ventilation with a fall in P_aCO_2 causing deep sighing (Kussmaul) respiration. The respiratory centre is depressed by severe hypoxaemia and sedatives (e.g. opiates) and stimulated by doxapram, large doses of aspirin and pyrexia. COPD, chronic obstructive pulmonary disease. Derived from Manning HL, Schwartzstein RM. *New England Journal of Medicine* 1995; **333**: 1547–1553.

Breathlessness on physical exertion is normal and not considered a symptom unless the level of exertion is very light, such as when walking slowly. Surveys of healthy western populations reveal that over 20% of the general population report themselves as breathless on relatively minor exertion. Although breathlessness is a very common symptom, the sensory and neural mechanisms underlying it remain obscure. The sensation of breathlessness is derived from at least three sources:

- *Changes in lung volume*. These are sensed by receptors in thoracic wall muscles signalling changes in their length.
- *Tension developed by contracting muscles*. This is sensed by Golgi tendon organs.
- *Central perception of the sense of effort*.

The airways of the lungs

From the trachea to the periphery, the airways become smaller in size (although greater in number). The cross-sectional area available for airflow increases as the total number of airways increases. The flow of air is greatest in the trachea and slows progressively towards the periphery (as the velocity of airflow depends on the ratio of flow to cross-sectional area). In the terminal airways, gas flow occurs solely by diffusion. The resistance to airflow is very low (0.1–0.2 kPa/L in a normal tracheobronchial tree), steadily increasing from the small to the large airways.

Airways expand as lung volume is increased, and at full inspiration (*total lung capacity*, TLC) they are 30–40% larger in calibre than at full expiration (*residual volume*, RV). In chronic obstructive pulmonary disease (COPD) the small airways are narrowed and this can be partially compensated by breathing at a larger lung volume.

Control of airway tone

Bronchomotor tone is maintained by vagal efferent nerves and, even in a normal subject, is reduced by atropine or β-adrenoceptor agonists. Adrenoceptors on the surface of bronchial muscles respond to circulating catecholamines; there is no direct sympathetic innervation. Airway tone shows a *circadian rhythm*, which is greatest at 04.00 and lowest in the mid-afternoon. Tone can be increased transiently by inhaled stimuli acting on epithelial nerve endings, which trigger reflex bronchoconstriction via the vagus. These stimuli include cigarette smoke, solvents, inert dust and cold air. Airway responsiveness to these stimuli increases following respiratory tract infections even in healthy subjects. In asthma, the airways are very irritable and as the circadian rhythm remains the same, asthmatic symptoms are usually worse in the early morning.

Airflow

Movement of air through the airways results from a difference between the pressure in the alveoli and the atmospheric pressure; alveolar pressure is negative in inspiration and positive in expiration. During quiet breathing the pleural pressure is negative throughout the breathing cycle. With vigorous expiratory efforts (e.g. cough), the pleural pressure becomes positive (up to 10 kPa). This compresses the central

airways, but the smaller airways do not close off because the driving pressure for expiratory flow (alveolar pressure) is also increased.

Alveolar pressure (P_{ALV}) is equal to the pleural pressure (P_{PL}) plus the elastic recoil pressure (P_{EL}) of the lung.

When there is no airflow (i.e. during a pause in breathing) the tendency of the lungs to collapse (the positive recoil pressure) is exactly balanced by an equivalent negative pleural pressure.

As air flows from the alveoli towards the mouth there is a gradual loss of pressure owing to flow resistance (Fig. 14.7a).

In forced expiration, as mentioned above, the driving pressure raises both the alveolar pressure and the intrapleural pressure. Between the alveolus and the mouth, there is a point (C in Fig. 14.7b) where the airway pressure equals the intrapleural pressure, and the airway collapses. However, this collapse is temporary, as the transient occlusion of the airway results in an increase in pressure behind it (i.e. upstream) and this raises the intra-airway pressure so that the airways open and flow is restored. The airways thus tend to vibrate at this point of 'dynamic collapse'.

As lung volume decreases during expiration, the elastic recoil pressure of the lungs decreases and the 'collapse point' moves upstream (i.e. towards the smaller airways – see Fig. 14.7c). Where there is pathological loss of recoil pressure (as in chronic obstructive pulmonary disease, COPD), the 'collapse point' is located even further upstream and causes expiratory flow limitation. The measurement of the forced expiratory volume in 1 second (FEV_1) is a useful clinical index of this phenomenon. To compensate, these patients often 'purse their lips' in order to increase airway pressure so that their peripheral airways do not collapse. On inspiration, the intrapleural pressure is always less than the intraluminal pressure within the intrathoracic airways, so there is no limitation to airflow with increasing effort. Inspiratory flow is limited only by the power of the inspiratory muscles.

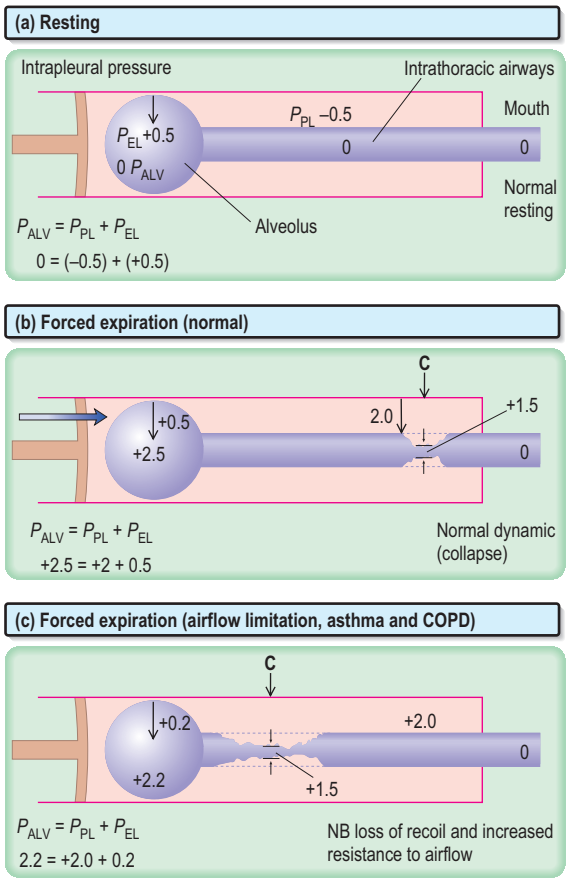


Fig. 14.7 **Diagrams showing ventilatory forces. (a)** During resting at functional residual capacity. **(b)** During forced expiration in normal subjects. **(c)** During forced expiration in a patient with COPD. The respiratory system is represented as a piston with a single alveolus and the collapsible part of the airways within the piston (see text). C, collapse point; P_{ALV}, alveolar pressure; P_{EL}, elastic recoil pressure; P_{PL}, pleural pressure.

Flow–volume loops

The relationship between maximal flow rates on expiration and inspiration is demonstrated by the maximal flow–volume (MFV) loop. Figure 14.8a shows this in a normal subject.

In subjects with healthy lungs, effects of flow limitation will not be apparent, since maximal flow rates are rarely

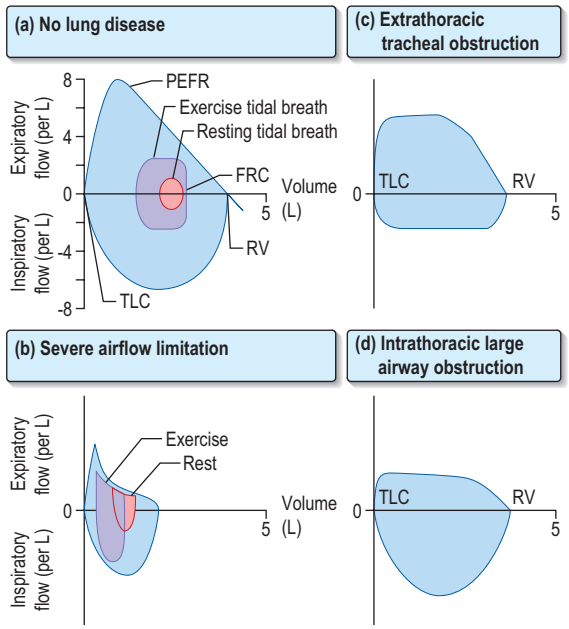


Fig. 14.8 **(a and b) Maximal flow–volume loops, showing the relationship between maximal flow rates on expiration and inspiration. (a)** In a normal subject. **(b)** In a patient with severe airflow limitation. Flow–volume loops during tidal breathing at rest (starting from the functional residual capacity (FRC)) and during exercise are also shown. The highest flow rates are achieved when forced expiration begins at total lung capacity (TLC) and represent the peak expiratory flow rate (PEFR). As air is blown out of the lung, so the flow rate decreases until no more air can be forced out, a point known as the residual volume (RV). Because inspiratory airflow is only dependent on effort, the shape of the maximal inspiratory flow–volume loop is quite different, and inspiratory flow remains at a high rate throughout the manoeuvre. **(c and d) Flow–volume loops of patients with large airway (tracheal) obstruction, showing plateauing of maximal expiratory flow. (c)** Extrathoracic tracheal obstruction with a proportionally greater reduction of maximal inspiratory (as opposed to expiratory) flow rate. **(d)** Intrathoracic large airway obstruction; the expiratory plateau is more pronounced and inspiratory flow rate is less reduced than in (c). In severe airflow limitation the ventilatory demands of exercise cannot be met (cf. a and b), greatly reducing effort tolerance.

achieved even during vigorous exercise. However, in patients with severe COPD, limitation of expiratory flow occurs even during tidal breathing at rest (see Fig. 14.8b). To increase ventilation these patients have to breathe at higher lung volumes and allow more time for expiration, both of which reduce the tendency for airway collapse. To compensate they increase flow rates during inspiration, where there is relatively less flow limitation.

The volume that can be forced in from the residual volume in 1 second (FIV_1) will always be greater than that which can be forced out from TLC in 1 second (FEV_1). Thus, the ratio of FEV_1 to FIV_1 is below 1. The only exception to this occurs when there is significant obstruction to the airways outside the thorax, such as tracheal tumour or retrosternal goitre. Expiratory airway narrowing is prevented by tracheal resistance and expiratory airflow becomes more effort-dependent. During forced inspiration this same resistance causes such negative intraluminal pressure that the trachea is compressed by the surrounding atmospheric pressure. Inspiratory flow thus becomes less effort-dependent, and the ratio of FEV_1 to FIV_1 exceeds 1. This phenomenon, and the characteristic flow–volume loop, is diagnostic of extrathoracic airways obstruction (Fig. 14.8c).

Ventilation and perfusion relationships

For optimum gas exchange there must be a match between ventilation of the alveoli ($\dot{V}_A$) and their perfusion ($\dot{Q}$). However, in reality there is variation in the $\dot{V}_A/\dot{Q}$ ratio in both normal and diseased lungs (Fig. 14.9). In the normal lung both ventilation and perfusion are greater at the bases than at the apices, but the gradient for perfusion is steeper, so the net effect is that ventilation exceeds perfusion towards the apices, while perfusion exceeds ventilation at the bases. Other causes of $\dot{V}_A/\dot{Q}$ mismatch include direct shunting of blood through the lung without passing through alveoli (e.g. the bronchial circulation) and areas of lung that receive no blood (e.g. anatomical deadspace, bullae and areas of underperfusion during acceleration and deceleration, e.g. in aircraft and high performance cars).

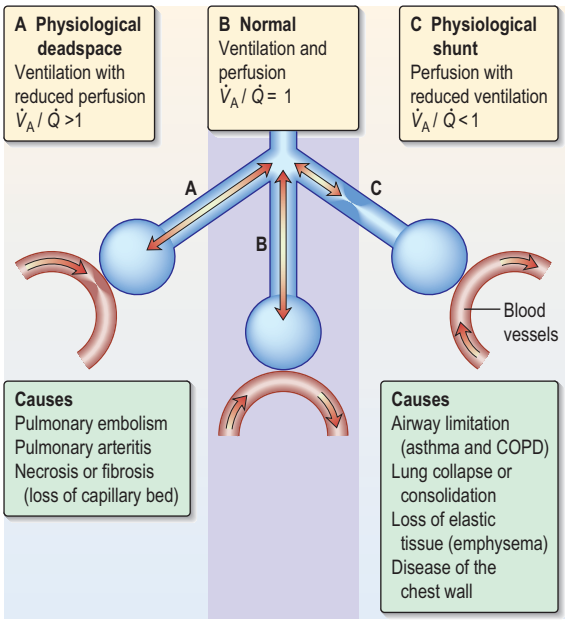


Fig. 14.9 **Relationships between ventilation and perfusion: a schematic diagram showing the alveolar–capillary interface.** The centre (B) shows normal ventilation and perfusion. On the left (A) there is a block in perfusion (physiological deadspace), while on the right (C) there is reduced ventilation (physiological shunting).

An increased physiological shunt results in arterial hypoxaemia since it is not possible to compensate for some of the blood being underoxygenated by increasing ventilation elsewhere. An increased physiological deadspace just increases the work of breathing and has less impact on blood gases since the normally perfused alveoli are well ventilated. In more advanced disease this compensation cannot occur, leading to increased alveolar and arterial P_{CO_2} (P_aCO_2), together with hypoxaemia which cannot be compensated by increasing ventilation.

Hypoxaemia occurs more readily than hypercapnia because of the different ways in which oxygen and carbon dioxide are carried in the blood. Carbon dioxide can be considered to be in simple solution in the plasma, the volume carried being proportional to the partial pressure. Oxygen is carried in chemical combination with haemoglobin in the red blood cells, with a non-linear relationship between the volume carried and the partial pressure (see Fig. 15.5, p. 892). Alveolar hyperventilation reduces the alveolar P_{CO_2} (P_aCO_2) and diffusion leads to a proportional fall in the carbon dioxide content of the blood. However, as the haemoglobin is already saturated with oxygen, there is no significant increase in the blood oxygen content as a result of increasing the alveolar P_{O_2} through hyperventilation. The hypoxaemia of even a small amount of physiological shunting cannot therefore be compensated for by hyperventilation.

In individuals who have mild degrees of $\dot{V}_A/\dot{Q}$ mismatch, the P_aO_2 and P_aCO_2 may still be normal. Increasing the requirements for gas exchange by exercise will widen the $\dot{V}_A/\dot{Q}$ mismatch and the P_aO_2 will fall. $\dot{V}_A/\dot{Q}$ mismatch is by far the most common cause of arterial hypoxaemia.

Alveolar stability

Pulmonary alveoli are essentially hollow spheres. Surface tension acting at the curved internal surface tends to cause the sphere to decrease in size. The surface tension within the alveoli would make the lungs extremely difficult to distend were it not for the presence of surfactant, an insoluble lipoprotein largely consisting of dipalmitoyl lecithin, which forms a thin monomolecular layer at the air–fluid interface. Surfactant is secreted by type II pneumocytes within the alveolus and reduces surface tension so that alveoli remain stable.

Fluid surfaces covered with surfactant exhibit a phenomenon known as hysteresis; that is, the surface-tension-lowering effect of the surfactant can be improved by a transient increase in the size of the surface area of the alveoli. During quiet breathing, small areas of the lung undergo collapse, but it is possible to re-expand these rapidly by a deep breath; hence the importance of sighs or deep breaths as a feature of normal breathing. Failure of this mechanism – e.g. in patients with fractured ribs – gives rise to patchy basal lung collapse. Surfactant levels may be reduced in a number of diseases that cause damage to the lung (e.g. pneumonia). Lack of surfactant plays a central role in the respiratory distress syndrome of the newborn. Severe reduction in perfu-

FURTHER READING

Booth S, Dudgeon D. *Dyspnoea in Advanced Disease. A Guide to Clinical Management.* Oxford: Oxford University Press, 2005.

sion of the lung causes impairment of surfactant activity and may explain the characteristic areas of collapse associated with pulmonary embolism.

DEFENCE MECHANISMS OF THE RESPIRATORY TRACT

Pulmonary disease often results from a failure of the normal host defence mechanisms of the healthy lung (Fig. 14.10). These can be divided into physical and physiological mechanisms, and humoral and cellular mechanisms.

Physical and physiological mechanisms

Humidification

This prevents dehydration of the epithelium.

Particle removal

Over 90% of particles greater than 10 μm diameter are removed in the nostril or nasopharynx. This includes most pollen grains, which are typically >20 μm in diameter. Particles between 5 and 10 μm become impacted in the carina. Particles smaller than 1 μm tend to remain airborne, thus the particles capable of reaching the deep lung are confined to the 1–5 μm range.

Particle expulsion

This is effected by coughing, sneezing or gagging.

Respiratory tract secretions (Fig. 14.10)

The mucus of the respiratory tract is a gelatinous substance consisting chiefly of acid and neutral polysaccharides. The mucus is a thick gel that is relatively impermeable to water. It floats on a liquid or sol layer that is present around the cilia of the epithelial cells. The gel layer is secreted from goblet cells and mucous glands as distinct globules that coalesce increasingly in the central airways to form a more or less continuous mucus blanket. Under normal conditions the tips of the cilia are in contact with the undersurface of the gel phase and coordinate their movement to push the mucus blanket towards the mouth. Whilst it only takes 30–60 minutes for mucus to be cleared from the large bronchi, it can be several days before mucus is cleared from respiratory bronchioles. One of the major long-term effects of cigarette smoking is a reduction in mucociliary transport. This contributes to recurrent infection and in the larger airways it prolongs contact with carcinogens. Air pollutants, local and general anaesthetics and products of bacterial and viral infection also reduce mucociliary clearance.

Congenital defects in mucociliary transport lead to recurrent infections and eventually to bronchiectasis. For example, in the 'immotile cilia' syndrome there is an absence of the dynein arms in the cilia themselves, while in cystic fibrosis there is ciliary dyskinesia and abnormally thick mucus.

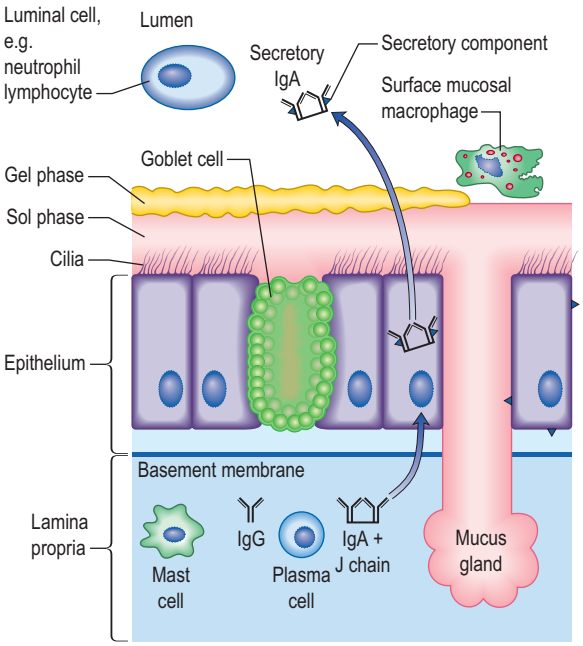


Fig. 14.10 **Defence mechanisms present at the epithelial surface.**

Humoral and cellular mechanisms

Non-specific soluble factors

- *α_1-Antitrypsin* (α_1-antiprotease, see p. 358) in lung secretions is derived from plasma. It inhibits chymotrypsin and trypsin and neutralizes proteases including neutrophil elastase.
- *Antioxidant defences* include enzymes such as superoxide dismutase and low-molecular-weight antioxidant molecules (ascorbate, urate) in the epithelial lining fluid. In addition, lung cells are protected by an extensive range of intracellular defences, especially members of the glutathione S-transferase (GST) superfamily.
- *Lysozyme* is an enzyme found in granulocytes that has bactericidal properties.
- *Lactoferrin* is synthesized from epithelial cells and neutrophil granulocytes and has bactericidal properties.
- *Interferons* are produced by most cells in response to viral infection and are potent modulators of lymphocyte function.
- *Complement* in secretions is also derived from plasma. In association with antibodies, it plays a major cytotoxic role.
- *Surfactant protein A* (SPA) is one of four species of surfactant proteins which opsonizes bacteria/particles, enhancing phagocytosis by macrophages.
- *Defensins* are bactericidal peptides present in the azurophil granules of neutrophils.
- *Dimeric secretory IgA* with secretory component acts against antigens (p. 275).

Innate and adapted immunity

The above mechanisms act as a defence against microbes, inorganic substances, e.g. asbestos, particulate matter, such as dust, and other antigens. They act by aiding opsonization whereby macrophages can engulf the antigens.

With infection, neutrophils migrate out of pulmonary capillaries into the air spaces and phagocytose and kill microbes with, for example, antimicrobial proteins (lactoferrin), degra-

dative enzymes (elastase) and antioxidants. In addition, neutrophil extracellular traps (NET) ensnare and kill extracellular bacteria. Alveolar T lymphocytes stimulate the macrophages with cytokines, e.g. interleukins.

Neutrophils also generate a variety of mediators, e.g. TNF-α, IL-1 and chemokines which activate dendritic cells and B cells and produce the T-cell-activating cytokine IL-12. The latter enhances neutrophil-mediated defence during pneumonia. Dendritic cells are antigen presenting cells and are key to the adaptive immune response (p. 60).

Microbes are detected by host cells by pattern recognition receptors, e.g. toll-like receptors. These, as well as the stimulated cells above, act on NF-κB transcription factors in the epithelial cells to produce adhesion molecules, chemokines and colony stimulating factors to initiate inflammation. Inflammation is necessary for innate immunity and host defence but leads to lung damage; there is a fine line between defence and injury.

SYMPTOMS (see also p. 688)

Runny, blocked nose and sneezing

Nasal symptoms are extremely common. It can be difficult to distinguish between the common cold or allergic rhinitis as a cause of 'runny nose' (rhinorrhoea), nasal blockage and attacks of sneezing. In allergic rhinitis, symptoms may be intermittent, following contact with pollens or animal danders, or persistent, especially when the house-dust mite is the allergen. Colds are frequent during the winter but if the symptoms persist for weeks the patient is probably suffering from perennial rhinitis rather than from persistent viral infection.

Nasal secretions are usually thin and runny in rhinitis but thicker and yellowish green with viral infections. Nose bleeds and blood-stained nasal discharge are common and rarely indicate serious pathology. However, a blood-stained nasal discharge associated with nasal obstruction and pain may be the presenting feature of a nasal tumour (p. 1082). Nasal polyps typically present with nasal blockage and loss of smell.

Cough (see also p. 845)

Cough is the commonest symptom of lower respiratory tract disease. It is caused by mechanical or chemical stimulation of cough receptors in the epithelium of the pharynx, larynx, trachea, bronchi and diaphragm. Afferent receptors go to the cough centre in the medulla where efferent signals are generated to the expiratory musculature. Smokers often have a morning cough with little sputum. A productive cough is the cardinal feature of chronic bronchitis, while dry coughing, particularly at night, can be a symptom of asthma. Cough also occurs in asthmatics after mild exertion or following forced expiration. Cough can also occur for psychological reasons without any definable pathology.

A worsening cough is the most common presenting symptom of a bronchial carcinoma. The explosive character of a normal cough is lost when a vocal cord is paralysed – a bovine cough – usually as a result of lung cancer infiltrating the left recurrent laryngeal nerve. Cough can be accompanied by stridor in whooping cough and in the presence of laryngeal or tracheal obstruction.

Sputum

Approximately 100 mL of mucus is produced daily in a healthy, non-smoking individual. This flows at a regular pace up the airways, through the larynx, and is then swallowed. Excess mucus is expectorated as sputum. The most common cause of excess mucus production is cigarette smoking.

Mucoid sputum is clear and white but can contain black specks resulting from the inhalation of carbon. Yellow or green sputum is due to the presence of cellular material, including bronchial epithelial cells, or neutrophil or eosinophil granulocytes. Yellow sputum is not necessarily due to infection, as eosinophils in the sputum, as seen in asthma, can give the same appearance. The production of large quantities of yellow or green sputum is characteristic of bronchiectasis.

Haemoptysis (blood-stained sputum) varies from small streaks of blood to massive bleeding.

- The most common cause of mild haemoptysis is acute infection, particularly in exacerbations of chronic obstructive pulmonary disease (COPD) but it should not be attributed to this without investigation.
- Other common causes are pulmonary infarction, bronchial carcinoma and tuberculosis.
- In lobar pneumonia, the sputum is usually rusty in appearance rather than frankly blood-stained.
- Pink, frothy sputum is seen in pulmonary oedema.
- In bronchiectasis, the blood is often mixed with purulent sputum.
- Massive haemoptyses (>200 mL of blood in 24 hours) are usually due to bronchiectasis or tuberculosis.
- Uncommon causes of haemoptyses are idiopathic pulmonary haemosiderosis, Goodpasture's syndrome, microscopic polyangiitis, trauma, blood disorders and benign tumours.

Haemoptysis should always be investigated. Often, the diagnosis can be made from a chest X-ray, but a normal chest X-ray does not exclude disease. However, bronchoscopy is only diagnostic in about 5% of patients with haemoptysis and a normal chest X-ray.

Firm plugs of sputum may be coughed up by patients suffering from an exacerbation of allergic bronchopulmonary aspergillosis. Sometimes such sputum may look like casts of inflamed bronchi.

Breathlessness

Dyspnoea is a sense of awareness of increased respiratory effort that is unpleasant and that is recognized by the patient as being inappropriate. It is highly unlikely that this term will be used by the patient. Patients may complain of tightness in the chest; this must be differentiated from angina.

Breathlessness should be assessed in relation to the patient's lifestyle. For example, a moderate degree of breathlessness may be totally disabling if the patient has to climb many flights of stairs to reach home.

FURTHER READING

Tarlo SM. Cough: occupational and environmental considerations: ACCP evidence-based clinical practice guidelines. *Chest* 2006; **129** suppl 1: 186S–196S.

Orthopnoea (see p. 688) is breathlessness on lying down and is partly due to the weight of the abdominal contents pushing the diaphragm up into the thorax. Such patients may also become breathless on bending over.

Tachypnoea and *hyperpnoea* refer, respectively, to an increased rate of breathing and an increased level of ventilation, which may be appropriate to the situation (e.g. during exercise).

Hyperventilation is inappropriate overbreathing. This may occur at rest or on exertion and results in a lowering of the alveolar and arterial P_{CO_2} (see p. 1208).

Paroxysmal nocturnal dyspnoea is described on page 688.

Wheezing

Wheezing is a common complaint and is the result of airflow limitation due to any cause. The symptom of wheezing is not diagnostic of asthma; it may be absent in the early stages of this disease, and may also occur in patients with bronchiolitis or chronic obstructive pulmonary disease.

Chest pain

The most common type of chest pain encountered in respiratory disease is a localized sharp pain, often referred to as pleuritic. It is made worse by deep breathing or coughing and can be precisely localized by the patient. Localized anterior chest pain may be accompanied by tenderness of a costochondral junction as a symptom of costochondritis. Pain in the shoulder tip suggests irritation of the diaphragmatic pleura, whereas central chest pain radiating to the neck and arms is more typically cardiac in origin. Retrosternal soreness may occur in patients with tracheitis, and a constant, severe, dull pain may be the result of invasion of the thoracic wall by carcinoma.

EXAMINATION OF THE RESPIRATORY SYSTEM

The nose (p. 1080)

The chest

Examination of the chest

Inspection

Look for mental alertness, cyanosis, breathlessness at rest, use of accessory muscles, any deformity or scars on the chest and movement on both sides. A coarse tremor or flap of the outstretched hands indicates CO_2 intoxication. Prominent veins on the chest may imply obstruction of the superior vena cava.

Cyanosis (see p. 690) is a dusky colour of the skin and mucous membranes, due to the presence of more than 5 g/dL of desaturated haemoglobin. When due to central causes, cyanosis is visible on the tongue (especially the underside) and lips, and indicates a P_{aO_2} below about 6 kPa. Patients with central cyanosis will also be cyanosed peripherally. Peripheral cyanosis without central cyanosis is caused by a reduced peripheral circulation and is noted on the fingernails and skin of the extremities with associated coolness of the skin.

Finger clubbing is present when the normal angle between the base of the nail and the nail fold is lost. The base of the nail is fluctuant owing to increased vascularity, and there is an increased curvature of the nail in all directions, with expansion of the end of the digit. Some causes of clubbing are given in Table 14.1. Clubbing is not a feature of uncomplicated COPD.

Palpation and percussion

Check the position of the trachea and apex beat. Examine the supraclavicular fossa for enlarged lymph nodes. The distance between the sternal notch and the cricoid cartilage (three to four finger breadths in full expiration) is reduced in patients with severe airflow limitation. Check chest expansion. A tape measure may be used if precise measurements are needed, e.g. in ankylosing spondylitis. Local discomfort over the sternochondral joints may suggest costochondritis. Compression of the chest laterally and anteroposteriorly may produce localized pain suggesting a rib fracture. On percussion liver dullness is usually detected anteriorly at the level of the sixth rib. Liver and cardiac dullness are lost with overinflated lungs (Table 14.2).

Auscultation

Ask the patient to take deep breaths through the mouth. Inspiration is more prolonged than expiration. Healthy lungs filter off most of the high-frequency component, which is mainly due to turbulent flow in the larynx. Normal breath sounds are harsher anteriorly over the upper lobes (particularly on the right) and described as vesicular.

Bronchial breathing is heard best over consolidated or collapsed lung and sometimes over areas of localized fibrosis or bronchiectasis. Such areas conduct the high-frequency hissing component of breath sounds well. Characteristically, the noise heard during inspiration and expiration is equally long but separated by a short silent phase. Whispering pectoriloquy (whispered, higher-pitched sounds heard distinctly through a stethoscope) invariably accompanies bronchial breathing.

Table 14.1 Some causes of finger clubbing

Respiratory
Bronchial carcinoma, especially epidermoid (squamous cell) type (major cause)
Chronic suppurative lung disease:
Bronchiectasis
Lung abscess
Empyema
Pulmonary fibrosis (e.g. idiopathic lung fibrosis)
Pleural and mediastinal tumours (e.g. mesothelioma)
Cryptogenic organizing pneumonia
Cardiovascular
Cyanotic heart disease
Subacute infective endocarditis
Atrial myxoma
Miscellaneous
Congenital – no disease
Cirrhosis
Inflammatory bowel disease

Table 14.2 Physical signs of respiratory disease

Pathological process	Chest wall movement (reduced)	Mediastinal displacement	Percussion note	Breath sounds	Vocal resonance	Added sounds
Consolidation (i.e. lobar pneumonia)	Affected side	None	Dull	Bronchial	Increased	Fine crackles
Collapse						
Major bronchus	Affected side	Towards lesion	Dull	Diminished or absent	Reduced or absent	None
Peripheral bronchus	Affected side	Towards lesion	Dull	Bronchial	Increased	Fine crackles
Fibrosis						
Localized	Affected side	Towards lesion	Dull	Bronchial	Increased	Coarse crackles
Generalized (e.g. idiopathic lung fibrosis)	Both sides	None	Normal	Vesicular	Increased	Fine crackles
Pleural effusion (>500 mL)	Affected side	Away from lesion (in massive effusion)	Stony dull	Vesicular reduced or absent	Reduced or absent	None
Large pneumothorax	Affected side	Away from lesion	Normal or hyperresonant	Reduced or absent	Reduced or absent	None
Asthma	Both sides	None	Normal	Vesicular Prolonged expiration	Normal	Expiratory polyphonic wheeze
Chronic obstructive pulmonary disease	Both sides	None	Normal	Vesicular Prolonged expiration	Normal	Expiratory polyphonic wheeze and coarse crackles

Added sounds

Wheeze. Wheeze is usually heard during expiration and results from vibrations in the collapsible part of the airways when apposition occurs as a result of the flow-limiting mechanisms. Wheezes are heard in asthma and in chronic obstructive pulmonary disease, but are not invariably present. In the most severe cases of asthma a wheeze may not be heard, as airflow may be insufficient to generate the sound. Wheezes may be monophonic (single large airway obstruction) or polyphonic (narrowing of many small airways). An end-inspiratory (as opposed to expiratory) 'squeak' may be heard in obliterative bronchiolitis.

Crackles. These brief crackling sounds are probably produced by opening of previously closed bronchioles, and their timing during breathing is of significance – early inspiratory crackles are associated with diffuse airflow limitation, whereas late inspiratory crackles are characteristically heard in pulmonary oedema, fibrosis of the lung and bronchiectasis.

Pleural rub. This is a creaking or groaning sound that is usually well localized. It indicates inflammation and roughening of the pleural surfaces, which normally glide silently over one another.

Vocal resonance. Healthy lung attenuates high-frequency notes, as compared to the lower-pitched components of speech. Consolidated lung has the reverse effect, transmitting high frequencies well; the spoken word then takes on a bleating quality. Whispered (and therefore high-pitched) speech can be clearly heard over consolidated areas, as compared to healthy lung. Low-frequency sounds such as 'ninety-nine' are well transmitted across healthy lung to produce vibration that can be felt over the chest wall. Consolidated lung transmits these low-frequency noises less well, and pleural fluid severely dampens or obliterates the vibrations altogether. Tactile vocal fremitus is the palpation of this vibration, usually by placing the edge of the hand on the chest wall. For all practical purposes this duplicates the assessment of vocal resonance and is not routinely performed as part of the chest examination.

Cardiovascular system examination (p. 690) gives additional information about the lungs.

Additional bedside tests

Since so many patients with respiratory disease have airflow limitation, airflow should be routinely measured using a peak flow meter or spirometer. This will provide a much more accurate assessment of airflow limitation than any physical sign.

INVESTIGATION OF RESPIRATORY DISEASE

Imaging

Radiology plays an essential part in investigating most chest symptoms. Some diseases such as tuberculosis or lung cancer may be undetectable on clinical examination but are obvious on the chest X-ray. Conversely, asthma or chronic bronchitis may be associated with a normal chest X-ray. Always try to get previous films for comparison.

Chest X-ray

When viewing films consider:

- *Centring of the film*. The distance between each clavicular head and the spinal processes should be equal.
- *Penetration* (check film is not too dark).
- *The view*. Routine films are taken PA, i.e. the film is placed in front of the patient with the X-ray source behind. Anteroposterior (AP) films are taken only in very ill patients who are unable to stand up or be taken to the radiology department; the cardiac outline appears bigger and the scapulae cannot be moved out of the way. Lateral chest X-ray were often performed in the past to localize pathology, but CT scans are generally more useful.

Look at:

- the shape and bony structure of the chest wall
- whether the trachea is central
- whether the diaphragm is elevated or flat
- the shape, size and position of the heart
- the shape and size of the hilar shadows
- the size and shape of any lung abnormalities and vascular shadowing.

X-ray abnormalities

Collapse and consolidation

Simple pneumonia is easy to recognize (see Fig. 14.35) but look carefully for any evidence of collapse (Figure 14.11, Table 14.3). Loss of volume or crowding of the ribs are the best indicators of lobar collapse. The lung lobes collapse in characteristic directions. The lower lobes collapse downwards and towards the mediastinum, the left upper lobe collapses forwards against the anterior chest wall, while the right upper lobe collapses upwards and outwards, forming the appearance of an arch over the remaining lung. The right middle lobe collapses anteriorly and inward, obscuring the right heart border. If a whole lung collapses, the mediastinum will shift towards the side of the collapse. Uncomplicated consolidation does not cause mediastinal shift or loss of lung volume, so any of these features should raise the suspicion of an endobronchial obstruction.

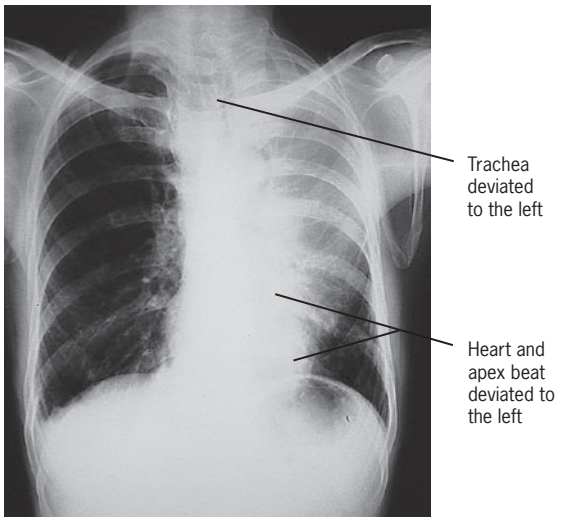


Fig. 14.11 **Collapse of the left lung.** Chest X-ray showing hazy shadowing in the left upper zone with a raised left diaphragm. There is compensatory emphysema of the right lung.

Pleural effusion (see Fig. 14.45)

Pleural effusions need to be larger than 500 mL to cause much more than blunting of the costophrenic angle. On an erect film they produce a characteristic shadow with a curved upper edge rising into the axilla. If very large, the whole of one side of the thorax may be opaque, with shift of the mediastinum to the opposite side.

Fibrosis

Localized fibrosis causes streaky shadowing, and the accompanying loss of lung volume causes mediastinal structures to move to the same side. More generalized fibrosis in the lung can lead to a honeycomb appearance (see p. 873), seen as diffuse shadows containing multiple circular translucencies a few millimetres in diameter.

Round shadows

Lung cancer is the commonest cause of large round shadows but many other causes are recognized (Table 14.4).

Miliary mottling

This term describes numerous minute opacities, 1–3 mm in size, which are caused by many pathological processes. The most common causes are miliary tuberculosis, pneumoconiosis, sarcoidosis, fibrosing alveolitis (idiopathic pulmonary

Table 14.3 Causes of collapse of the lung

Enlarged tracheobronchial lymph nodes due to malignant disease, tuberculosis
Inhaled foreign bodies (e.g. peanuts) in children, usually in the right main bronchus
Bronchial casts or plugs (e.g. allergic bronchopulmonary aspergillosis)
Retained secretions – postoperatively and in debilitated patients

Table 14.4 Causes of round shadows (>3 cm) in the lung

Carcinoma
Metastatic tumours (usually multiple shadows)
Lung abscess (usually with fluid level)
Encysted interlobar effusion (usually in horizontal fissure)
Hydatid cysts (often with a fluid level)
Arteriovenous malformations (usually adjacent to a vascular shadow)
Aspergilloma
Rheumatoid nodules
Tuberculoma (may be calcification within the lesion)
Rare causes: Bronchial carcinoid Cylindroma Chondroma Lipoma
Other shadows related to mediastinum: Pericardium, Oesophagus, Spinal cord } Seen on lateral chest X-ray

fibrosis) and pulmonary oedema (see Fig. 13.16), though the last is usually perihilar and accompanied by larger, fluffy shadows. A rare but striking cause of miliary mottling is pulmonary microlithiasis.

Computed tomography

CT scanners provide excellent images of the lungs and mediastinal structures (Fig. 14.12). Several different protocols are used, depending on the suspected pathology. Mediastinal structures are shown more clearly by injecting intravenous contrast medium to enhance the vascular structures.

CT is essential in staging bronchial carcinoma by demonstrating tumour size, nodal involvement, metastases and invasion of mediastinum, pleura or chest wall. CT-guided needle biopsy allows samples to be obtained from peripheral masses.

Enlarged mediastinal nodes (1 cm) may be either malignant or reactive and may require biopsy. Staging scans should include assessment of liver and adrenals, which are common sites for metastatic disease.

High-resolution CT (HRCT) scanning (sampling lung parenchyma with 1–2 mm thickness scans at 10–20 mm intervals) allows assessment of diffuse inflammatory and infective parenchymal processes. It is valuable in the following situations:

- Detection of diffuse disease of the lung parenchyma, including sarcoidosis, cryptogenic and extrinsic allergic alveolitis, occupational lung disease, and any other form of interstitial pulmonary fibrosis.
- Diagnosis of bronchiectasis. High-resolution CT has a sensitivity and specificity of greater than 90%. This technique has replaced bronchography.
- Distinguishing emphysema from diffuse parenchymal lung disease or pulmonary vascular disease as a cause of a low gas transfer factor with otherwise normal lung function.
- Diagnosis of lymphangitis carcinomatosa.

Multi-slice (up to 64) CT scanners are used to produce detailed images in two or three dimensions in any plane. This detail is particularly useful for the detection of pulmonary emboli. Pulmonary nodules and airway disease are more easily defined and the technique makes HRCT less necessary.

Magnetic resonance imaging

Magnetic resonance imaging (MRI) is less valuable than CT in assessing the lung parenchyma. In the mediastinum, MRI with ECG-gating allows accurate images of the heart and aortic aneurysms to be obtained, and MRI retains a place in staging lung cancer, for assessing tumour invasion in the mediastinum, chest wall and at the lung apex, because it produces good images in the sagittal and coronal planes. Vascular structures can be clearly differentiated as flowing blood produces a signal void on MRI.

Positron emission tomography (PET)

Tumours take up a tracer, e.g. labelled fluorodeoxyglucose (FDG), to highlight the glycolytic pathway of tumour cells. This emits positrons that can be imaged, which helps to differentiate benign from malignant tumours due to the avid

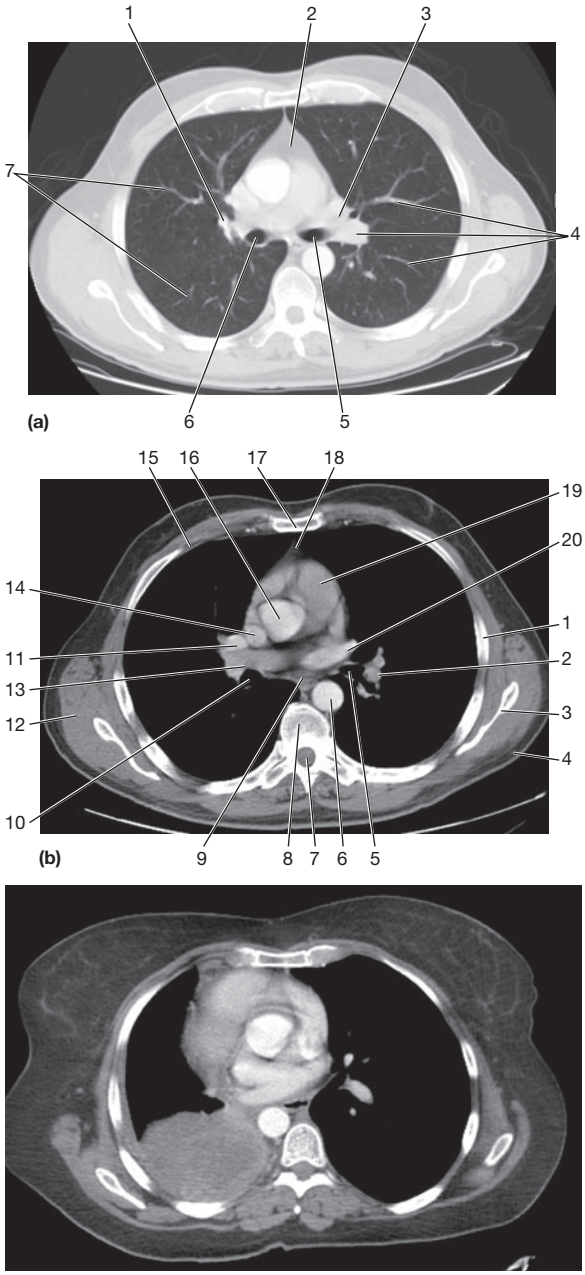


Fig. 14.12 **CT scan of the lung. (a) Lung setting – showing normal lung markings.** 1, right hilum; 2, mediastinum; 3, left hilum; 4, lung vessels; 5, left main bronchus; 6, right main bronchus; 7, peripheral lung vessels. **(b) Mediastinal (soft tissue) setting** – showing normal mediastinal structures following intravenous contrast enhancement. 1, rib; 2, descending left pulmonary artery; 3, scapula; 4, subcutaneous fat; 5, left main bronchus; 6, descending aorta; 7, spinal canal; 8, vertebral body; 9, oesophagus; 10, right main bronchus; 11, right pulmonary artery; 12, muscle; 13, right superior pulmonary vein; 14, superior vena cava; 15, costal cartilage; 16, ascending aorta; 17, sternum; 18, thymic remnant; 19, pulmonary trunk; 20, left superior pulmonary vein. **(c) Post-contrast scan** showing large right upper zone carcinoma with enlarged lymph nodes in the mediastinum surrounding the trachea.

FURTHER READING

Hansell DM, Armstrong P, Lynch DA et al. *Imaging of Diseases of the Chest,* 4th edn. Chicago: Elsevier Mosby, 2004.

ability of malignant tumours to take up glucose. PET scanning is the investigation of choice for assessing lymph nodes and metastatic disease in bronchial carcinoma. PET-CT improves localization of tumours.

Scintigraphic imaging

This technique is used widely for the detection of pulmonary emboli although it is now performed less often owing to widespread use of D-dimer measurements and CT pulmonary angiography.

Perfusion scan

Macro-aggregated human albumin labelled with technetium-99m is injected intravenously. The particles are of such a size that they impact in pulmonary capillaries, where they remain for a few hours. A gamma camera is then used to detect the position of the macro-aggregated human albumin. The resultant pattern indicates the distribution of pulmonary blood flow; cold areas occur where there is defective blood flow (e.g. in pulmonary emboli).

Ventilation–perfusion scan (see p. 785)

Xenon-133 gas is inhaled into the lung and its distribution is detected at the same time as the perfusion scan. Using the two scans, a pulmonary embolus can be seen to cause a striking diminution of perfusion relative to ventilation. Other lung diseases (e.g. asthma or pneumonia) impair both ventilation and perfusion. Unfortunately, however, a pulmonary embolus often produces substantial changes in the lung substance (e.g. atelectasis) so that such a clear distinction is not always obvious. Nevertheless, this is a better technique than perfusion scan alone.

Ultrasound (USS)

Ultrasound is not useful for lung parenchymal disease as ultrasound energy is scattered by air. Ultrasound is useful for pleuritic lesions, particularly for diagnosing and aspirating small pleural effusions.

Respiratory function tests (Table 14.5)

In clinical practice, airflow limitation can be assessed by relatively simple tests that have good intra-subject repeatability. Normal values are required for their interpretation since these tests vary considerably, not only with sex, age and height, but also between individuals of the same age, sex and height. The standard deviation about the mean for a group of individuals is therefore very high; for example, the standard deviation for the peak expiratory flow rate is approximately 50 L/min, and for the FEV_1 it is approximately 0.4 L. Repeated measurements of lung function are useful for assessing the progression of disease in an individual patient.

Tests of ventilatory function

These tests are used mainly to assess the degree of airflow limitation present during expiration.

Table 14.5 Respiratory function tests and exercise tests

Test	Use	Advantages	Disadvantages
PEFR	Monitoring changes in airflow limitation in asthma	Portable Can be used at the bedside	Effort-dependent Poor measure of chronic airflow limitation
FEV, FVC, FEV_1/FVC	Assessment of airflow limitation The best single test	Reproducible Relatively effort-independent	Bulky equipment but smaller portable machines available
Flow–volume curves	Assessment of flow at lower lung volumes Detection of large airway obstruction both intra- and extrathoracic (e.g. tracheal stenosis, tumour)	Recognition of patterns of flow–volume curves for different diseases	Sophisticated equipment needed for full test but expiratory loop now possible with compact spirometry
Airways resistance	Assessment of airflow limitation	Sensitive	Technique difficult to perform
Lung volumes	Differentiation between restrictive and obstructive lung disease	Effort-independent, complements FEV_1	Sophisticated equipment needed
Gas transfer	Assessment and monitoring of extent of interstitial lung disease and emphysema	Non-invasive (compared with lung biopsy or radiation from repeated chest X-rays and CT)	Sophisticated equipment needed
Blood gases	Assessment of respiratory failure	Can detect early lung disease when measured during exercise	Invasive
Pulse oximetry	Postoperative, sleep studies and respiratory failure	Continuous monitoring Non-invasive	Measures saturation only
Exercise tests (6-min walk)	Practical assessment for disability and effects of therapy	No equipment required	Time-consuming Learning effect At least two walks required
Cardiorespiratory assessment	Early detection of lung/heart disease Fitness assessment	Differentiates breathlessness due to lung or heart disease	Expensive and complicated equipment required

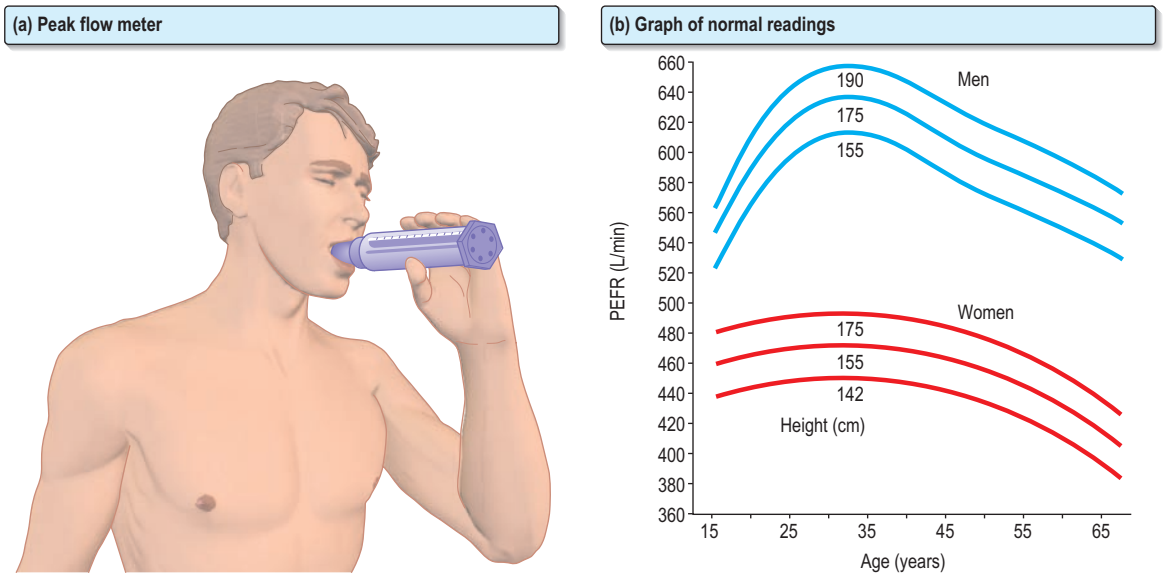


Fig. 14.13 **Peak flow measurements. (a)** Peak flow meter: the lips should be tight around the mouthpiece. **(b)** Graph of normal readings for men and women.

Peak expiratory flow rate (PEFR)

This is an extremely simple and cheap test. Subjects are asked to take a full inspiration to total lung capacity and then blow out forcefully into the peak flow meter (Fig. 14.13), which is held horizontally. The lips must be placed tightly around the mouthpiece. The best of three tests is recorded.

Although reproducible, PEFR is not a good measure of airflow limitation since it measures the expiratory flow rate only in the first 2 ms of expiration and overestimates lung function in patients with moderate airflow limitation. PEFR is best used to monitor progression of disease and its treatment. Regular measurements of peak flow rates on waking, during the afternoon, and before bed demonstrate the wide diurnal variations in airflow limitation that characterize asthma and allow an objective assessment of treatment to be made (Fig. 14.14).

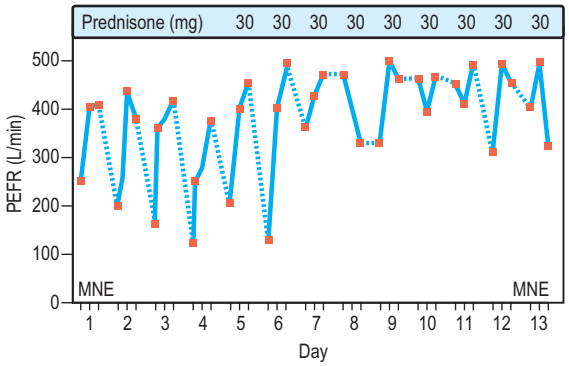


Fig. 14.14 **Diurnal variability in PEFR in asthma, showing the effect of steroids.** M, morning; N, noon; E, evening.

Spirometry

The spirometer measures the FEV_1 and the forced vital capacity (FVC). Both the FEV_1 and FVC are related to height, age and sex. The technique involves a maximum inspiration followed by a forced expiration (for as long as possible) into the spirometer. The act of expiration triggers the moving record chart, which measures volume against time. The record chart moves for a total of 5 s, but expiration should continue until all the air has been expelled from the lungs, as patients with severe airflow limitation may have a very prolonged forced expiratory time. This is demonstrated on the record chart in Figure 14.15.

The FEV_1 expressed as a percentage of the FVC is an excellent measure of airflow limitation. In normal subjects it is around 75%. With increasing airflow limitation the FEV_1 falls proportionately more than the FVC, so that the FEV_1/FVC ratio is reduced. With restrictive lung disease the FEV_1 and the FVC are reduced in the same proportion and the FEV_1/FVC ratio remains normal or may even increase because of the enhanced elastic recoil.

In chronic airflow limitation (particularly in COPD and asthma) the total lung capacity (TLC) is usually increased, yet there is nearly always some reduction in the FVC. This is the result of disease in the small airways causing obstruction to airflow before the normal RV is reached. This trapping of air within the lung (giving an increased RV) is a characteristic feature of these diseases.

Other ventilatory function tests

Measurement of airways resistance in a body plethysmograph is more sensitive but the equipment is expensive and the necessary manoeuvres are too exhausting for many patients with chronic airflow limitation.

Flow–volume loops

The ability to measure flow rates against volume (flow–volume loops, see Fig. 14.8) enables a more sophisticated analysis to be made of the site of airflow limitation within the lung. At the start of expiration from TLC, the site of maximum resistance is the large airways, and this accounts for the flow reduction in the first 25% of the curve. As the lung volume

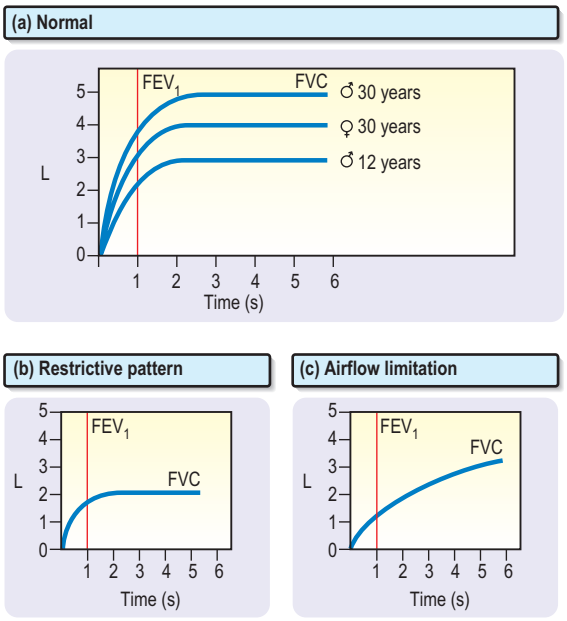


Fig. 14.15 **Spirometry.** Volume–time curves showing **(a)** normal patterns for age and sex, **(b)** restrictive pattern (FEV_1 and FVC reduced), **(c)** airflow limitation (FEV_1 only reduced).

reduces further, so the elastic pressures within the lung holding open the smaller airways reduce, and disease of the lung parenchyma or the small airways themselves becomes apparent. For example, in diseases such as chronic obstructive pulmonary disease (COPD), where the brunt of the disease falls upon the smaller airways, expiratory flow rates at 50% or 25% of the vital capacity may be disproportionately reduced when compared with flow rates at larger lung volumes.

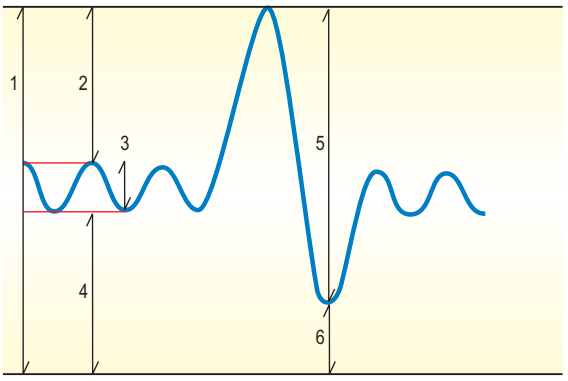
Lung volume

The subdivisions of the lung volume are shown in Figure 14.16. Tidal volume and vital capacity can be measured using a simple spirometer, but the TLC and RV need to be measured by an alternative technique. TLC is measured by connecting the lungs to a reservoir containing a known amount of non-absorbable gas (helium) that can readily be measured. If the concentration of the gas in the reservoir is known at the start of the test and is measured after equilibration of the gas has occurred (when the patient has breathed in and out of the reservoir), the dilution of the gas will reflect the TLC. This technique is known as *helium dilution*. RV can be calculated by subtracting the vital capacity from the TLC.

The TLC measured using this technique is inaccurate if large cystic spaces are present in the lung, because the helium cannot diffuse into them. Under these circumstances the thoracic gas volume can be measured more accurately using a body plethysmograph. The difference between the two measurements can be used to define the extent of non-communicating air space within the lungs.

Transfer factor

This measures the transfer of gas across the alveolar–capillary membrane and reflects the uptake of oxygen from the alveoli into the red cells. A low concentration of carbon monoxide is inhaled and is avidly taken up in a linear fashion



1 Total lung capacity
2 Inspiratory reserve volume
3 Tidal volume
4 Functional residual capacity
5 Vital capacity
6 Residual volume

Fig. 14.16 **The subdivisions of the lung volume.**

by circulating haemoglobin, the amount of which must be known when the test is performed. In normal lungs the transfer factor is a true measure of the diffusing capacity of the lungs for oxygen and depends on the thickness of the alveolar–capillary membrane. In lung disease the diffusing capacity (D_{CO}) also depends on the $\dot{V}_A/\dot{Q}$ relationship as well as on the area and thickness of the alveolar membrane. To control for differences in lung volume, the uptake of carbon monoxide is related to the lung volume; this is known as the transfer coefficient (K_{CO}).

Gas transfer is usually reduced in patients with severe degrees of emphysema and fibrosis. Overall gas transfer can be thought of as a relatively non-specific test of lung function but one that can be particularly used in the early detection and assessment of progress of diseases affecting the lung parenchyma (e.g. idiopathic pulmonary fibrosis, sarcoidosis, asbestosis).

Measurement of blood gases

This technique is described on page 911.

Measurement of the partial pressures of oxygen and carbon dioxide within arterial blood is essential in the management of cases of respiratory failure and severe asthma, when repeated measurements are often the best guide to therapy.

Arterial oxygen saturation (S_aO_2) can be continuously measured using an oximeter with either ear or finger probes. The oximeter measures the differential absorption of light by oxy- and deoxyhaemoglobin and measures saturation to within 5% of that obtained by blood gas analysis.

Nitric oxide exhalation test

Endogenous nitric oxide is present in the breath and increased in asthma. It is an indicator of airway inflammation. It is increased in asthma and can be readily analysed.

Haematological and biochemical tests

It is useful to measure:

- Haemoglobin, to detect anaemia or polycythaemia.
- Packed cell volume (secondary polycythaemia occurs with chronic hypoxia).

www.studentconsult.com
Ask the authors

- Routine biochemistry (often disturbed in carcinoma and infection).
- D-dimer to detect intravascular coagulation. A negative test makes pulmonary embolism very unlikely.

Other blood investigations sometimes required include α_1-antitrypsin levels, *Aspergillus* antibodies, viral and mycoplasma serology, autoantibody profiles and specific IgE measurements.

Sputum

Sputum should be inspected for colour:

- yellowish green indicates inflammation (infection or allergy)
- the presence of blood suggests neoplasm or pulmonary infarct (see haemoptysis, p. 819).

Microbiological studies (e.g. Gram stain and culture) are not helpful in upper respiratory tract infections or in acute or chronic bronchitis. They are of value in:

- pneumonia
- tuberculosis
- unusual clinical problems
- *Aspergillus* lung disease.

Sputum cytology

This is useful in the diagnosis of bronchial carcinoma and asthma (the presence of eosinophils). Its advantages are its speed, cheapness and non-invasive nature.

However, its value depends on the production of sputum and the presence of a reliable cytologist. Sputum can be induced following the inhalation of nebulized hypertonic saline (5%). Bronchoscopy and bronchial washings (see p. 828), or, occasionally, transtracheal aspiration are performed.

Transtracheal aspiration

This technique involves pushing a needle through the cricothyroid membrane, through which a catheter is threaded to a position just above the carina. This procedure induces coughing, and specimens are collected by aspiration or by the introduction and subsequent aspiration of sterile saline. Although not often required, it is an excellent technique for assessing infection in the lower respiratory tract because it avoids contamination of the specimen with bacteria from the pharynx and mouth.

Exercise tests

The predominant symptom in respiratory medicine is breathlessness. The degree of disability produced by breathlessness can be assessed before and after treatment by asking the patient to walk for 6 minutes along a measured track. This has been shown to be a reproducible and useful test once the patient has undergone an initial training walk to overcome the learning effect. Additional information can be obtained by attaching a pulse oximeter during exercise to assess desaturation on exertion.

More sophisticated exercise tests assessing both lung and heart function are useful in investigating unexplained breathlessness. Measurement of uptake of oxygen ($\dot{V}_{O_2}$), work performed, heart rate and blood pressure together with serial ECGs allows:

- the early detection of lung disease
- the detection of myocardial ischaemia
- the distinction between lung and heart disease
- assessment of fitness.

Pleural aspiration

Diagnostic aspiration is necessary for all but very small effusions. A needle attached to a 20 mL syringe is inserted under local anaesthesia through an intercostal space towards the top of an area of dullness. Fluid is withdrawn and the presence of any blood is noted. Samples are sent for protein estimation, lactate dehydrogenase, cytology and bacteriological examination, including culture and Ziehl–Neelsen/auramine stain for tuberculosis. Large amounts of fluid can be aspirated through a large-bore needle to help relieve extreme breathlessness. To avoid introducing infection into the pleural space, and causing an empyema, full aseptic precautions should be used.

For small or loculated effusions, ultrasound is useful to guide aspiration and drainage.

Pleural biopsy

Blind pleural biopsy (Fig. 14.17) is a useful technique if performed by an experienced operator, and can yield positive results in up to 80% of cases of tuberculosis and in 60% of cases of malignancy provided multiple biopsies are taken (Practical box 14.1). Better yields can be obtained by biopsy under direct vision using video-assisted thoracoscopy, especially in mesothelioma.

Intercostal drainage

This is carried out when large effusions are present, producing severe breathlessness, or for drainage of an empyema

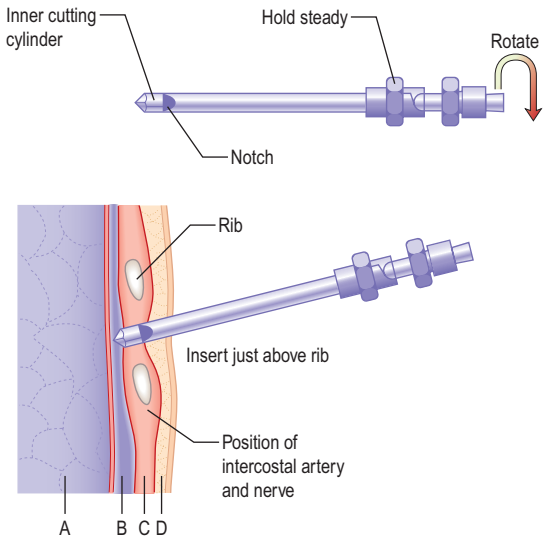


Fig. 14.17 **Technique of pleural biopsy.** The biopsy needle is shown penetrating the chest wall. A, lung parenchyma; B, pleural space; C, muscle; D, skin and subcutaneous tissue.

Practical Box 14.1 Pleural biopsy

- Explain to the patient the nature of the procedure.
- Get written consent.

1. Pleural biopsy is best performed after the aspiration of diagnostic fluid samples but before draining large volumes of fluid.
2. A small skin incision is made, as the end of the Abrams' pleural biopsy needle is blunt.
3. Once in place through the pleura, the back part of the needle is rotated to open the notch; this is kept pointing forward.
4. With lateral pressure the needle is withdrawn so that the notch will snag against the pleura.
5. The needle is held firmly and the hexagonal grip is twisted clockwise to cut the biopsy. To avoid damage to the intercostal vessel or nerve, the notch should never be directed upwards when the biopsy is taken.
6. Several biopsies should be taken at different angles by repeated insertion of the needle.
7. Specimens should be put in sterile saline for culture for tuberculosis and into 10% formol saline for histological examination.

Practical Box 14.2 Intercostal drainage

- Explain to the patient the nature of the procedure.
- Get written consent.

1. Carefully sterilize the skin over the aspiration site. Sterile gloves, cap, gown and mask must be worn.
2. Anaesthetize the skin, muscle and pleura with 2% lidocaine.
3. Make a small incision, then push a 28 French gauge Argyle catheter into the pleural space.
4. Attach to a three-way tap and 50 mL syringe.
5. Aspirate up to 1000 mL. Stop aspiration if the patient becomes uncomfortable – shock may ensue if too much fluid is withdrawn too quickly.
6. A Silastic pigtail catheter can be inserted under X-ray/ultrasound control and attached to tubing and bag for slower aspiration. For drainage and effusions, an 8–12 French gauge pigtail is inserted using the Seldinger technique in which the needle used to enter the pleural space is withdrawn over a control wire along which the catheter is then passed. A 14–16 French gauge pigtail catheter is used for drainage of empyema.

For pleurodesis

Tetracycline 500 mg or bleomycin 15 units in 30–50 mL sodium chloride 0.9% solution is instilled into the pleural cavity to achieve pleurodesis in recurrent/malignant effusion.

(Practical box 14.2). Pleurodesis is performed for recurrent/malignant effusion.

Mediastinoscopy and scalene node biopsy

Mediastinoscopy is used in the diagnosis of mediastinal masses and in staging nodal disease in carcinoma of the bronchus. An incision is made just above the sternum and a mediastinoscope inserted by blunt dissection.

Fibreoptic bronchoscopy

(Practical box 14.3 and Fig. 14.18)

Under local anaesthesia and sedation, the central airways can be visualized down to subsegmental level and biopsies

Practical Box 14.3 Fibreoptic bronchoscopy

This enables the direct visualization of the bronchial tree as far as the subsegmental bronchi under a local anaesthetic. Informed written consent should be obtained after explaining the nature of the procedure.

Indications

- Lesions requiring biopsy seen on chest X-ray.
- Haemoptysis.
- Stridor.
- Positive sputum cytology for malignant cells with no chest X-ray abnormality.
- Collection of bronchial secretions for bacteriology, especially tuberculosis.
- Recurrent laryngeal nerve paralysis of unknown aetiology.
- Infiltrative lung disease (to obtain a transbronchial biopsy).
- Investigation of collapsed lobes or segments and aspiration of mucus plugs.

Procedure

- The patient is starved overnight.
- Atropine 0.6 mg i.m. is given 30 min before the procedure.
- Topical anaesthesia (lidocaine 2% gel) is applied to the nose, nasopharynx and pharynx.
- Intravenous sedation (e.g. diazepam 10 mg or midazolam 2.5–10 mg) is given.
- The bronchoscope is passed through the nose, nasopharynx and pharynx under direct vision to minimize trauma.
- Lidocaine (2 mL of 4%) is dropped through the instrument on to the vocal cords.
- The bronchoscope is passed through the cords into the trachea.
- All segmental and subsegmental orifices should be identified.
- Biopsies and brushings should be taken of macroscopic abnormalities or occasionally from peripheral lesions under radiographic control.

Disadvantages

- All patients require sedation to tolerate the procedure.
- Minor and transient cardiac dysrhythmias occur in up to 40% of patients on passage of the bronchoscope through the larynx. Monitoring is required.
- Oxygen supplementation is required in patients with P_aO_2 below 8 kPa.
- Fibreoptic bronchoscopy should be performed with care in the very sick, and transbronchial biopsies avoided in ventilated patients owing to the increased risk of pneumothorax.
- Massive bleeding may occur on accidental biopsy of vascular lesions or carcinoid tumours. Rigid bronchoscopy may be required to allow adequate access to the bleeding point for haemostasis.

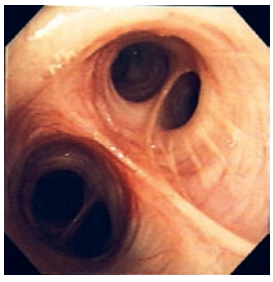


Fig. 14.18 **Normal endobronchial appearances as seen at fibreoptic bronchoscopy.**

taken for histology. More distal lesions may be sampled by washing or blind brushing. Diffuse inflammatory and infective lung processes may be sampled by bronchoalveolar lavage and transbronchial biopsy. The yield is best in sarcoidosis, lymphangitis carcinomatosa and extrinsic alveolitis. Other fibrotic lung diseases usually yield non-diagnostic samples so it may be more relevant to proceed directly to open or thoracoscopic lung biopsy.

Video-assisted thoracoscopic (VATS) lung biopsy

This technique reduces the need for open thoracotomy when a lung biopsy is required (p. 873).

Skin-prick tests

Allergen solutions are placed on the skin (usually the volar surface of the forearm) and the epidermis is broken using a 1 mm tipped lancet. A separate lancet should be used for each allergen. If the patient is sensitive to the allergen a weal develops and the diameter of the induration should be measured after 10 minutes. A weal of at least 3 mm diameter is regarded as positive provided that the control test is negative. The results should always be interpreted in the light of the history. Skin tests are not affected by bronchodilators or corticosteroids but antihistamines should be discontinued at least 48 hours before testing.

SMOKING AND AIR POLLUTION

SMOKING

Prevalence

Cigarette smoking has declined in the western world over the last decade. In the UK, 28% of men and women aged 16 years and over smoke, compared to 22% of adults in the USA. This compares with two-thirds of Chinese men who smoke. Smoking is on the increase in many developing countries, particularly in women.

Toxic effects

Cigarette smoke contains polycyclic aromatic hydrocarbons and nitrosamines, which are potent carcinogens and mutagens in animals. It causes release of enzymes from neutrophil granulocytes and macrophages that are capable of destroying elastin and leading to lung damage. Pulmonary epithelial permeability increases even in symptomless cigarette smokers, and correlates with the concentration of carboxyhaemoglobin in blood. This altered permeability may allow easier access for carcinogens.

The dangers

Cigarette smoking is addictive and harmful to health (Table 14.6). People usually start smoking in adolescence for psychosocial reasons and, once they smoke regularly, the pharmacological properties of nicotine play a major part in persistence, conferring some advantage to the smoker's mood. Very few cigarette smokers (less than 2%) can limit themselves to occasional or intermittent smoking.

There is a significant dose–response relationship between the number of cigarettes smoked daily, airway inflammation (Table 14.7) and lung cancer mortality. Sputum production and airflow limitation increase with daily cigarette consumption, and effort tolerance decreases. Smoking and asbestos exposure are synergistic in producing bronchial carcinoma, increasing the risk in asbestos workers by five to eight times that of non-smokers exposed to asbestos.

Cigarette smokers who change to other forms of tobacco can reduce the risk, even if they continue to inhale, and are better off changing to cigars or pipes. However, pipe and cigar smokers also have a greater risk of lung cancer than lifelong non-smokers or former smokers.

Environmental tobacco smoke ('passive smoking') has been shown to cause more frequent and more severe attacks of asthma in children and possibly increases the number of cases of asthma. It is also associated with a small but definite increase in lung cancer.

Table 14.6 The dangers of cigarette smoking

General
Lung cancer
COPD
Carcinoma of the oesophagus
Ischaemic heart disease
Peripheral vascular disease
Bladder cancer
An increase in abnormal spermatozoa
Memory problems
Maternal smoking
A decrease in birthweight of the infant
An increase in fetal and neonatal mortality
An increase in asthma
Passive smoking
Risk of asthma, pneumonia and bronchitis in infants of smoking parents
An increase in cough and breathlessness in smokers and non-smokers with COPD and asthma
Increased cancer risk

Table 14.7 Effects of smoking on the lung

Large airways
Increase in submucosal gland volume
Increase in number of goblet cells
Chronic inflammation
Metaplasia and dysplasia of the surface epithelium
Small airways
Increase in number and distribution of goblet cells
Airway inflammation and fibrosis
Epithelial metaplasia/dysplasia
Carcinoma
Parenchyma
Proximal acinar scarring
Increase in alveolar macrophage numbers
Emphysema (centri-acinar, pan-acinar)

Stopping smoking

If the entire population could be persuaded to stop smoking, the effect on healthcare use would be enormous. National campaigns, bans on advertising and a substantial increase in the cost of cigarettes are the best ways of achieving this at the population level. Smoking bans in the workplace, pubs and public spaces have also helped. Meanwhile, active encouragement to stop smoking remains a useful approach for individuals. Smokers who want to stop should have access to smoking cessation clinics to provide behavioural support. Nicotine replacement therapy (NRT) and bupropion are effective aids to smoking cessation in those smoking more than 10 cigarettes per day. Both should only be used in smokers who commit to a target stop date, and the initial prescription should be for 2 weeks beyond the target stop date. NRT is the preferred choice; there is no evidence that combined therapy offers any advantage. Therapy should be changed after 3 months if abstinence is not achieved.

Varenicline is an oral partial agonist on the $\alpha_4\beta_2$ subtype of the nicotinic acetylcholine receptor. It stimulates the nicotine receptor and reduces withdrawal symptoms and also the craving for cigarettes. A 12-week course increases the chance of stopping smoking four times; its main side-effect is nausea and severe depression has been recorded.

AIR POLLUTION AND EPIDEMIOLOGY

Atmospheric air pollution, due to the burning of coal for energy and heat, has been a feature of urban living in developed countries for at least two centuries. It consists of black smoke and sulphur dioxide (SO_2). Air pollution of this type peaked in the 1950s in the UK, until legislation led to restrictions on coal burning. Such pollution remains common in Eastern Europe and Russia and is increasing in newly industrialized countries (especially India and China). The combustion of petroleum and diesel oil in motor vehicles has led to new air pollution, consisting of primary pollutants such as nitrogen oxides (NO and NO_2), diesel particulates, polyaromatic hydrocarbons and the secondary pollutant ozone (O_3) generated by photochemical reactions in the atmosphere (ozone levels are highest in sunny, rural areas). Levels of NO_2 can be high in poorly ventilated kitchens and living rooms where gas is used for cooking and in fires.

Particulate matter, categorized by size, consists of coarse particles (10–2.5 μm in aerodynamic diameter), produced by construction work and farming, and fine particles (<2.5 μm) generated from burning fossil fuels. The latter, designated $PM_{2.5}$, are largely responsible for air pollution causing health problems, e.g. respiratory and cardiac disease. These very small particles remain airborne for long periods and are carried into rural areas.

The WHO global air-quality guidelines suggest 24 h values of <25 μg/m^3 for $PM_{2.5}$ for the short term and 10 μg/m^3 in the long term. In Europe 70% of the particulates present in urban air result from the combustion of diesel fuel, providing a background concentration of 3–5 μg/m^3. The WHO estimates that air pollution causes 800 000 premature deaths world-wide every year.

Deaths from respiratory and cardiovascular disease occur mainly in older populations, and air pollution causes symptoms of bronchitis in children. Pollution from motor vehicles has also been shown to cause increased hospital admissions, reduced lung function in children and younger adults and an increase in lung cancer (polyaromatic hydrocarbons).

It has been proposed at various times that air pollution is one of the causes of the recent increase in asthma and other allergic diseases (Table 14.8). There is no current evidence that this is true, but both NO_2 and ozone have been shown to enhance the nasal and lung airway responses to inhaled allergen in those with established allergic disease. Air pollution does, however, have an adverse effect on lung development in teenage children.

Management

When air quality is poor, asthmatics are advised to avoid exercising outdoors and to increase their anti-inflammatory medication (i.e. inhaled corticosteroids).

Short- and long-term measures are required to reduce air pollution, particularly diesel particulates (which are predicted to increase as more diesel engines are used). Such measures include increased motor engine efficiency, catalytic converters, diesel particulate traps and decreased reliance on cars and trucks.

FURTHER READING

Hales S, Howden-Chapman P. Effects of air pollution on health. *British Medical Journal* 2007; **335**: 314–315.

Gu Dongfeng, Kelly TN, Wu Xigui, et al. Mortality attributable to smoking in China. *New England Journal of Medicine* 2009; **360**: 150–159.

Lippmann M. Health effects of airborne particulate matter. *New England Journal of Medicine* 2007; **57**: 2395–2397.

WHO. *WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Global update 2005.* Copenhagen: World Health Organization, 2005.

Table 14.8 Air pollutants and their health effects

	Average concentration	Poor air quality	Susceptible individuals	Mechanisms of health effects
Sulphur dioxide (SO_2)	5–15 ppb	>125 ppb	Asthmatics	Bronchoconstriction through neurogenic mechanism
Ozone (O_3)	10–30 ppb	>90 ppb	All affected, particularly during exercise	Restrictive lung defect Airway inflammation Enhanced response to allergen
Nitrogen dioxide (NO_2)	25–40 ppb	>100 ppb	Allergic individuals	Airway inflammation Enhanced response to allergen
Particulate matter (PM_{10})	25–30 μg/m^3	>65 μg/m^3	Elderly Allergic individuals	Airway and alveolar inflammation Enhanced production selectively of the allergy antibody (IgE)
($PM_{2.5}$)	3–5 μg/m^3	>25 μg/m^3	With cardiac and respiratory disease	Airway inflammation

ppb, parts per billion.

DISEASES OF THE UPPER RESPIRATORY TRACT

The common cold (acute coryza)

This highly infectious illness causes a mild systemic upset and prominent nasal symptoms. It is due to infection by a wide range of respiratory viruses, of which the rhinoviruses are the most common. Other common cold viruses include coronaviruses and adenoviruses. Infectivity from close personal contact (nasal mucus on hands) or droplets is high in the early stages of the infection, and spread is facilitated by overcrowding and poor ventilation. There are at least 100 different antigenic strains of rhinovirus, making it difficult for the immune system to confer protection. On average, individuals suffer two to three colds per year, but the incidence lessens with age, presumably as a result of accumulating immunity to different virus strains. The incubation period varies from 12 hours to 5 days.

The clinical features are tiredness, slight pyrexia, malaise and a sore nose and pharynx. Sneezing and profuse, watery nasal discharge are followed by thick mucopurulent secretions which may persist for up to a week. Secondary bacterial infection occurs only in a minority of cases.

Sinusitis (see p. 1081)

Rhinitis

Rhinitis is defined clinically as sneezing attacks, nasal discharge or blockage occurring for more than an hour on most days for:

- a limited period of the year (seasonal or intermittent rhinitis)
- throughout the whole year (perennial or persistent rhinitis).

Seasonal rhinitis

This is often called 'hayfever' and is the most common allergic disorder. Since the term hayfever implies that only grass pollen is responsible, it is better described as seasonal (or intermittent) allergic rhinitis. World-wide prevalence rates vary from 2% to 20%. Prevalence is maximal in the second decade, and up to 30% of young British people suffer symptoms in June and July.

Nasal irritation, sneezing and watery rhinorrhoea are the most troublesome symptoms, but many also suffer from itching of the eyes and soft palate and occasionally even itching of the ears because of the common innervation of the pharyngeal mucosa and the ear. In addition, approximately 20% suffer from seasonal wheezing. The common seasonal allergens are shown in Figure 14.19. Since pollination of plants that give rise to high pollen counts varies from country to country, seasonal rhinoconjunctivitis and accompanying wheeze may occur at different times of year in different regions.

Perennial rhinitis

Patients with perennial rhinitis rarely have symptoms that affect the eyes or throat. Half have symptoms predominantly of sneezing and watery rhinorrhoea, whilst the other half complain mostly of nasal blockage. The patient may lose the sense of smell and taste. Sinusitis occurs in about 50% of cases, due to mucosal swelling that obstructs drainage from the sinuses. Perennial rhinitis is most frequent in the second and third decades, decreasing with age, and can be divided into four main types.

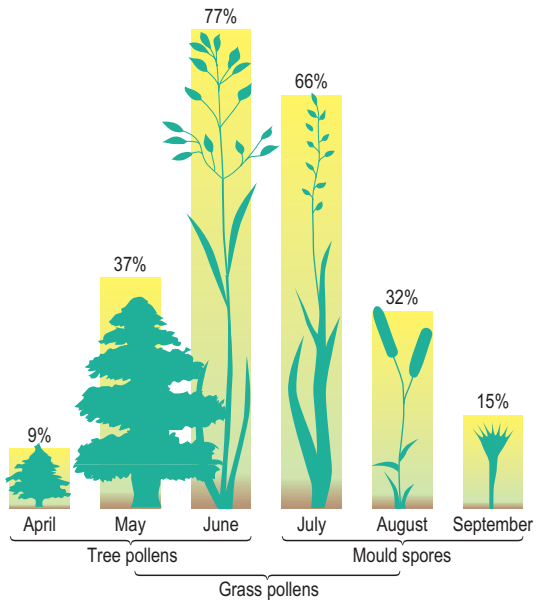


Fig. 14.19 **Seasonal allergic rhinitis.** Bar graph showing the proportion of patients whose symptoms are worst in the month or months indicated. The causative agents are also shown.

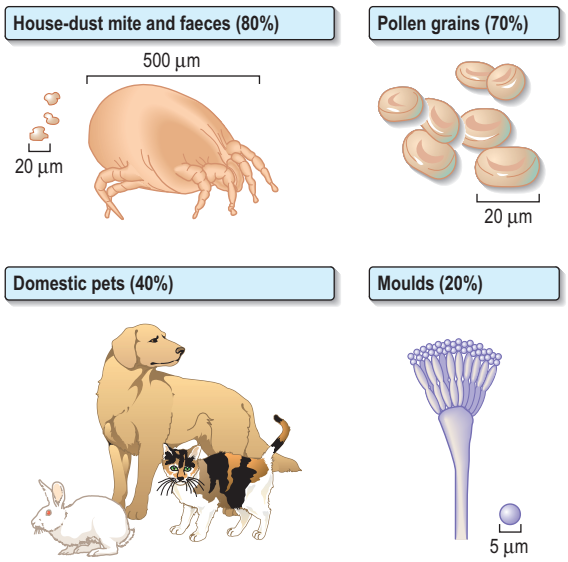


Fig. 14.20 **Common allergens causing allergic rhinitis and asthma.** The house-dust mite, faeces of house-dust mites, pollen grains, domestic pets and moulds. Percentages are those of positive skin-prick tests to these allergens in patients with allergic rhinitis.

Perennial allergic rhinitis

The major cause of this is allergy to the faecal particles of the house-dust mite *Dermatophagoides pteronyssinus* or *D. farinae*; these particles are approximately 20 µm in diameter (Fig. 14.20), not dissimilar in size to pollen grains. The house-

dust mite itself is under 0.5 mm in size, invisible to the naked eye (Fig. 14.20), and is found in dust throughout the house, particularly in older, damp dwellings. It lives off desquamated human skin scales and the highest concentrations (4000 mites per gram of surface dust) are found in human bedding.

The next most common allergens come from domestic pets (especially cats) and are proteins derived from urine or saliva spread over the surface of the animal as well as skin protein. Allergy to urinary protein from small mammals is a major cause of morbidity amongst laboratory workers.

Industrial dust, vapours and fumes cause occupationally related perennial rhinitis more often than asthma.

The presence of perennial rhinitis makes the nose more reactive to non-specific stimuli such as cigarette smoke, washing powders, household detergents, strong perfumes and traffic fumes. Although patients often think they are allergic to these stimuli, these are irritant responses and do not involve allergic immune reactions.

Perennial non-allergic rhinitis with eosinophilia

No extrinsic allergic cause can be identified, either from the history or on skin testing, but, as in patients with perennial allergic rhinitis, eosinophilic granulocytes are present in nasal secretions. Most of these patients are intolerant of aspirin/NSAIDs.

Vasomotor rhinitis

These patients with perennial rhinitis have no demonstrable allergy or nasal eosinophilia. Watery secretions and nasal congestion are triggered by, for example, cold air, smoke, perfume, newsprint, possibly because of an imbalance of the autonomic nerves controlling the erectile tissue (sinusoids) in the nasal mucosa.

Nasal polyps

These are round, smooth, soft, semi-translucent, pale or yellow, glistening structures attached to the sinus mucosa by a relatively narrow stalk or pedicle, occurring in patients with both allergic and vasomotor rhinitis. They contain mast cells, eosinophils and mononuclear cells in large numbers and cause nasal obstruction, loss of smell and taste, and mouth breathing, but rarely sneezing, since the mucosa of the polyp is largely denervated. The mechanism(s) of their formation is not known.

Pathogenesis

Sneezing, increased secretion and changes in mucosal blood flow are mediated both by efferent nerve fibres and by released mediators (see p. 850). Mucus production results largely from parasympathetic stimulation. Blood vessels are under both sympathetic and parasympathetic control. Sympathetic fibres maintain tonic contraction of blood vessels, keeping the sinusoids of the nose partially constricted with good nasal patency. Stimulation of the parasympathetic system dilates these blood vessels. This stimulation varies spontaneously in a cyclical fashion so that air intake alternates slowly over several hours from one nostril to the other. The erectile cavernous nasal sinusoids can be influenced by emotion, which, in turn, can affect nasal patency.

Allergic rhinitis develops as a result of interaction between the inhaled allergen and IgE molecules present on the surface of mast cells which are found in increased numbers in nasal secretions and within the nasal epithelium. Release of preformed mediators, in particular histamine, causes an increase in permeability of the epithelium. Sneezing results from stimulation of afferent nerve endings (mostly via histamine) and begins within minutes of the allergen entering the nose. This is followed by nasal exudation and secretion and eventually nasal blockage peaking 15–20 minutes after contact with the allergen.

B cells produce IgE antibody against the allergen. IgE binds to mast cells via high affinity cell surface receptors, causing degranulation and release of histamine, proteases (tryptase, chymase), prostaglandins (PGDs), cysteinyl leukotriene (LTC4), and cytokines. This causes the acute symptoms of sneezing, itch, rhinorrhoea and nasal congestion.

Additionally, allergens are also presented to T cells via antigen presenting cells (dendritic cells). This causes a release of IL-4 and IL-13 which further stimulate the B cells and also IL-5, IL-9 and GM-CSF, switching from a Th1 to a Th2 response to activate and recruit eosinophils, basophils, neutrophils and T lymphocytes. These cause the late phase nasal reactions of chronicity with nasal obstruction, hyperactivity and anosmia.

Investigations and diagnosis

The allergic factors causing rhinitis are usually obtained from the history.

Skin-prick testing indicates that the mechanisms leading to allergic rhinitis (or asthma) are present in human skin. A positive test does not necessarily mean that the particular allergen producing the weal causes the respiratory disease. However, if there is a positive clinical history for that allergen, a causative role is likely. Measuring specific serum IgE antibodies against the particular allergen provides the same information as the skin-prick test. Blood tests are much more expensive and should be reserved for use in patients who cannot be skin tested for some reason (e.g. dermatographism, active eczema or using antihistamines and unable to stop for 3 days before skin tests).

Treatment

Allergen avoidance

Removal of a household pet or total enclosure of industrial processes releasing sensitizing agents can lead to cure of rhinitis and, indeed, asthma.

Pollen avoidance is impossible. Contact may be diminished by wearing sunglasses, driving with the car windows shut, avoiding walks in the countryside (particularly in the late afternoon when the number of pollen grains is highest at ground level), and keeping the bedroom window shut at night. These measures are rarely sufficient in themselves to control symptoms. Exposure to pollen is generally lower in coastal regions, where sea breezes keep pollen grains inland.

The house-dust mite infests most areas of the house, but particularly the bedroom. Mite counts are extremely low in hospitals where carpets are absent, floors are cleaned frequently and mattresses and pillows are covered in plastic sheeting that can be wiped down. Mite allergen exposure can be reduced by enclosing bedding in fabric specifically designed to prevent the passage of mite allergen, while allowing water vapour through. This is both comfortable and reduces symptoms. Acaricides are less effective and cannot be recommended. Increased room ventilation and reduced soft furnishings including carpets, curtains and soft toys are all helpful in reducing the mite load. However,

a Cochrane collaboration review in asthma has shown that house-dust mite control measures are of no value in asthma (p. 853).

H_1 antihistamines

Antihistamines remain the most common therapy for rhinitis, and many can be purchased directly over the counter in the UK. They are particularly effective against sneezing, but are less effective against rhinorrhoea and have little influence on nasal blockage. First-generation antihistamines (chlorphenamine, hydroxyzine) cause sedation and loss of concentration in all patients (including those who are not aware of the problem) and should no longer be used. Second-generation drugs such as loratadine (10 mg once daily), desloratadine (5 mg daily), cetirizine 10 mg daily and L-cetirizine 5 mg daily, and fexofenadine (120 mg daily) are at least as potent; they do not cause sedation. Antihistamines also control itching in the eyes and palate.

Decongestants

Drugs with sympathomimetic activity (α-adrenergic agents) are widely used for the treatment of nasal obstruction. They may be taken orally or more commonly as nasal drops or sprays (e.g. ephedrine nasal drops). Xylometazoline and oxymetazoline are widely used because they have a prolonged action and tachyphylaxis does not develop. Secondary nasal hyperaemia can occur some hours later as a rebound effect and rhinitis medicamentosa can develop if patients go on taking increasing quantities of the local decongestant to overcome this phenomenon. Although local decongestants are an effective treatment for vasomotor rhinitis, patients must be warned about rebound nasal obstruction and should use the drug carefully. Usually, such preparations should be prescribed for only a limited period to open the nasal airways for administration of other therapy, such as topical corticosteroids.

Anti-inflammatory drugs

Sodium cromoglicate and nedocromil sodium influence a number of aspects of inflammation, including mast cell and eosinophil activation and nerve function. They act by blocking an intracellular chloride channel and preventing cell activation. Topical sodium cromoglicate is of limited value in the treatment of allergic rhinitis, although both it and nedocromil sodium are very effective in the management of allergic conjunctivitis.

Corticosteroids

The most effective treatment for rhinitis is a topical corticosteroid preparation (e.g. beclometasone spray twice daily or fluticasone propionate spray once daily). The amount used is insufficient to cause systemic effects and the effect is primarily anti-inflammatory. Topical steroids should be started before the beginning of seasonal symptoms. The combination of a topical corticosteroid with a non-sedating antihistamine taken regularly is particularly effective. Patients should be carefully instructed in how to use the nasal steroid device to achieve optimal drug deposition. In selected cases, an α-adrenergic agonist may help to decongest the nose prior to taking the topical corticosteroid.

If other therapy has failed, seasonal and perennial rhinitis respond readily to a short course (2 weeks) of treatment with oral prednisolone 5–10 mg daily. Nasal polyps may respond to oral corticosteroids and their recurrence may be prevented by continuous application of topical corticosteroids.

Leukotriene antagonists

In patients who do not respond to antihistamines or topical steroids, a leukotriene antagonist (e.g. montelukast 10 mg daily in the evening or zafirlukast 20 mg twice daily) may be helpful, especially in those with a history of NSAID sensitivity or concomitant asthma.

Immunotherapy

This is used for patients with seasonal allergic rhinitis who have not responded to the above. An oral preparation of grass pollen extract is available, as well as an extract that is given in gradually increasing doses subcutaneously (p. 72).

Vaccines containing allergic peptides coupled to synthetic bacterial DNA sequences have been developed. These compounds, which bind to toll-like receptor-9, have shown clinical efficacy, e.g. against ragweed allergy.

Pharyngitis

The most common viruses causing pharyngitis are adenoviruses, of which there are about 32 serotypes. Endemic adenovirus infection causes the common sore throat, in which the oropharynx and soft palate are reddened and the tonsils are inflamed and swollen. Within 1–2 days the tonsillar lymph nodes enlarge. Occasionally, localized epidemics due to adenovirus serotype 8 occur, particularly in schools during summer, with episodes of fever, conjunctivitis, pharyngitis and lymphadenitis of the neck glands. The disease is self-limiting and only requires symptomatic treatment without antibiotics.

In recent years the proportion of sore throats due to bacterial infections, e.g. haemolytic streptococcus, has fallen. Persistent and severe tonsillitis requires antibiotic therapy (phenoxymethylpenicillin 500 mg four times a day or cefaclor 250 mg three times daily). Amoxicillin and ampicillin should be avoided if there is a possibility of infectious mononucleosis (p. 107).

Acute laryngotracheobronchitis

Acute laryngitis is an occasional but striking complication of upper respiratory tract infections, particularly those caused by parainfluenza viruses and measles. Inflammatory oedema extends to the vocal cords and the epiglottis, causing considerable narrowing of the airway; in addition, there may be associated *tracheitis* or *tracheobronchitis*. Children under the age of 3 years are most severely affected. The voice becomes hoarse, with a barking cough (croup) and there is audible laryngeal stridor. Progressive airways obstruction may occur, with recession of the soft tissue of the neck and abdomen during inspiration, and in severe cases, central cyanosis. Inhalation of steam is not helpful. Oral or i.m. corticosteroids should be given (e.g. dexamethasone) and nebulized adrenaline (epinephrine) gives short-term relief. With the use of steroids, endotracheal intubation is rarely necessary. Oxygen and adequate fluids should be given. Rarely, a tracheostomy is required.

FURTHER READING

Bjornson CL, Johnson DW. Croup. *Lancet* 2008; **371**: 329–339.

Acute epiglottitis

H. influenzae type b (Hib) can cause life-threatening infection of the epiglottis, usually in children under 5 years of age. The child becomes extremely ill with a high fever, and severe airflow obstruction may rapidly occur. This is a life-threatening emergency and requires urgent endotracheal intubation and intravenous antibiotics (e.g. ceftazidime 25–150 mg/kg). Chloramphenicol (50–100 mg/kg) can also be used. The epiglottis, which is red and swollen, should not be inspected until facilities to maintain the airways are available.

Other manifestations of Hib infection are meningitis, septic arthritis and osteomyelitis. A highly effective vaccine is now available, which is given to infants at 2, 3 and 4 months with their primary immunizations against diphtheria, tetanus and pertussis (DTP). This programme has reduced death rates from Hib infections virtually to zero in many countries.

Influenza (see also p. 116)

The influenza virus belongs to the orthomyxovirus group and exists in two main forms, A and B. Influenza B is associated with localized outbreaks of milder nature, whereas influenza A causes world-wide pandemics. Influenza A can develop new antigenic variants at irregular intervals. Human immunity develops against the haemagglutinin (H) antigen and the neuraminidase (N) antigen on the viral surface. Major shifts in the antigenic make-up of influenza A viruses provide the necessary conditions for major pandemics, whereas minor antigenic drifts give rise to less severe epidemics because immunity in the population is less blunted.

The most serious influenza pandemic occurred in 1918, and was associated with more than 20 million deaths world-wide. In 1957, a major shift in the antigenic make-up of the virus led to the appearance of influenza A2 type H2N2, which caused a world-wide pandemic. A further pandemic occurred in 1968 owing to the emergence of Hong Kong influenza type H3N2, and minor antigenic drifts have caused outbreaks around the world ever since. A further pandemic is likely in the next few years. Recent concerns have focused on the H5N1 strain of influenza A which was previously confined to birds but can be contracted by humans who are in close contact with infected poultry.

Clinical features

The incubation period of influenza is usually 1–3 days. The illness starts abruptly with a fever, shivering and generalized aching in the limbs. This is associated with severe headache, soreness of the throat and a dry cough that can persist for several weeks. Diarrhoea occurs in 70% of cases of H5N1. Influenza infection can be followed by a prolonged period of debility and depression that may take weeks or months to clear.

Complications

Secondary bacterial infection is common following influenza virus infection, particularly with *Strep. pneumoniae* and *H. influenzae*. Secondary pneumonia caused by *Staph. aureus* is rarer, but more serious, and carries a mortality of up to 20%. Postinfectious encephalomyelitis rarely occurs after influenza infection.

Diagnosis and treatment

Laboratory diagnosis is not usually necessary, but a definitive diagnosis can be established by demonstrating a fourfold increase in the complement-fixing antibody or the haemagglutinin antibody measured at onset and after 1–2 weeks or by demonstrating the virus in throat or nasal secretions.

Treatment is by bed rest and paracetamol, with antibiotics to prevent secondary infection in those with chronic bronchitis, cardiac or renal disease.

Neuraminidase inhibitors may help to shorten the duration of symptoms in patients with influenza, if given within 48 hours of the first symptom. The cost–benefit of zanamivir and oseltamivir remains unproven but these are currently recommended in the UK for patients with suspected influenza over the age of 65 and 'at-risk' adults, as part of a strategy to reduce admissions to hospital when influenza is circulating in the community.

Prophylaxis

Protection by influenza vaccines is only effective in up to 70% of people and only lasts for about a year. Influenza vaccine should not be given to individuals who are allergic to egg protein as some are manufactured in chick embryos. New vaccines have to be prepared to cover each change in viral antigenicity and are therefore in limited supply at the start of an epidemic. Nevertheless, routine vaccination is recommended for all individuals over 65 years of age and also for younger people with chronic heart disease, chronic lung disease (including asthma), chronic kidney disease, diabetes mellitus and those who are immunosuppressed. During pandemics key hospital and health service personnel should also be vaccinated.

Inhalation of foreign bodies

Children inhale foreign bodies, e.g. peanuts, more commonly than do adults. In the adult, inhalation may occur after excess alcohol or under general anaesthesia (loose teeth or dentures).

When the foreign body is large it may impact in the trachea. The person chokes and then becomes silent; death ensues unless the material is quickly removed (see Emergency box 14.1).

Emergency Box 14.1

Treatment of inhaled foreign bodies (Heimlich manoeuvre)

Emergency
The Heimlich manoeuvre is used to expel the obstructing object:

1. Stand behind the patient.
2. Encircle your arms around the upper part of the abdomen just below the patient's rib cage.
3. Give a sharp, forceful squeeze, forcing the diaphragm sharply into the thorax. This should expel sufficient air from the lungs to force the foreign body out of the trachea.

Non-emergency
Rigid bronchoscopy should be performed.

Impaction usually occurs in the right main bronchus and produces:

- choking
- persistent monophonic wheeze
- later, persistent suppurative pneumonia
- lung abscess (common).

DISEASES OF THE LOWER RESPIRATORY TRACT

Lower respiratory tract infection accounts for approximately 10% of the world-wide burden of morbidity and mortality. Seventy five per cent of all antibiotic usage is for these diseases, despite the fact that they are mainly due to viruses.

Acute bronchitis

Acute bronchitis in previously healthy subjects is often viral. Bacterial infection with *Strep. pneumoniae* or *H. influenzae* is a common sequel to viral infections, and is more likely to occur in cigarette smokers or people with chronic obstructive pulmonary disease (COPD).

The illness begins with an irritating, unproductive cough, together with discomfort behind the sternum. This may be associated with tightness in the chest, wheezing and shortness of breath. The cough becomes productive, with yellow or green sputum. There is a mild fever and a neutrophil leucocytosis; wheeze with occasional crackles can be heard on auscultation. In otherwise healthy adults the disease improves spontaneously in 4–8 days without the patient becoming seriously ill.

Antibiotics are usually given (e.g. amoxicillin 250 mg three times daily), although it is not known whether this hastens recovery in otherwise healthy individuals.

Chronic obstructive pulmonary disease (COPD)

COPD is predicted to become the third most common cause of death and fifth most common cause of disability world-wide by 2020.

The term 'chronic obstructive pulmonary disease' (COPD) was introduced to bring together a variety of clinical syndromes associated with destruction of the lung and airflow obstruction. The terms 'chronic obstructive airways disease' (COAD) and 'chronic obstructive lung disease' (COLD) have been used as synonyms in different parts of the world. Prior to 1979, patients with these conditions were often classified by symptoms (chronic bronchitis, chronic asthma), by pathological changes (emphysema) or by physiological correlates (pink puffers, blue bloaters). Recognition that these entities overlapped and often coexisted led to the term COPD.

Recently it has been realized that COPD is associated with a number of co-morbidities, e.g. ischaemic heart disease, hypertension, diabetes, heart failure and cancer, suggesting that there is a generalized systemic inflammatory process.

Definition of COPD

A recent operational definition states: 'COPD is a disease state characterized by airflow limitation that is not fully reversible. The airflow limitation is usually both progressive and associated with an abnormal inflammatory response of the lungs to noxious particles or gases'.

Epidemiology and aetiology

COPD is caused by long-term exposure to toxic particles and gases. In developed countries, cigarette smoking accounts for over 90% of cases. In developing countries other factors, such as the inhalation of smoke from biomass fuels used in heating and cooking in poorly ventilated areas, are also implicated. However, only 10–20% of heavy smokers develop COPD, indicating individual susceptibility. The development of COPD is proportional to the number of cigarettes smoked per day; the risk of death from COPD in patients smoking 30 cigarettes daily is 20 times that of a non-smoker. Autopsy studies have shown substantial numbers of centri-acinar emphysematous spaces in the lungs of 50% of British smokers over the age of 60 years independent of the diagnosis of significant respiratory disease before death.

Climate and air pollution play a smaller role, but the mortality from COPD increases dramatically during periods of heavy atmospheric pollution (p. 830). Urbanization, social class and occupation may also play a part in aetiology, but these effects are difficult to separate from that of smoking. Some animal studies suggest that diet could be a risk factor for COPD, but this has not been proven in humans.

The socio-economic burden of COPD is considerable. In the UK, COPD causes approximately 18 million lost working days annually for men and 2.1 million lost working days for women, accounting for approximately 7% of all days of absence from work due to sickness. Nevertheless, the number of admissions to hospital in the UK with COPD has been falling steadily over the last 25 years.

Pathophysiology

The most consistent pathological finding in COPD is increased numbers of mucus-secreting goblet cells in the bronchial mucosa, especially in the larger bronchi (Fig. 14.21). In more advanced cases, the bronchi become overtly inflamed and pus is seen in the lumen.

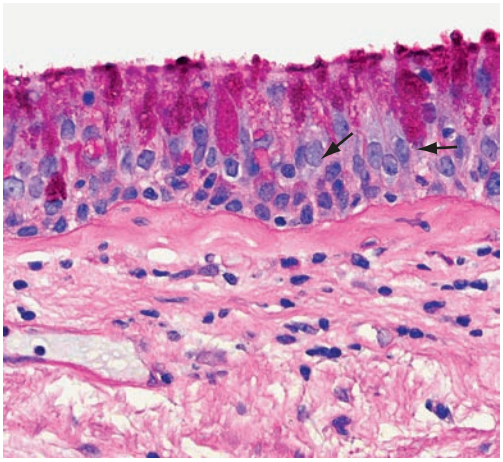


Fig. 14.21 **COPD.** Section of bronchial mucosa stained for mucus glands by PAS showing increase in mucus-secreting goblet cells. Courtesy of Dr J Wilson and Dr S Wilson, University of Southampton.

FURTHER READING

Mizgerd JP. Acute lower respiratory tract infection. *New England Journal of Medicine* 2008; **358**: 716–727.

Microscopically there is infiltration of the walls of the bronchi and bronchioles with acute and chronic inflammatory cells; lymphoid follicles may develop in severe disease. In contrast to asthma, the lymphocytic infiltrate is predominantly $CD8^+$. The epithelial layer may become ulcerated and, when the ulcers heal, squamous epithelium replaces the columnar cells. The inflammation is followed by scarring and thickening of the walls which leads to widespread narrowing in the small airways (Fig. 14.22).

The small airways are particularly affected early in the disease, initially without the development of any significant breathlessness. This initial inflammation of the small airways is reversible and accounts for the improvement in airway function if smoking is stopped early. In later stages the inflammation continues, even if smoking is stopped.

Further progression of the airways disease leads to progressive squamous cell metaplasia, and fibrosis of the bronchial walls. The physiological consequence of these changes is the development of airflow limitation. If the airway narrowing is combined with emphysema (causing loss of the elastic recoil of the lung with collapse of small airways during expiration) the resulting airflow limitation is even more severe.

Emphysema is defined pathologically as dilatation and destruction of the lung tissue distal to the terminal bronchiole. It is classified according to the site of damage:

- *Centri-acinar emphysema*. Distension and damage of lung tissue is concentrated around the respiratory bronchioles, whilst the more distal alveolar ducts and alveoli tend to be well preserved. This form of emphysema is extremely common; when of modest extent, it is not necessarily associated with disability. Severe centri-acinar emphysema is associated with substantial airflow limitation.
- *Pan-acinar emphysema*. This is less common. Distension and destruction appear to involve the whole of the acinus, and in the extreme form the lung becomes a mass of bullae. Severe airflow limitation and $\dot{V}_A/\dot{Q}$ mismatch occur. This type of emphysema occurs in α_1-antitrypsin deficiency (see p. 837).
- *Irregular emphysema*. There is scarring and damage affecting the lung parenchyma patchily without particular regard for acinar structure.

Emphysema leads to expiratory airflow limitation and air trapping. The loss of lung elastic recoil results in an increase in TLC, and premature closure of airways limits expiratory flow while the loss of alveoli results in decreased capacity for gas transfer.

$\dot{V}_A/\dot{Q}$ mismatch occurs partly because of damage and mucus plugging of smaller airways from chronic inflammation, and partly because of the rapid expiratory closure of the smaller airways owing to loss of elastic recoil from emphysema. This leads to a fall in P_aO_2 and an increase in the work of respiration.

CO_2 excretion is not impaired to the same extent and indeed many patients will show low normal P_aCO_2 values due to increasing their ventilation in an attempt to maintain normal blood gases by increasing their respiratory effort. Other patients fail to maintain their respiratory effort and as a consequence their CO_2 levels increase. In the short term, this rise in CO_2 leads to stimulation of respiration but in the long term, these patients often become insensitive to CO_2 and come to depend on hypoxaemia to drive their ventilation. These patients appear less breathless and because of their reduced O_2 saturation, they start to retain fluid and stimulate erythrocyte production (leading to polycythaemia). In consequence they become bloated, plethoric and cyanosed, the typical appearance of the 'blue bloater'. Attempts to abolish hypoxaemia by administering oxygen can make the situation much worse by decreasing respiratory drive in these patients who depend on hypoxia to drive their ventilation.

The classic Fletcher and Peto studies (Fig. 14.23) show that there is a loss of 50 mL per year in FEV_1 in COPD compared to 20 mL per year in healthy people.

In *summary*, three mechanisms have been suggested for this limitation of airflow in small airways (<2 mm in diameter).

- Loss of elasticity and alveolar attachments of airways due to emphysema. This reduces the elastic recoil and the airways collapse during expiration.
- Inflammation and scarring cause the small airways to narrow.
- Mucus secretion which blocks the airways.

Each mechanism narrows the small airways and causes air trapping leading to hyperinflation of the lungs and breathlessness.

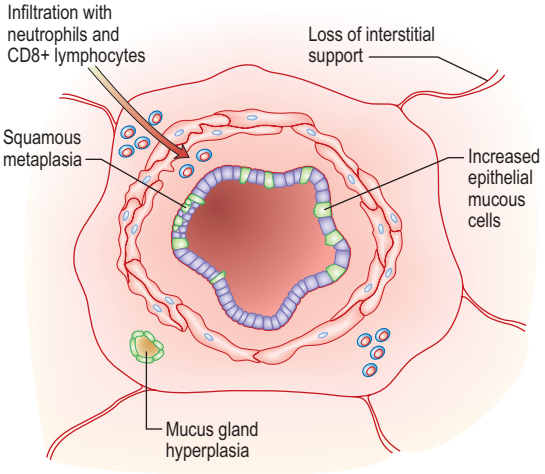


Fig. 14.22 **Pathological changes in the airways in chronic bronchitis and emphysema.**

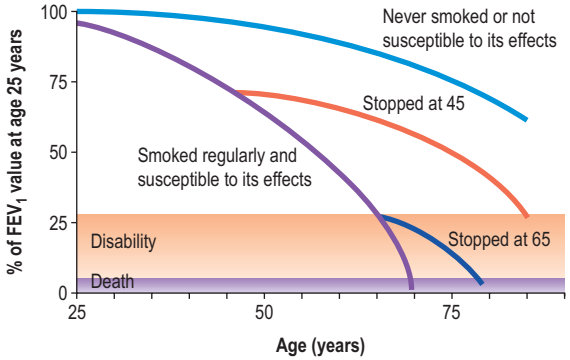


Fig. 14.23 **Influence of smoking on airflow limitation.** From Fletcher CM, Peto R. *British Medical Journal* 1977; **1**: 1645.

Pathogenesis

Cigarette smoking

Bronchoalveolar lavage and biopsies of the airways of smokers show increased numbers of neutrophil granulocytes. These granulocytes can release elastases and proteases, which may help to produce emphysema. It has been suggested that an imbalance between protease and antiprotease activity causes the damage. α_1-Antitrypsin is a major serum antiprotease which can be inactivated by cigarette smoke (see below).

Mucous gland hypertrophy in the larger airways is thought to be a direct response to persistent irritation resulting from the inhalation of cigarette smoke. The smoke has an adverse effect on surfactant, favouring overdistension of the lungs.

Infections

Patients with COPD cope badly with respiratory infections, which are often the precipitating cause of acute exacerbations of the disease. However, it is less clear whether infection is relevant in the development of the progressive airflow limitation that characterizes disabling COPD. Prompt use of antibiotics and routine vaccinations against influenza and pneumococci are appropriate.

α_1-Antitrypsin deficiency (see also p. 358)

α_1-Antitrypsin is a proteinase inhibitor which is produced in the liver, secreted into the blood and diffuses into the lung. Here it inhibits proteolytic enzymes such as neutrophil elastase, which are capable of destroying alveolar wall connective tissue.

More than 75 alleles of the α_1-antitrypsin gene have been described. The three main phenotypes are MM (normal), MZ (heterozygous deficiency) and ZZ (homozygous deficiency). About 1 child in 5000 in Britain is born with the homozygous deficiency, but not all develop chest disease. Those who do develop breathlessness under the age of 40 years have radiographic evidence of basal emphysema and are usually, but not always, cigarette smokers. Hereditary deficiency of α_1-antitrypsin accounts for about 2% of emphysema cases. A small minority develop liver disease (see p. 358).

Clinical features

Symptoms

The characteristic symptoms of COPD are productive cough with white or clear sputum, wheeze and breathlessness, usually following many years of a smoker's cough. Colds seem to 'settle on the chest' and frequent infective exacerbations occur, giving purulent sputum. Symptoms can be worsened by factors such as cold, foggy weather and atmospheric pollution. With advanced disease, breathlessness becomes severe even after mild exercise such as getting dressed. Apart from the pulmonary features, there are systemic effects on cardiovascular function e.g. hypertension, as well as osteoporosis, depression, metabolic problems leading to weight loss and loss of muscle mass with weakness.

Signs

In mild disease there may be no signs or nothing more than quiet wheezes throughout the chest. In severe disease, the patient is tachypnoeic, with prolonged expiration. The accessory muscles of respiration are used and there may be intercostal indrawing on inspiration and pursing of the lips on expiration (see p. 816). Chest expansion is poor, the lungs are hyperinflated, and there is loss of the normal cardiac and liver dullness.

Patients who remain responsive to CO_2 are usually breathless and rarely cyanosed. Heart failure and oedema are rare features except as terminal events. In contrast, patients who become insensitive to CO_2 are often oedematous and cyanosed but not particularly breathless. Those with hypercapnia may have peripheral vasodilatation, a bounding pulse, and when the P_{aCO_2} is above about 10 kPa, a coarse flapping tremor of the outstretched hands. Severe hypercapnia will lead to confusion and progressive drowsiness. At this stage papilloedema may be present but this is neither specific nor sensitive as a diagnostic feature.

Respiratory failure

The later stages of COPD are characterized by the development of respiratory failure. For practical purposes this is said to occur when there is either a P_{aO_2} of less than 8 kPa (60 mmHg) or a P_{aCO_2} of more than 7 kPa (55 mmHg) (see Ch. 15).

The persistence of chronic alveolar hypoxia and hypercapnia leads to constriction of the pulmonary arterioles and subsequent pulmonary arterial hypertension. Cardiac output is normal or increased but salt and fluid retention occurs as a result of renal hypoxia.

Cor pulmonale

Patients with advanced COPD may develop cor pulmonale (see p. 783), which is defined as heart disease secondary to disease of the lung. It is characterized by pulmonary hypertension, right ventricular hypertrophy, and eventually right heart failure. On examination, the patient is centrally cyanosed (owing to the lung disease) and, when heart failure develops, the patient becomes more breathless and ankle oedema occurs. Initially there may be a prominent parasternal heave, due to right ventricular hypertrophy, and a loud pulmonary second sound. In very severe pulmonary hypertension the pulmonary valve becomes incompetent. With right heart failure, tricuspid incompetence may develop with a greatly elevated jugular venous pressure (JVP), ascites and upper abdominal discomfort owing to swelling of the liver.

Diagnosis

This is usually clinical (GOLD criteria, Table 14.9). There is a history of breathlessness and sputum production in a lifetime smoker. In the absence of a history of cigarette smoking an initial working diagnosis of asthma is usual unless there is a family history of COPD suggesting α_1-antitrypsin deficiency.

The patient may have signs of hyperinflation and typical pursed lip respiration. No individual clinical feature is diagnostic. Emphysema is often incorrectly diagnosed on signs of overinflation of the lungs (e.g. loss of liver dullness on percussion), but this may occur with other diseases such as asthma. Furthermore, centri-acinar emphysema may be present without signs of overinflation. Some elderly men (without emphysema) develop a barrel-shaped chest as a result of osteoporosis of the spine, and a consequent decrease in height.

Investigations

- *Lung function tests* show evidence of airflow limitation (see Figs 14.8 and 14.15). The FEV_1:FVC ratio is reduced and the PEFR is low. In many patients the airflow limitation is reversible to some extent (usually a change in FEV_1 of <15%), and it can be difficult to distinguish between asthma and COPD. Lung volumes may be normal or increased, and the gas transfer coefficient of carbon monoxide is low when significant emphysema is present.
- *Chest X-ray* is often normal, even when the disease is advanced. The classic features are overinflation of the lungs with low, flattened diaphragms, and sometimes the presence of large bullae. Blood vessels may be 'pruned' with large proximal vessels and relatively little blood visible in the peripheral lung fields.
- *High resolution CT scans* are used, particularly to show emphysematous bullae.
- *Haemoglobin level and PCV* can be elevated as a result of persistent hypoxaemia (secondary polycythaemia, see p. 420). The CRP is raised.
- *Blood gases* are often normal at rest. In more advanced cases there is hypoxaemia and there may also be hypercapnia.
- *Sputum examination* is not useful in ordinary cases as *Strep. pneumoniae* and *H. influenzae* are the only common organisms to produce acute exacerbations. Occasionally *Moraxella catarrhalis* may cause infective exacerbations.
- *Electrocardiogram*. In advanced cor pulmonale the P wave is taller (P pulmonale) and there may be right bundle branch block (RSR′ complex) and the changes of right ventricular hypertrophy (see p. 784).
- *Echocardiogram* – useful to assess cardiac function where there is disproportionate dyspnoea.
- *α_1-Antitrypsin* levels and genotype, especially in premature disease or non-smokers.

Table 14.9 COPD – Global Initiative in Obstructive Lung Disease (GOLD) criteria

Stage of COPD		Function	Symptoms of breathlessness
I	Mild	FEV_1/FVC <70% FEV_1 ≥ 80% predicted	None or mild
II	Moderate	FEV_1/FVC <70% 50% of predicted ≤ FEV_1<80% of predicted	On exertion
III	Severe	FEV_1/FVC <70% 30% of predicted ≤ FEV_1<50% of predicted	On minimal exertion, e.g. dressing
IV	Very severe	FEV_1/FVC <70% FEV_1<30% predicted or FEV_1<50% predicted plus respiratory failure	At rest

Modified from the Global Strategy for the Diagnosis, Management and Prevention of COPD, 2006. www.goldcopd.com.

Management

See Figure 14.24 for management strategies.

Smoking cessation

It is essential to persuade the patient to stop smoking. Even at a late stage of the disease this may slow down the rate of deterioration and prolong the time before disability and death occur (see Fig. 14.24).

Drug therapy

This is used both for the short-term management of exacerbations and for the long-term relief of symptoms and with an

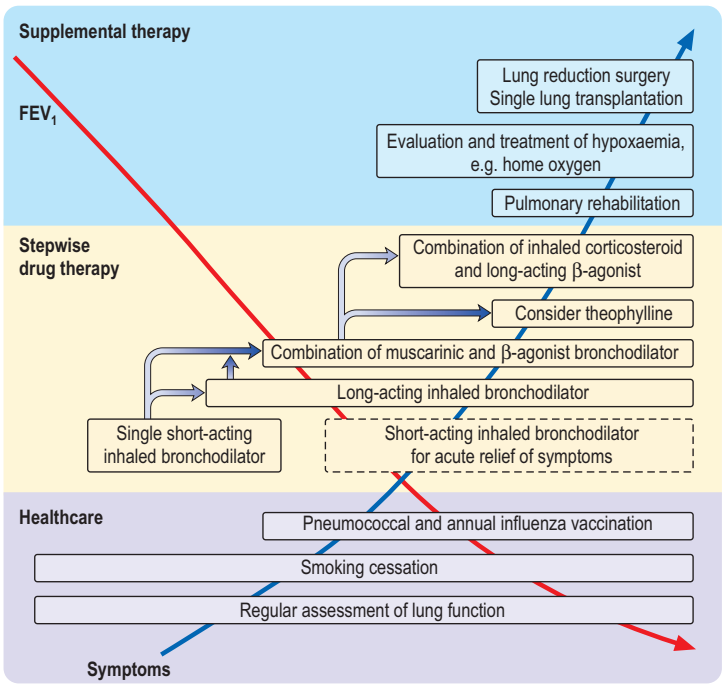


Fig. 14.24 **Algorithm for the treatment of COPD.** The various components of management are shown as the FEV_1 decreases and the symptoms become more severe. After Sutherland ER, Cherniack RM. Management of chronic obstructive pulmonary disease. *New England Journal of Medicine* 2004; **350**: 2689–2697. Copyright © 2004 Massachusetts Medical Society. All rights reserved.

FEV_1 of <60% predicted. Many of the drugs used are similar to those used in asthma (see p. 854).

Bronchodilators

Many patients with mild COPD feel less breathless after inhaling a β-adrenergic agonist such as salbutamol (200 μg every 4–6 hours). In more severe airway limitation (moderate and severe COPD), a long-acting β_2 agonist should be used, e.g. formoterol 12 μg powder inhaled twice daily or salmeterol 50 μg twice daily.

More prolonged and greater bronchodilatation is achieved with antimuscarinic agents: tiotropium (long-acting) (18 μg daily), ipratropium (40 μg four times daily) or oxitropium (200 μg twice daily). Tiotropium has been shown to improve function and quality of life but the decline in FEV_1 is unaffected. Patients find inhalers difficult to use and spacer devices improve delivery. Objective evidence of improvement in the peak flow or FEV_1 may be small and decisions to continue or stop therapy should be based mainly on the patient's reported symptoms. Long-acting preparations of theophylline are of little benefit.

Corticosteroids

In symptomatic patients with moderate/severe COPD, a trial of corticosteroids is always indicated, since a proportion of patients have a large, unsuspected, reversible element to their disease and airway function may improve considerably. Prednisolone 30 mg daily should be given for 2 weeks, with measurements of lung function before and after the treatment period. If there is objective evidence of a substantial degree of improvement in airflow limitation (FEV_1 increase >15%), prednisolone should be discontinued and replaced by inhaled corticosteroids (beclometasone 40 μg twice daily in the first instance, adjusted according to response). The long-term value of regular inhaled corticosteroids in all patients with COPD has not been proven.

Combination of a corticosteroid with a long-acting β_2 agonist has been shown in a trial to protect against a decline in lung function but there was no reduction in overall mortality.

Antibiotics

Prompt antibiotic treatment shortens exacerbations and should always be given in acute episodes as it may prevent hospital admission and further lung damage. Patients can be given a supply of antibiotics to keep at home to start as soon as their sputum turns yellow or green. Although amoxicillin-resistant *H. influenzae* is increasing (occurring in 10–20% of isolates from sputum) it is not a serious clinical problem. Resistance to cefaclor (500 mg 8-hourly) or cefixime (400 mg once daily) is significantly less frequent; co-amoxiclav is a useful alternative.

Long-term treatment with antibiotics remains controversial. They were once thought to be of no value, but eradication of infection and keeping the lower respiratory tract free of bacteria may help to prevent deterioration in lung function.

Antimucolytic agents

These reduce sputum viscosity and can reduce the number of acute exacerbations. A 4-week trial of carbocisteine 2.25 g daily can be tried.

Diuretic therapy (see p. 657)

This is necessary for all oedematous patients. Daily weights should be recorded during acute inpatient episodes.

Oxygen therapy

Two controlled trials (chiefly in males) have indicated that life can be prolonged by the continuous administration of oxygen at 2 L/min via nasal prongs to achieve an oxygen saturation of greater than 90% for large proportions of the day and night. Survival curves from these two studies are shown in Figure 14.25.

Only 30% of those not receiving long-term oxygen therapy survived for more than 5 years. A fall in pulmonary artery pressure was achieved if oxygen was given for 15 hours daily, but substantial improvement in mortality was only achieved by the administration of oxygen for 19 hours daily. These results suggest that long-term continuous domiciliary oxygen therapy will benefit patients who have:

- A P_aO_2 of <7.3 kPa (55 mmHg) when breathing air. Measurements should be taken on two occasions at least 3 weeks apart after appropriate bronchodilator therapy (Box 14.1).
- A P_aO_2 7.3–8 kPa with secondary polycythaemia, nocturnal hypoxaemia, peripheral oedema or evidence of pulmonary hypertension.
- Carboxyhaemoglobin of less than 3% (i.e. patients who have stopped smoking).

The provision of 19 hours of oxygen daily at a flow rate of 1–3 L/min using a 28% oxygen mask is best achieved using an oxygen concentrator. To achieve this with oxygen cylinders would require 20 standard cylinders per week, which is unacceptably expensive.

Nocturnal hypoxia

COPD patients with severe arterial hypoxaemia also suffer from profound nocturnal hypoxaemia which may drop the P_aO_2 as low as 2.5 kPa (19 mmHg), particularly during the rapid eye movement (REM) phase of sleep.

Because patients with COPD are already hypoxic, the fall in P_aO_2 produces a much larger fall in oxygen saturation

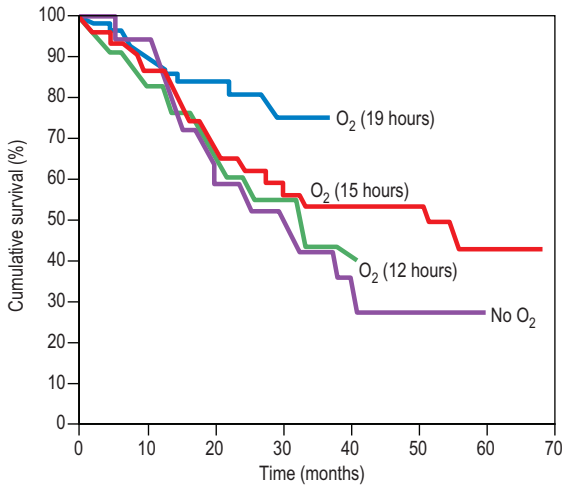


Fig. 14.25 **Cumulative survival curves for patients receiving oxygen.** Oxygen doses are in hours per day.

Box 14.1 Guidelines for domiciliary oxygen (Royal College of Physicians, 1999)

- Chronic obstructive pulmonary disease with P_aO_2 <7.3 kPa when breathing air during a period of clinical stability.
- Chronic obstructive pulmonary disease with P_aO_2 7.3–8 kPa in the presence of secondary polycythaemia, nocturnal hypoxaemia, peripheral oedema or evidence of pulmonary hypertension.
- Severe chronic asthma with a P_aCO_2 <7.3 kPa or persistent disabling breathlessness.
- Diffuse lung disease with P_aO_2 <8 kPa and in patients with P_aO_2 > 8 kPa with disabling dyspnoea.
- Cystic fibrosis when P_aO_2 <7.3 kPa or if P_aO_2 7.3–8 kPa in the presence of secondary polycythaemia, nocturnal hypoxaemia, pulmonary hypertension or peripheral oedema.
- Pulmonary hypertension without parenchymal lung involvement when P_aO_2 <8 kPa.
- Neuromuscular or skeletal disorders, after specialist assessment.
- Obstructive sleep apnoea despite continuous positive airways pressure therapy, after specialist assessment.
- Pulmonary malignancy or other terminal disease with disabling dyspnoea.
- Heart failure with daytime P_aO_2 < 7.3 kPa (on air) or with nocturnal hypoxaemia.

(owing to the steepness of the oxygen–haemoglobin dissociation curve) and desaturation of up to 50% occurs. The mechanism is alveolar hypoventilation due to:

- inhibition of intercostal and accessory muscles in REM sleep
- shallow breathing in REM sleep, which reduces ventilation, particularly in severe COPD
- an increase in upper airway resistance because of a reduction in muscle tone.

These nocturnal hypoxaemic episodes are associated with a further rise in pulmonary arterial pressure owing to vasoconstriction, and the majority of deaths in patients with COPD occur during the night, possibly from cardiac arrhythmias. These patients additionally show severe secondary polycythaemia, partly as a result of the severe nocturnal hypoxaemia.

Each episode of desaturation is usually terminated by arousal from sleep, so that normal sleep is reduced and the patient suffers from daytime sleepiness.

Treatment. Patients with arterial hypoxaemia should not be given sleeping tablets, as these will further depress respiratory drive. Treatment is with nocturnal administration of oxygen and ventilatory support.

Positive-pressure ventilation can be administered non-invasively through a tightly fitting nasal mask with bilevel positive airway pressure (BiPAP) – inspiratory to provide inspiratory assistance and expiratory to prevent alveolar closure, each adjusted independently. Although these devices to maintain adequate ventilation during sleep and to allow respiratory muscles to rest at night are effective in chronic chest wall disease (e.g. kyphoscoliosis) or neuromuscular disease (e.g. previous poliomyelitis), they do not improve respiratory function, respiratory muscle strength, exercise tolerance or breathlessness in patients with COPD.

Pulmonary rehabilitation

A modest increase in exercise capacity with diminution in the sense of breathlessness and improved general well-being can result from exercise training. Regular training periods can be instituted at home; climbing stairs or walking fixed distances can be combined with regular clinic visits for encouragement. Breathing exercises are probably of less value. Quality of life can be improved by a multidisciplinary approach involving physiotherapy, exercise and education, although this does not alter life expectancy or the rate of decline in lung function; stopping smoking is still the most beneficial. Nutritional advice, psychological, social and behavioural intervention is helpful.

Additional measures

- *Vaccines*. Patients with COPD should receive yearly influenza vaccine and one dose of the polyvalent pneumococcal polysaccharide vaccine (a single dose usually provides lifelong immunity).
- *α_1-Antitrypsin replacement*. Weekly or monthly infusions of α_1-antitrypsin have been recommended for patients with serum levels of this compound below 310 mg/L and abnormal lung function. Whether this modifies the long-term progression of the disease has still to be determined.
- *Heart failure* should be treated (p. 737).
- *Secondary polycythaemia* – venesection is recommended if the PCV is >55%.
- *Pulmonary hypertension* can be partially relieved by the use of oral β-adrenergic stimulants such as salbutamol (4 mg three times daily), but whether this is useful in the long term is unknown.
- The *sensation of breathlessness* can be reduced by the use of either promethazine 125 mg daily or dihydrocodeine 1 mg/kg by mouth. Reduced breathlessness and increased exercise tolerance also result from the combination of oral dihydrocodeine and oxygen from a portable cylinder. Although opiates are the most effective treatment for intractable breathlessness they depress ventilation and carry the risk of increasing respiratory failure.
- *Antileukotriene agents* are being used in trials (p. 856).
- *Air travel*. Commercial cabin pressures are equivalent to an altitude of 2000–2400 m. In healthy people, the P_aO_2 will fall from 13.5 to 10 kPa with a fall of 3% in oxygen saturation. In patients with moderate COPD, significant desaturation can occur. The desaturation associated with air travel can be simulated by measuring oxygen saturation while breathing 15% oxygen at sea level. Patients whose saturation drops below 85% within 15 minutes should be advised to contact their airline to request supplemental oxygen during flight.
- *Surgery*. Some patients with large emphysematous bullae (which reduce lung capacity) can benefit from bullectomy, enabling adjacent areas of collapsed lung to re-expand and start functioning again. In addition, carefully selected patients with severe COPD (FEV_1 < 1 L) may benefit from lung volume reduction surgery. This surgery increases elastic recoil, which reduces the expiratory collapse of the airway and decreases expiratory airflow limitation. It also enables the diaphragm to work at a better mechanical advantage. Initial studies suggested that ventilation was improved and patients felt less breathless, but mortality was

unchanged. However, a controlled trial in *severe* emphysema showed increased mortality and no improvement in the patients' condition. *Single lung transplantation* (see p. 846) is used for end-stage emphysema, with 3-year survival rates of 75%. The principal benefit is to improve quality of life but it does not extend survival.

Acute exacerbation of COPD (type II respiratory failure)

- *Oxygen therapy*. There are many causes of respiratory failure but COPD is by far the most common (Fig. 14.26). The primary aim of the management of respiratory failure is to improve the P_aO_2 by continuous oxygen therapy. In type II respiratory failure the P_aCO_2 is elevated and the patient is dependent on hypoxic drive. Consequently, giving additional oxygen will nearly always lead to a further rise in the P_aCO_2. Small increases in P_aCO_2 can be tolerated but not if the pH falls dramatically. The pH should not be allowed to fall below 7.25; if it does, increased ventilation must be achieved either by artificial ventilation or by using a respiratory stimulant. Figure 14.27 shows a fixed-performance mask (Venturi mask) for the administration of oxygen. This style of mask is used to deliver low concentrations of oxygen. It should be compared with the variable-performance face mask (see Fig. 15.15). Initially, 24% oxygen is given, which is only slightly greater than the concentration of oxygen in air. However, because of the shape of the oxygen–haemoglobin dissociation curve (see Fig. 15.5), this small increase in oxygen is valuable. The concentration of inspired oxygen can be gradually increased provided the P_aCO_2 does not rise unacceptably.
- *Removal of retained secretions*. The patient should be encouraged to cough to remove secretions. Physiotherapy is helpful. If this fails, bronchoscopy and/or aspiration via an endotracheal tube may be necessary. A tracheostomy is only rarely required.
- *Respiratory support* (see p. 916). Non-invasive ventilatory techniques can be very helpful in avoiding the need for endotracheal intubation. The best current technique uses tight-fitting facial masks to deliver bilevel positive airway pressure ventilatory support (BiPAP). Assisted ventilation with an endotracheal tube (continuous positive airway pressure (CPAP)) is occasionally used for patients with COPD with severe respiratory failure when there is a definite precipitating factor and the overall prognosis is reasonable. Assessing the relative reversibility in an acute setting can present a difficult ethical problem.
- *Respiratory stimulants*. Respiratory stimulants are less used nowadays, largely because of the increasing availability of respiratory support services. Doxapram, 1.5–4.0 mg/min by slow i.v. infusion, may help in the short term to arouse the patient and to stimulate coughing, with clearance of some secretions.
- *Corticosteroids, antibiotics and bronchodilators* should be administered in the acute phase and then reassessed once the patient has recovered (see above).

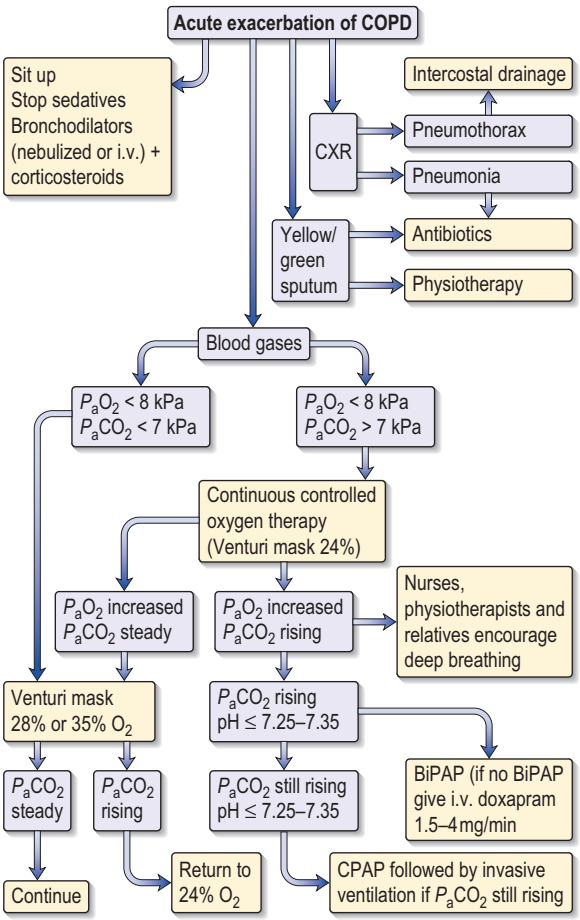


Fig. 14.26 **Algorithm for the treatment of respiratory failure in COPD.** LVF, left ventricular failure; CPAP, continuous positive airway pressure; BiPAP, bilevel positive airway pressure.

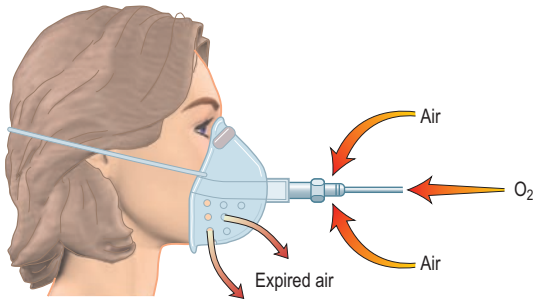


Fig. 14.27 **'Fixed-performance' device for administration of oxygen to spontaneously breathing patients (Venturi mask).** Oxygen is delivered through the injector of the Venturi mask at a given flow rate. A fixed amount of air is entrapped and the inspired oxygen can be predicted accurately. Masks are available to deliver 24%, 28% and 35% oxygen.

Prognosis of COPD

Predictors of a poor prognosis are increasing age and worsening airflow limitation, i.e. a fall in FEV_1. A predictive index (BODE – **b**ody mass index, degree of airflow **o**bstruction, **d**yspnoea and **e**xercise capacity) is shown in Box 14.2.

A patient with a BODE index of 0–2 has a mortality rate of 10%; one with a BODE index of 7–10, a mortality rate of 80% at 4 years.

Ask the authors
www.studentconsult.com

Box 14.2 BODE index (body mass index, degree of airflow obstruction, dyspnoea and exercise capacity)

Variable	Points on BODE index 0	1	2	3
FEV_1 (percent of predicted)	≥65	50–64	36–49	≥35
Distance walked in 6 minutes (metres)	≥350	250–349	150–249	≤149
MMRC dyspnoea scale	0–1	2	3	4
Body mass index	>21	≤21		

Scores on the modified Medical Research Council (MMRC) dyspnoea scale range from 0 to 4, with a score of 4 indicating that the patient is breathless when dressing.

Obstructive sleep apnoea

This condition affects 1–2% of the population and occurs most often in overweight middle-aged men. It can occur in children, particularly those with enlarged tonsils. The major symptoms and their frequency are listed in Table 14.10. During sleep, activity of the respiratory muscles is reduced, especially during REM sleep when the diaphragm is virtually the only active muscle. Apnoeas occur when the airway at the back of the throat is sucked closed when breathing in during sleep. When awake, this tendency is overcome by the action of opening muscles of the upper airway, the genioglossus and palatal muscles, but these become hypotonic during sleep (Fig. 14.28). Partial narrowing results in snoring, complete occlusion causes apnoea and critical narrowing causes hypopnoeas. Apnoea leads to hypoxia and increasingly strenuous respiratory efforts until the patient overcomes the resistance. The combination of the central hypoxic stimulation and the effort to overcome obstruction wakes the patient from sleep. These awakenings are so brief that the patient remains unaware of them but may be woken hundreds of times per night leading to sleep deprivation, especially a reduction in REM sleep, with consequent daytime sleepiness and impaired intellectual performance. Contributory factors are obesity, narrow pharyngeal opening and coexistent COPD.

Correctable factors occur in about one-third of cases and include:

- encroachment on pharynx – obesity, acromegaly, enlarged tonsils
- nasal obstruction – nasal deformities, rhinitis, polyps, adenoids
- respiratory depressant drugs – alcohol, sedatives, strong analgesics.

Table 14.10 Symptoms of obstructive sleep apnoea

Loud snoring (95%)	Nocturnal choking (30%)
Daytime sleepiness (90%)	Reduced libido (20%)
Unrefreshed sleep (40%)	Morning drunkenness (5%)
Restless sleep (40%)	Ankle swelling (5%)
Morning headache (30%)	

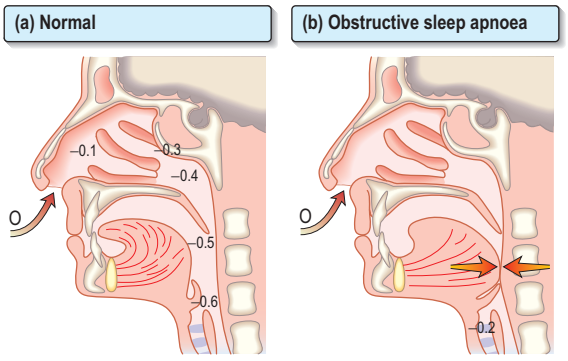


Fig. 14.28 **Section through head, showing pressure changes (in kPa) in (a) the normal situation and (b) obstructive sleep apnoea.** There is a pressure drop during inspiration as air is sucked through the turbinates. In patients with obstructive sleep apnoea this is sufficient to collapse the pharynx, obstructing inspiration.

Diagnosis

Relatives often provide a good history of the snore–silence–snore cycle. The Epworth Sleepiness Scale (see Table 20.5) helps discriminate apnoea from simple snoring. The diagnosis is supported by non-invasive ear or finger oximetry performed at home. Characteristically, oxygen saturation falls in a cyclical manner giving a sawtooth appearance to the tracing. If oximetry is negative or equivocal, inpatient assessment with oximetry and video-recording is indicated, preferably in a room specifically adapted for sleep studies. Full polysomnographic studies are rarely necessary for clinical diagnosis but are useful in research labs. These involve oximetry, direct measurements of thoracic and abdominal movement to assess breathing, and electroencephalography to record patterns of sleep and arousal. Some centres also measure oronasal airflow.

The diagnosis of sleep apnoea/hypopnoea is confirmed if there are more than 10–15 apnoeas or hypopnoeas in any 1 hour of sleep. There is, however, overlap with *central sleep apnoea* (see p. 1145).

Management

Management consists of correction of treatable factors (see above) with, if necessary, nasal continuous positive airway pressure (CPAP) delivered by a nasal mask during sleep. Such systems raise the pressure in the pharynx by about 1 kPa, keeping the walls apart. CPAP results in improvement of symptoms, quality of life, daytime alertness and survival. Up to 50%, however, cannot tolerate CPAP. Modafinil (a CNS stimulant) is useful in the short term.

Bronchiectasis

The term ‘bronchiectasis’ is used to describe abnormal and permanently dilated airways. Bronchial walls become inflamed, thickened and irreversibly damaged. The mucociliary transport mechanism is impaired and frequent bacterial infections ensue. Clinically, the disease is characterized by cough production of large amounts of sputum and dilated and thickened bronchi, detected on CT.

Aetiology

The causes are shown in Table 14.11. Cystic fibrosis is the most common cause in developed countries.

Table 14.11 Causes of bronchiectasis

Congenital
Deficiency of bronchial wall elements
Pulmonary sequestration

Mechanical bronchial obstruction
Intrinsic:
- Foreign body
- Inspissated mucus
- Post-tuberculous stenosis
- Tumour

Extrinsic:
- Lymph node
- Tumour

Postinfective bronchial damage
Bacterial and viral pneumonia, including pertussis, measles and aspiration pneumonia

Granuloma
Tuberculosis, sarcoidosis

Diffuse diseases of the lung parenchyma
e.g. idiopathic pulmonary fibrosis

Immunological over-response
Allergic bronchopulmonary aspergillosis
Post-lung transplant

Immune deficiency
Primary:
- Panhypogammaglobulinaemia
- Selective immunoglobulin deficiencies (IgA and IgG_2)

Secondary:
- HIV and malignancy

Mucociliary clearance defects
Genetic:
- Primary ciliary dyskinesia (Kartagener's syndrome with dextrocardia and situs inversus)
- Cystic fibrosis

Acquired:
- Young's syndrome – azoospermia, sinusitis

Clinical features

Patients with mild bronchiectasis only produce yellow or green sputum after an infection. Localized areas of the lung may be particularly affected, when sputum production will depend on position. As the condition worsens, the patient suffers from persistent halitosis, recurrent febrile episodes with malaise, and episodes of pneumonia. Clubbing occurs, and coarse crackles can be heard over the infected areas, usually the bases of the lungs. When the condition is severe there is continuous production of foul-smelling, thick, khaki-coloured sputum. Haemoptysis can occur either as blood-stained sputum or as a massive haemorrhage. Breathlessness may result from airflow limitation.

Investigations

- **Chest X-ray** may be normal or may show dilated bronchi with thickened bronchial walls and sometimes multiple cysts containing fluid.
- **High-resolution CT scanning** (see p. 823) shows bronchial dilatation with loss of airway tapering at the periphery, bronchial wall thickening and cysts at the end of the bronchioles (Fig. 14.29) with a sensitivity of 97%.
- **Sputum** examination with culture and sensitivity of the organisms is essential for adequate treatment. The major pathogens are *Staph. aureus, Pseudomonas aeruginosa, H. influenzae* and anaerobes. Other pathogens include *Strep. pneumoniae* and *Klebsiella pneumoniae. Aspergillus fumigatus* can be isolated from 10% of sputum specimens in cystic fibrosis, but the role of this organism in producing infection is uncertain. Mycobacterium avium-intracellulare complex (MAI) is being increasingly found.
- **Sinus X-rays.** Thirty per cent have concomitant rhinosinusitis.
- **Serum immunoglobulins.** Ten per cent of adults with bronchiectasis have antibody deficiency (mainly IgA).
- **Sweat electrolytes** – if cystic fibrosis is suspected (see p. 845).
- **Mucociliary clearance** (nasal clearance of saccharin). A 1 mm cube of saccharin is placed on the inferior turbinate and the time to taste measured (normally less than 30 minutes).

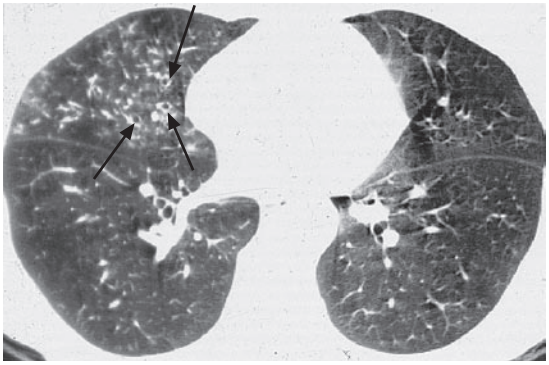


Fig. 14.29 **CT scan showing bronchiectasis in the right middle lobe.** Note dilated bronchi with thickened wall and adjacent artery giving a signet ring appearance.

Treatment

Postural drainage

Postural drainage is essential and patients must be trained by physiotherapists to tip themselves into a position in which the affected lobe(s) are uppermost at least three times daily for 10–20 minutes. Most patients find that lying over the side of the bed with head and thorax down is effective.

Antibiotics

Experience from the treatment of cystic fibrosis suggests that bronchopulmonary infections need to be eradicated if progression of the disease is to be halted. In mild cases, intermittent chemotherapy with cefaclor 500 mg three times daily or ciprofloxacin 500 mg twice daily may be sufficient. Flucloxacillin 500 mg 6-hourly is needed if *Staph. aureus* is isolated.

If the sputum remains yellow or green despite regular physiotherapy and intermittent chemotherapy, or if lung function deteriorates despite treatment with bronchodilators, it is likely that there is infection with *Pseudomonas aeruginosa*. Treatment requires parenteral or aerosol chemotherapy at regular 3-month intervals. Ceftazidime 2 g intravenously 8-hourly or by inhalation (1 g twice daily) has been shown to be effective. Ciprofloxacin 750 mg twice daily orally may work in the short term, but resistance can develop rapidly. High sputum levels of some antibiotics, e.g. tobramycin, can be achieved by inhalation.

Treatment for aspergillus is described on page 876 and MAI on page 204.

Bronchodilators
Bronchodilators are useful in patients with demonstrable airflow limitation.

Anti-inflammatory agents
Inhaled or oral steroids can decrease the rate of progression.

Surgery
Unfortunately, it is rare for bronchiectasis to be sufficiently localized for resection to be practical. Lung or heart–lung transplantation is sometimes required.

Complications
The incidence of complications has fallen with antibiotic therapy. Pneumonia, pneumothorax, empyema and metastatic cerebral abscess can occur. Severe, life-threatening haemoptysis can also occur, particularly in patients with cystic fibrosis.

Massive haemoptysis originates from the high-pressure systemic bronchial arteries and has a mortality of 25%. Other causes of massive haemoptysis are pulmonary tuberculosis (most common), aspergilloma, lung abscess and primary and secondary malignant tumours.

Treatment of the haemoptysis consists of bed rest and antibiotics, when most stop bleeding. Blood transfusion is given if required. Urgent fibreoptic bronchoscopy is occasionally necessary to detect the source of bleeding. If the haemoptysis does not settle rapidly, the treatment of choice is bronchial artery embolization. Surgical resection may be required if embolization fails.

Prognosis
The advent of effective antibiotic therapy has greatly improved the prognosis. Ultimately, most patients with severe bronchiectasis will develop respiratory failure. Cor pulmonale is another well-recognized complication. The three pathogens that can cause infective episodes and are difficult to eradicate are *Pseudomonas aeruginosa, Aspergillus fumigatus* and MAI.

Cystic fibrosis
In cystic fibrosis (CF) there is an alteration in the viscosity and tenacity of mucus produced at epithelial surfaces. The classical form of the syndrome includes bronchopulmonary infection and pancreatic insufficiency, with a high sweat sodium and chloride concentration. It is an autosomal recessive inherited disorder with a carrier frequency in Caucasians of 1 in 22 (see p. 44). There is a gene mutation on the long arm of chromosome 7 in the region of 7q31.2. The commonest abnormality is a specific deletion at position 508 in the amino acid sequence [ΔF_{508}] – which results in a defect in a transmembrane regulator protein (see p. 46). This is the cystic fibrosis transmembrane conductance regulator (CFTR), which is a critical chloride channel (Fig. 14.30). The mutation alters the secondary and tertiary structure of the protein, leading to a failure of opening of the chloride channel in response to elevated cyclic AMP in epithelial cells. This results in decreased excretion of chloride into the airway lumen and an increased reabsorption of sodium into the epithelial cells. With less excretion of salt there is less excretion of water and increased viscosity and tenacity of airway secretions. A possible reason for the high salt content of sweat is that there is a CFTR-independent mechanism of chloride secretion in the sweat gland with an impaired reabsorption of sodium chloride in the distal end of the duct. Many genetic variants are known. The frequency of ΔF_{508} mutation in CF is 70% in the USA and UK, under 50% in southern Europe and 30% in Ashkenazi families.

Clinical features
Respiratory effects
Although the lungs of babies born with CF are structurally normal at birth, frequent respiratory infections soon develop and are often the presenting feature. CF is the commonest cause of recurrent bronchopulmonary infection in childhood, and is a major cause in early adult life. Sinusitis is almost universal and nasal polyps are common. Breathlessness and haemoptysis occur in the later stages as airflow limitation and bronchiectasis develop. Spontaneous pneumothorax may occur. Respiratory failure and cor pulmonale eventually develop.

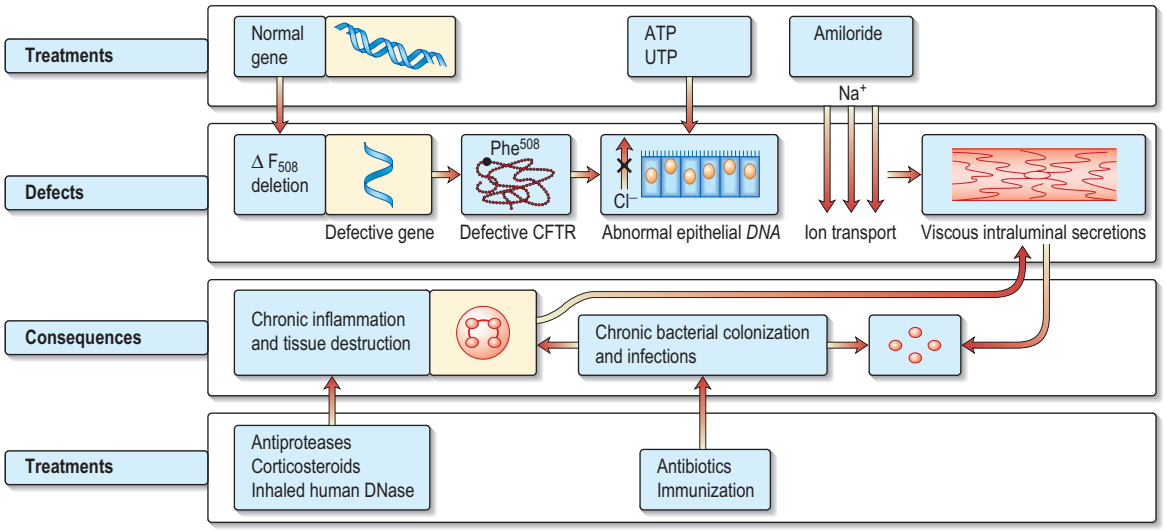


Fig. 14.30 **Cystic fibrosis.** Abnormalities and possible therapies (see text). CFTR, cystic fibrosis transmembrane conductance regulator.

Gastrointestinal effects

About 85% of patients have symptomatic steatorrhoea owing to pancreatic dysfunction (see p. 381). Children may be born with meconium ileus owing to the viscid consistency of meconium in CF, and later in life develop the meconium ileus equivalent syndrome, a form of small intestinal obstruction unique to CF. Cholesterol gallstones appear to occur with increased frequency. Cirrhosis develops in about 5% of older patients and there are increased incidences of peptic ulceration and gastrointestinal malignancy.

Nutritional effects

Many patients suffer from malnutrition due to a combination of malabsorption and maldigestion. Poor nutrition is associated with increased pulmonary sepsis.

Other features

Puberty and skeletal maturity are delayed in most CF patients. Males are almost always infertile owing to failure of development of the vas deferens and epididymis. Females are able to conceive, but often develop secondary amenorrhoea as the disease progresses. Arthropathy and diabetes mellitus (in 11% of adults with CF) also occur.

Diagnosis

The diagnosis of CF in older children and adults is based on the clinical history and:

- a family history of the disease
- a high sweat sodium concentration over 60 mmol/L (meticulous technique by laboratories performing regular sweat analysis is essential, but the test is still difficult to interpret in adults)
- blood DNA analysis of gene defect
- radiology showing features seen in bronchiectasis (see p. 843)
- absent vas deferens and epididymis
- blood immunoreactive trypsin levels – these are not useful diagnostically but may be useful in screening.

Treatment

CF patients should be managed by multidisciplinary specialized teams. The overall care involves education to improve their quality of life, good nutrition and the prompt treatment of exacerbations to avoid hospitalization.

- *General care* should include stopping smoking, vaccination with influenza and pneumococcal vaccines and pulmonary rehabilitation (p. 840).
- *Oxygen therapy* should be given as necessary (Box 14.1).
- *Antibiotic treatment* for respiratory infections is as described under bronchiectasis on page 843.
- Seventy per cent of adults with CF have *Pseudomonas* infection in their sputum. Nebulized anti-pseudomonal antibiotic therapy improves lung function and decreases the risk of infective exacerbations and hospitalization in these patients.
- *Drug therapy* includes β_2 agonists (p. 854) and inhaled corticosteroids (p. 855). These produce symptomatic relief but have no effect on long-term survival.
- *Airway clearance*. Inhalation of recombinant DNAse (dornase alfa 2.5 mg daily) has been shown to improve FEV_1 by 20% in some patients. Hypertonic saline by inhalation (in concentrations of up to 7%) gives short-term benefit. Amiloride, which inhibits sodium transport, has been used but no overall benefit has been shown in meta-analysis. *N*-acetylcysteine has been shown in vitro to liquefy CF sputum by cleaving disulfide bonds in mucus glycoproteins but clinical studies have been disappointing.
- *Non-invasive ventilation* (p. 916) improves symptoms in chronic respiratory failure but there is no evidence of a survival benefit. It acts as a bridge to *lung transplantation* (p. 846).
- Treatment for pancreatic insufficiency (p. 382) and malnutrition (p. 214).
- Human experimental studies have been conducted on the delivery to the epithelium of the normal CFTR gene using, as a vector, a replication-deficient adenovirus containing normal human CFTR complementary DNA which is trophic for epithelial cells. Gentamicin can suppress premature termination codons, and nasal administration has been shown to correct the physiological abnormality. These studies are in their early stages.

Prognosis and screening

Almost all CF patients develop progressive respiratory failure but the prognosis has improved considerably over the years. Ninety per cent of children now survive into their teens and the median survival for those born after 1990 is estimated at 40 years. A major problem is sputum infection with *Burkholderia cepacia* (formerly classified as a *Pseudomonas*), a plant pathogen which was previously considered a harmless commensal. Its acquisition can be associated with accelerated disease and rapid death. Multiple antibiotic resistance is common and spread occurs from person to person. Strategies to limit transmission include rigid segregation of both inpatients and outpatients and advice to CF sufferers not to socialize together. Sadly, groups formed for mutual support and education have been dissolved, leading to considerable distress.

Genetic screening is available for the four most common mutations and this identifies 85–95% of carriers. Screening for the carrier state should be offered to persons or couples with a family history of CF, together with counselling (see p. 44).

Chronic cough (see p. 819)

Pathological coughing results from two mechanisms:

- stimulation of sensory nerves in the epithelium by secretions, foreign bodies, cigarette smoke and tumours
- sensitization of the cough reflex in which there is an abnormal increase in the sensitivity of the cough receptors demonstrable by inhalation of capsaicin or saline solutions.

Sensitization of the cough reflex presents clinically as a persistent tickling sensation in the throat with paroxysms of coughing induced by changes in air temperature, aerosol sprays, perfumes and cigarette smoke. It is found in association with viral infections, oesophageal reflux, postnasal drip, cough-variant asthma, idiopathic cough, and in 15% of patients taking angiotensin-converting enzyme (ACE) inhibitors. The association with ACE inhibitors implicates

neuropeptides, prostaglandins E_2 and $F_2\alpha$ and bradykinin as a cause of the cough. In the absence of chest X-ray abnormalities, possible investigations include:

- ENT examination (p. 1080) and sinus CT for postnasal drip
- lung function tests and histamine bronchial provocation testing (p. 852) for cough-variant asthma
- ambulatory oesophageal pH monitoring and impedance for oesophageal reflux
- CT scan of thorax for diseases of lung parenchyma
- $\dot{V}_A/\dot{Q}$ scans for recurrent pulmonary embolism
- fibreoptic bronchoscopy for inhaled foreign body or tumour
- ECG, echocardiography and exercise testing and impedance for cardiac causes
- hyperventilation testing and psychiatric appraisal.

The absence of any pathology makes the management of unexplained cough difficult. Morphine will depress the sensitized cough reflex but its unwanted effects limit its use in the long term. Dihydrocodeine linctus may be of value in some patients. Demulcent preparations and cough sweets provide temporary relief only. Patients taking ACE inhibitors should be changed to an angiotensin-II receptor antagonist, e.g. losartan (see p. 738), which does not block bradykinin.

Lung and heart–lung transplantation

Indications and donor selection

The main diseases treated by transplantation are:

- pulmonary fibrosis
- primary pulmonary hypertension
- cystic fibrosis
- bronchiectasis
- emphysema – particularly α_1-antitrypsin inhibitor deficiency
- Eisenmenger's syndrome.

Indications for this treatment are patients under 60 years with a life expectancy of less than 18 months, no underlying cancer and no serious systemic disease.

Donor selection includes age under 40 years, good cardiac and lung function, and chest measurements slightly smaller than those of the recipient. Matching for ABO blood group is essential, but rhesus blood group compatibility is not essential. Since donor material is limited, single lung transplantation is preferred to double lung or heart–lung transplantation and can be successfully undertaken in pulmonary fibrosis, pulmonary hypertension and emphysema. Bilateral lung transplantation is required in infective conditions to prevent spillover of bacteria from the diseased lung to a single lung transplant. Eisenmenger's syndrome requires heart–lung transplant.

Complications and their treatment

- *Early* – post-transplant pulmonary oedema requires diuretics and respiratory support by ventilation.
- *Infections*, particularly within first 3 months:
  Bacterial pneumonia – antibiotics
  Cytomegalovirus – ganciclovir
  Herpes simplex – aciclovir.
- *P. jiroveci* – prophylactic co-trimoxazole.
- *Immunosuppression* is with ciclosporin (inhaled formulation has shown benefit) or tacrolimus, azathioprine or mycophenolate mofetil and prednisolone.
- *Rejection*:
  Early (first few weeks) – high-dose i.v. corticosteroids
  Late (after 3 months) – in obliterative bronchiolitis, high-dose i.v. corticosteroids are sometimes effective. Post-transplant lymphoproliferative disease may respond to rituximab, a monoclonal antibody which causes lysis of B lymphocytes.

Prognosis. Several studies show a major improvement in overall quality of life. One year survival rates have improved with a yearly mortality rate of about 10%. Death is mainly due to bronchiolitis. Actual survival varies with the original diagnosis but the median survival is approximately four years.

ASTHMA

Asthma is a common chronic inflammatory condition of the lung airways whose cause is incompletely understood. Symptoms are cough, wheeze, chest tightness and shortness of breath, often worse at night. The most frequent form has its onset in childhood between the ages of 3 and 5 years and may either worsen or improve during adolescence. Classically asthma has three characteristics:

- *airflow limitation* which is usually reversible spontaneously or with treatment
- *airway hyperresponsiveness* to a wide range of stimuli (see below)
- *inflammation of the bronchi* with T lymphocytes, mast cells, eosinophils with associated plasma exudation, oedema, smooth muscle hypertrophy, matrix deposition, mucus plugging and epithelial damage.

In chronic asthma, inflammation may be accompanied by irreversible airflow limitation as a result of airway wall remodelling that may involve large and small airways and mucus impaction.

Prevalence

In many countries the prevalence of asthma is increasing. This increase, with its accompanying allergy, is particularly in children and young adults where this disease may affect up to 15% of the population. There is also a geographical variation, with asthma being commoner in more developed countries. Some of the highest rates are in the UK, New Zealand and Australia, but the rates are lower in Far Eastern countries such as China and Malaysia, Africa and Central and Eastern Europe. However, long-term follow-up in developing countries suggests that the disease may become more frequent as individuals adopt a more 'westernized' lifestyle, but the environmental factors accounting for this remain unknown. Studies of occupational asthma suggest that a high percentage of the workforce, 15–20%, may become asthmatic if exposed to potent sensitizers. World-wide, approximately 300 million people have asthma and this is expected to rise to 400 million by 2025.

Classification

Asthma is a complex disorder of the conducting airways that most simply can be classified as:

FURTHER READING

Barnes PJ. Small airways in COPD. *New England Journal of Medicine* 2004; **350**: 2635–2637.

Garpestad E, Brennan J, Hill NS. Non-invasive ventilation for critical care. *Chest* 2007; **132**: 711–720.

Mannino DM, Buist AS. Global burden of COPD: risk factors, prevalence and future trends. *Lancet* 2007; **370**: 765–774.

Qaseem A, Snow V, Shekelle P et al. Diagnosis and management of stable chronic obstructive pulmonary disease: A practice guideline from the American College of Physicians. *Annals of Internal Medicine* 2007; **147**: 633–638.

Rowe SM, Miller S, Sorscher EJ. Cystic fibrosis. *New England Journal of Medicine* 2005; **252**: 1992–2001.

- *extrinsic* – implying a definite external cause
- *intrinsic* – when no causative agent can be identified.

Extrinsic asthma occurs most frequently in atopic individuals who show positive skin-prick reactions to common inhalant allergens such as dust mite, animal danders, pollens and fungi. Positive skin-prick tests to inhalant allergens are shown in 90% of children and 70% of adults with persistent asthma. Childhood asthma is often accompanied by eczema (atopic dermatitis) (see p. 1236). A frequently overlooked cause of late-onset asthma in adults is sensitization to chemicals or biological products in the workplace.

Intrinsic asthma often starts in middle age ('late onset'). Nevertheless, many patients with adult-onset asthma show positive allergen skin tests and on close questioning give a history of respiratory symptoms compatible with childhood asthma.

Non-atopic individuals may develop asthma in middle age from extrinsic causes such as sensitization to occupational agents such as toluene diisocyanate, intolerance to non-steroidal anti-inflammatory drugs such as aspirin or because they were given β-adrenoceptor-blocking agents for concurrent hypertension or angina that block the protective effect of endogenous adrenergic agonists. Extrinsic causes must be considered in all cases of asthma and, where possible, avoided.

Aetiology and pathogenesis

The two major factors involved in the development of asthma and many other stimuli that can precipitate attacks are shown in Figure 14.31.

Atopy and allergy

The term 'atopy' was used by clinicians at the beginning of the twentieth century to describe a group of disorders, including asthma and hayfever, that appeared:

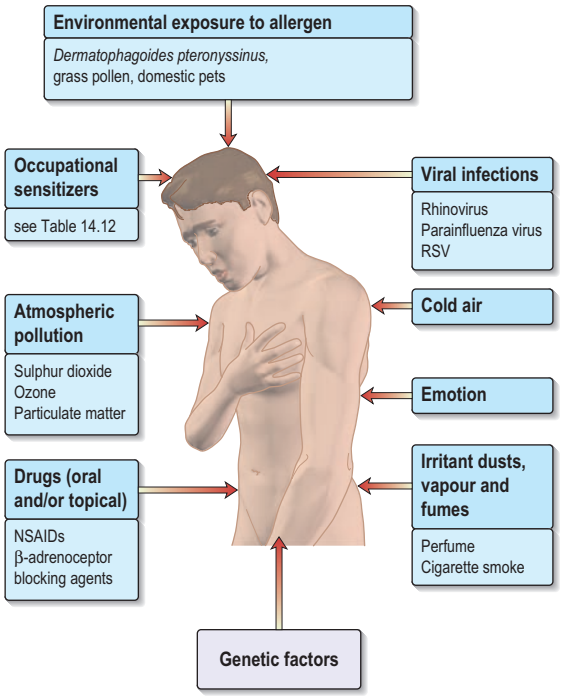


Fig. 14.31 **Causes and triggers of asthma.** RSV, respiratory syncytial virus; NSAIDs, non-steroidal anti-inflammatory drugs.

- to run in families
- to have characteristic wealing skin reactions to common allergens in the environment
- to have circulating allergen-specific IgE.

Allergen-specific IgE is present in 30–40% of the UK population, and there is a link between serum IgE levels and both the prevalence of asthma and airway hyperresponsiveness. Genetic and environmental factors affect serum IgE levels.

Genetic

Genes, in combination with environmental factors, may turn out to play a key role in the development of asthma.

- Genes controlling the production of the cytokines IL-3, IL-4, IL-5, IL-9, IL-13 and GM-CSF – which in turn affect mast and eosinophil cell development and longevity as well as IgE production – are present in a cluster on chromosome 5q31–33 (the IL-4 gene cluster).
- Polymorphic variation in proteins along the IL-4/-13 signalling pathway is strongly associated with allergy and asthma.
- Novel asthma genes identified by positional cloning from whole genome scans are the PHF11 locus on chromosome 2 (that includes genes SETDB2 and RCBTB1) and transcription factors, which are implicated in IgE synthesis and associated more with atopy than asthma.
- ADAM 33 (a disintegrin and metalloproteinase) on chromosome 20p13 is more strongly associated with airway hyperresponsiveness and tissue remodelling.
- Other recently discovered genes associated with asthma are those that encode neuropeptide S receptor (GPRA or GPR154) on chromosome 7p15, HLA-G on chromosome 6p21, dipeptidyl peptidase 10 on chromosome 2q14 and most recently on chromosome 17q21 ORMDL3, a member of a gene family that encodes transmembrane proteins anchored in the endoplasmic reticulum.

Environmental factors

Early childhood exposure to allergens and maternal smoking has a major influence on IgE production. Much current interest focuses on the role of intestinal bacteria and childhood infections in shaping the immune system in early life. It has been suggested that growing up in a relatively 'clean' environment may predispose towards an IgE response to allergens (the 'hygiene hypothesis'). Conversely, growing up in a 'dirtier' environment may allow the immune system to avoid developing allergic responses. Components of bacteria (e.g. lipopolysaccharide endotoxin; immunostimulatory CpG DNA sequences; flagellin), viruses (e.g. SS- and DS-RNA) and fungi (e.g. chiton, a cell wall component) are able to stimulate up to 10 different toll-like receptors (TLRs) expressed on immune and epithelial cells to direct the immune and inflammatory response away from the allergic (Th2) towards protective (Th1 and Treg) pathways. Th1 immunity is associated with antimicrobial protective immunity whereas regulatory T cells are strongly implicated in tolerance to allergens. Thus early life exposure to inhaled and ingested products of microorganisms, as occurs in livestock farming communities and developing countries, may be critical in helping shape the subsequent risk of a child becoming allergic and/or developing asthma.

The allergens involved in asthma are similar to those in rhinitis although pollen exposure causes hay fever to a greater extent than asthma. Allergens from the faecal particles of the house-dust mite are associated with most cases of asthma world-wide. Cockroach allergy has been implicated in asthma in US inner-city children, while allergens from furry pets (especially cats) are increasingly common causes. The fungal spores from *Aspergillus fumigatus* give rise to a complex series of lung disorders, including asthma (see p. 875). Many allergens, including those from *Aspergillus*, have intrinsic biological properties, e.g. proteolytic enzymes that facilitate their passage through the airway epithelium to increase their sensitizing capacity.

Chitins are cross-linked polysaccharides found in the exoskeleton of insects and cockroaches, fungi and in the eggs of helminths. They can be inhaled into the airways. Chitinase-family proteins may play a role in the pathogenesis of asthma as the levels in the lungs and the serum are high in asthma and correlate with disease activity.

Increased responsiveness of the airways of the lung (airway hyperresponsiveness)

Bronchial hyperresponsiveness (BHR) is demonstrated by asking the patient to inhale gradually increasing concentrations of either histamine or methacholine (*bronchial provocation tests*). This induces transient airflow limitation in susceptible individuals (approximately 20% of the population); the dose of the agonist (provocation dose, PD) necessary to produce a 20% fall in FEV_1 is known as the PD_{20} FEV_1 (or provocation concentration PC_{20} FEV_1). Patients with clinical symptoms of asthma respond to very low doses of methacholine, i.e. they have a low PD_{20} FEV_1 (<11 µmol). Exercise testing or inhalation of cold dry air, mannitol or hypertonic saline are other methods to assess BHR, but all of these involve the stimulus first releasing endogenous mediators such as histamine, prostaglandins and leukotrienes into the airways to cause the bronchoconstriction (indirect BHR). Measures of indirect BHR correlate more closely with symptoms and diurnal peak expiratory flow rate (PEFR) variation than PC_{20} histamine or methacholine and both are useful in diagnosing asthma if there is doubt and in guiding controller treatment.

Some patients also react to methacholine but at *higher doses* and include those with:

- attacks of asthma only on extreme exertion, e.g. winter sports enthusiasts
- wheezing or prolonged periods of coughing following a viral infection
- seasonal wheeze during the pollen season
- allergic rhinitis, but not complaining of any lower respiratory symptoms until specifically questioned
- some subjects with no respiratory symptoms.

Although the degree of hyperresponsiveness can itself be influenced by allergic mechanisms (see p. 851 and Fig. 14.34), its pathogenesis and mode of inheritance involve a combination of airway inflammation and tissue remodelling.

Precipitating factors

Occupational sensitizers (Table 14.12)

Over 250 materials encountered at the workplace, accounting for 15% of all asthma cases, give rise to occupational asthma. The causes are recognized occupational diseases in the UK, and patients in insurable employment are therefore eligible for statutory compensation provided they apply within 10 years of leaving the occupation in which the asthma developed.

Table 14.12 Occupational asthma

Cause	Source/Occupation
Low molecular weight (non-IgE related)	
Isocyanates	Polyurethane varnishes Industrial coatings Spray painting
Colophony fumes	Soldering/welders Electronics industry
Wood dust	
Drugs	
Bleaches and dyes	
Complex metal salts, e.g. nickel, platinum, chromium	
High molecular weight (IgE related)	
Allergens from animals and insects	Farmers, workers in poultry and seafood processing industry; laboratory workers
Antidotes	Nurses, health industry
Latex	Health workers
Proteolytic enzymes	Manufacture (but not use) of 'biological' washing powders
Complex salts of platinum	Metal refining
Acid anhydrides and polyamine hardening agents	Industrial coatings

Asthma can be due to:

- high molecular weight compounds, e.g. flour, organic dusts and other large protein molecules involving specific IgE antibodies, or
- low molecular weight compounds, e.g. reactive chemicals such as isocyanates and acid anhydrides that bond chemically to epithelial cells to activate them as well as provide haptens recognized by T cells.

The risk of developing some forms of occupational asthma increases in smokers. The proportion of employees developing occupational asthma depends primarily upon the level of exposure. Proper enclosure of industrial processes or appropriate ventilation greatly reduces the risk. Atopic individuals develop occupational asthma more rapidly when exposed to agents causing the development of specific IgE antibody. Non-atopic individuals can also develop asthma when exposed to such agents, but after a longer period of exposure.

Non-specific factors

The characteristic feature of BHR in asthma means that, as well as reacting to specific antigens, the airways will also respond to a wide variety of non-specific direct and indirect stimuli.

Cold air and exercise

Most asthmatics wheeze after prolonged exercise. Typically, the attack does not occur while exercising but afterwards. The inhalation of cold, dry air will also precipitate an attack. Exercise-induced wheeze is driven by release of histamine,

prostaglandins (PGs) and leukotrienes (LTs) from mast cells as well as stimulation of neural reflexes when the epithelial lining fluid of the bronchi becomes hyperosmolar owing to drying and cooling during exercise. The phenomenon can be shown by exercise, cold air and hypertonic (e.g. saline or mannitol) provocation tests.

Atmospheric pollution and irritant dusts, vapours and fumes

Many patients with asthma experience worsening of symptoms on contact with tobacco smoke, car exhaust fumes, solvents, strong perfumes or high concentrations of dust in the atmosphere. Major epidemics have been recorded when large amounts of allergens are released into the air, e.g. soybean epidemic in Barcelona. Asthma exacerbations increase in both summer and winter air pollution episodes associated with climatic temperature inversions. Epidemics of the disease have occurred in the presence of high concentrations of ozone, particulates and NO_2 in the summer and particulates, NO_2 and SO_2 in the winter.

Diet

Increased intakes of fresh fruit and vegetables have been shown to be protective, possibly owing to the increased intake of antioxidants or other protective molecules such as flavonoids. Genetic variation in antioxidant enzymes is associated with more severe asthma.

Emotion

It is well known that emotional factors may influence asthma both acutely and chronically, but there is no evidence that patients with the disease are any more psychologically disturbed than their non-asthmatic peers. An asthma attack is a frightening experience, especially when of sudden and unexpected onset. Patients at special risk of life-threatening attacks are understandably anxious.

Drugs

Non-steroid anti-inflammatory drugs (NSAIDs). NSAIDs, particularly aspirin and propionic acid derivatives, e.g. indometacin and ibuprofen, have a role in the development and precipitation of asthma in approximately 5% of patients. NSAID intolerance is especially prevalent in those with both nasal polyps and asthma and is not infrequently associated with a triad of asthma, rhinitis and flushing on drug exposure. In susceptible subjects exposure to NSAIDs reveals an imbalance in the metabolism of arachidonic acid. NSAIDs inhibit arachidonic acid metabolism via the cyclo-oxygenase (COX) pathway, preventing the synthesis of certain prostaglandins. In aspirin-intolerant asthma there is reduced production of PGE_2 which, in a sub-proportion of genetically susceptible subjects, induces the overproduction of cysteinyl leukotrienes by eosinophils, mast cells and macrophages. In such patients there is evidence for genetic polymorphisms involving the enzymes and receptors of the leukotriene generating pathway (Fig. 14.32). Interestingly, asthma in intolerant patients is not precipitated by COX-2 inhibitors, indicating that it is blockade of the COX-1 isoenzyme that is linked to impaired PGE_2 production.

Beta-blockers. The airways have a direct parasympathetic innervation that tends to produce bronchoconstriction. There

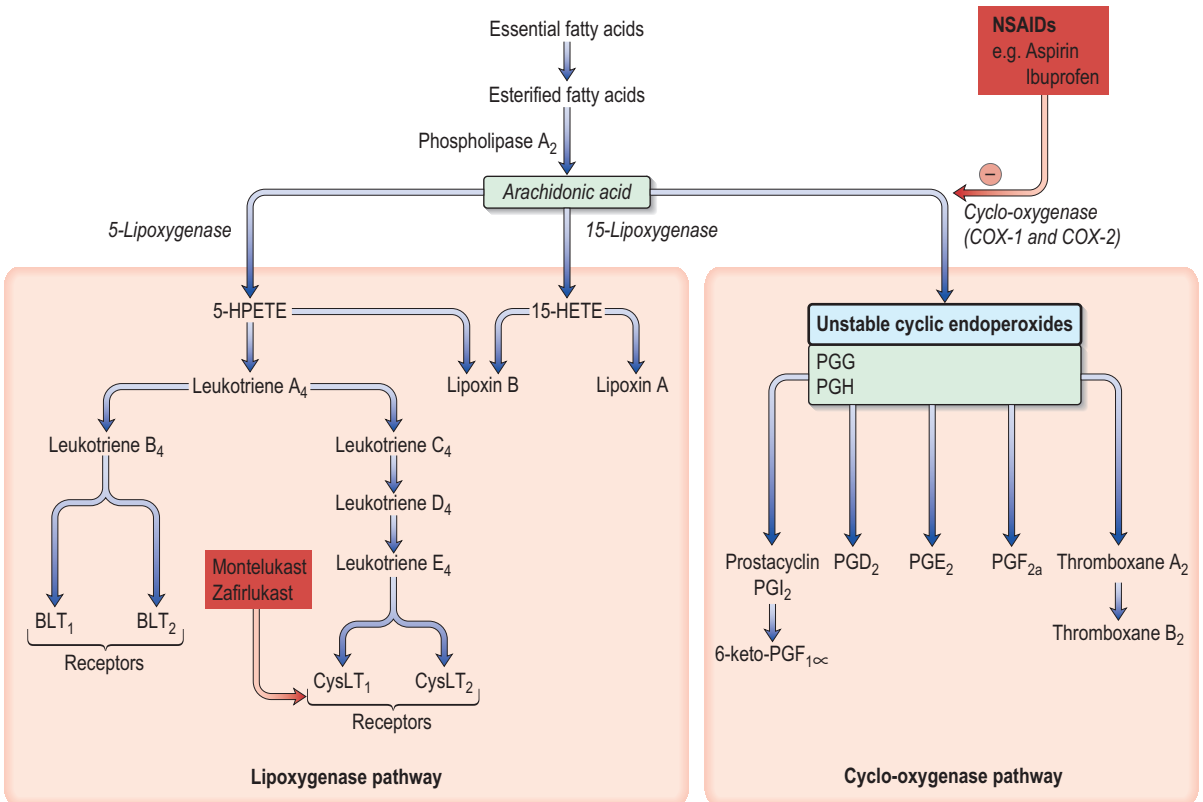


Fig. 14.32 **Arachidonic acid metabolism and the effect of drugs.** The sites of action of NSAIDs (e.g. aspirin, ibuprofen) are shown. The enzyme cyclo-oxygenase occurs in three isoforms, COX-1 (constitutive), COX-2 (inducible) and COX-3 (in brain). PG, prostaglandin; BLT, B leukotriene receptor; cysLT, cysteinyl leukotriene receptor.

is no direct sympathetic innervation of the smooth muscle of the bronchi, and antagonism of parasympathetically induced bronchoconstriction is critically dependent upon circulating epinephrine (adrenaline) acting through β_2-receptors on the surface of smooth muscle cells. Inhibition of this effect by β-adrenoceptor-blocking drugs such as propranolol leads to bronchoconstriction and airflow limitation, but only in asthmatic subjects. The so-called selective β_1-adrenergic-blocking drugs such as atenolol may still induce attacks of asthma; their use to treat hypertension or angina in asthmatic patients is best avoided.

Allergen-induced asthma

The experimental inhalation of allergen by atopic asthmatic individuals leads to the development of different types of reaction, as illustrated in Figure 14.33.

Immediate asthma (early reaction). Airflow limitation begins within minutes of contact with the allergen, reaches its maximum in 15–20 minutes and subsides by 1 hour.

Dual and late-phase reactions. Following an immediate reaction many asthmatics develop a more prolonged and sustained attack of airflow limitation that responds less well to inhalation of bronchodilator drugs such as salbutamol. Isolated late-phase reactions with no preceding immediate response can occur after the inhalation of some occupational sensitizers such as isocyanates. During and up to several weeks after the exposure, the airways are hyperresponsive, which may explain persisting symptoms after allergen exposure.

Pathogenesis

The pathogenesis of asthma is complex and not fully understood. It involves a number of cells, mediators, nerves and vascular leakage that can be activated by several different mechanisms, of which exposure to allergens is among the most significant (Fig. 14.34). The varying clinical severity and chronicity of asthma is dependent on an interplay between airway inflammation and airway wall remodelling. The inflammatory component is driven by Th2-type T lymphocytes which facilitate IgE synthesis through production of IL-4 and eosinophilic inflammation through IL-5 (Fig. 14.34). However, as the disease becomes more severe and chronic and loses its sensitivity to corticosteroids, there is greater evidence of a Th1 response with release of mediators such as TNF-α and associated tissue damage, mucous metaplasia and aberrant epithelial and mesenchymal repair.

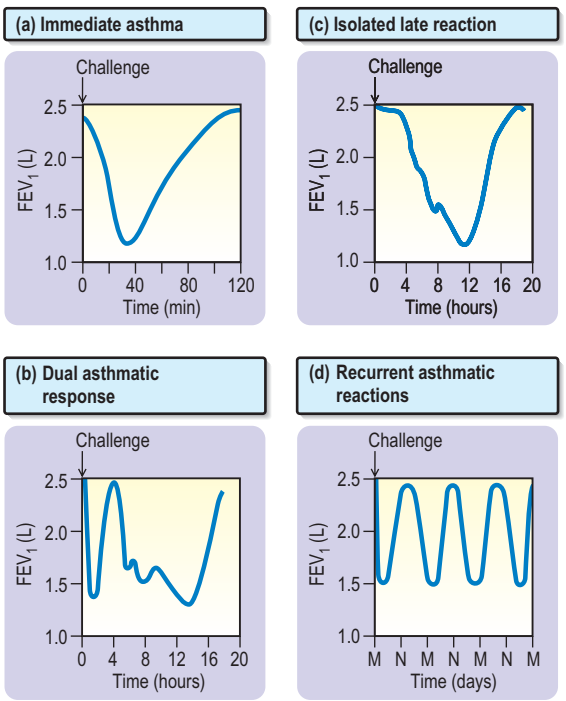


Fig. 14.33 **Different types of asthmatic reactions following challenge with allergen.** M, midnight; N, noon.

Inflammation

Several key cells are involved in the inflammatory response that characterizes all types of asthma.

Mast cells (see also p. 57). These are increased in the epithelium, smooth muscle and mucous glands in asthma and release powerful preformed and newly generated mediators that act on smooth muscle, small blood vessels, mucus-secreting cells and sensory nerves, such as histamine, tryptase, PGD_2 and LTC_4, and its metabolites LTD_4 and LTE_4 (previously known as slow reacting substance of anaphylaxis, SRS-A), which cause the immediate asthmatic reaction. Mast cells are inhibited by such drugs as sodium cromoglicate and β_2-agonists which might contribute to their therapeutic efficacy in preventing acute bronchoconstriction triggered by indirect challenges. Mast cells also release an array of cytokines, chemokines and growth factors that contribute to the late asthmatic response and more chronic aspects of asthma.

Eosinophils. These cells are found in large numbers in the bronchial wall and secretions of asthmatics. They are attracted to the airways by the eosinophilopoietic cytokines IL-3, IL-5 and GM-CSF as well as by chemokines which act on type 3 C-C chemokine receptors (CCR-3) (i.e. eotaxin, RANTES, MCP-1, MCP-3 and MCP-4). These mediators also prime eosinophils for enhanced mediator secretion. When activated, they release LTC_4, and basic proteins such as major basic protein (MBP), eosinophil cationic protein (ECP) and peroxidase (EPX) that are toxic to epithelial cells. Both the number and activation of eosinophils are rapidly decreased by corticosteroids. Sputum eosinophilia is of diagnostic help as well as providing a biomarker of response to controller therapy.

Dendritic cells and lymphocytes. These cells are abundant in the mucous membranes of the airways and the alveoli. Dendritic cells have a role in the initial uptake and presentation of allergens to lymphocytes. T helper lymphocytes ($CD4^+$) show evidence of activation (Fig. 14.34) and the release of their cytokines plays a key part in the migration and activation of mast cells (IL-3, IL-4, IL-9 and IL-13) and eosinophils (IL-3, IL-5, GM-CSF). In addition, production of IL-4 and IL-13 helps maintain the proallergic Th2 phenotype, favouring switching of antibody production by B lymphocytes to IgE. In mild/moderate asthma there occurs a selective upregulation of Th2 T cells with reduced evidence of the Th1 phenotype (producing gamma-interferon, TNF-α and IL-2), although additional Th1 prominence may accompany more severe disease. This polarization is mediated by dendritic cells and

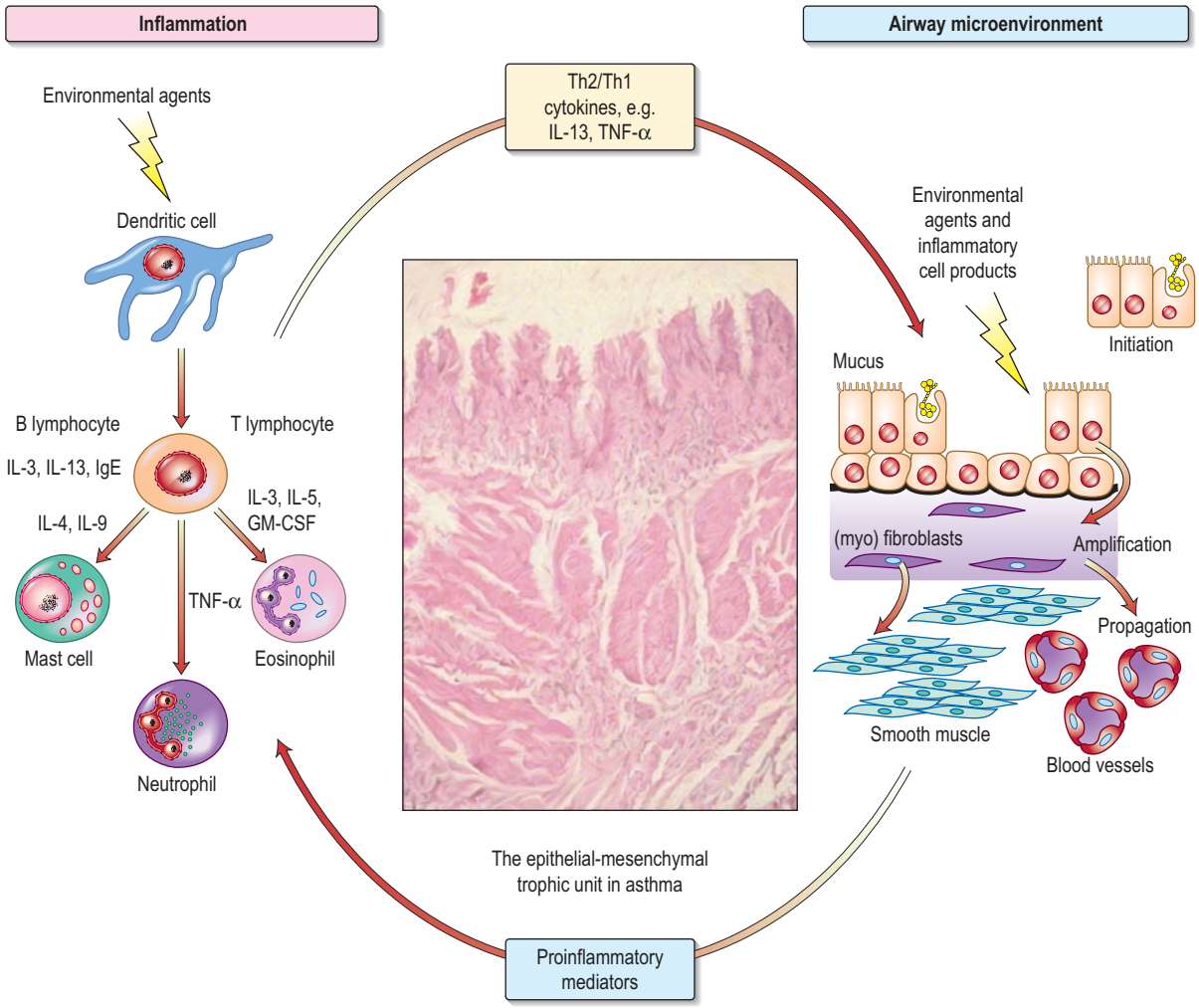


Fig. 14.34 **Inflammatory and remodelling responses in asthma with activation of the epithelial mesenchymal trophic unit.** Epithelial damage alters the set point for communication between bronchial epithelium and underlying mesenchymal cells, leading to myofibroblast activation, an increase in mesenchymal volume, and induction of structural changes throughout airway wall. Adapted from Holgate ST, Polosa R. The mechanisms, diagnosis, and management of severe asthma in adults. *Lancet* 2006; **368**: 780–793 with permission from Elsevier.

involves a combination of antigen presentation, costimulation and exposure to polarizing cytokines. The activity of both macrophages and lymphocytes is influenced by corticosteroids but not β_2-adrenoceptor agonists.

Remodelling

A characteristic feature of chronic asthma is an alteration of structure and functions of the formed elements of the airways. Together, these structural changes interact with inflammatory cells and mediators to cause the characteristic features of the disease. Deposition of matrix proteins, swelling and cellular infiltration cause an expansion of the submucosa beneath the epithelium so that for a given degree of smooth muscle shortening there is excess airway narrowing. Swelling outside the smooth muscle layer spreads the retractile forces exerted by the surrounding alveoli over a greater surface area so that the airways close more easily. Several factors contribute to these changes.

The epithelium. In asthma the epithelium of the conducting airways is stressed and damaged with loss of ciliated columnar cells on to the lumen. Metaplasia occurs with a resultant increase in the number and activity of mucus-secreting goblet cells. The epithelium is a major source of mediators, cytokines and growth factors that serve to enhance inflammation and promote tissue remodelling (Fig. 14.34). Damage and activation of the epithelium make it more vulnerable to infection by common respiratory viruses, e.g. rhinovirus, coronavirus, and to the effects of air pollutants. Increased production of nitric oxide (NO), due to the increased expression of inducible NO synthase, is a feature of epithelial damage and activation and the measurement of exhaled NO is proving useful as a non-invasive test of continuing inflammation (p. 826).

Epithelial basement membrane. A pathognomonic feature of asthma is the deposition of repair collagens (types I, III and V) and proteoglycans in the lamina reticularis beneath the basement membrane. This, along with the deposition of other matrix proteins such as laminin, tenascin and fibronectin, causes the appearance of a thickened basement membrane observed by light microscopy in asthma. This collagen deposition reflects activation of an underlying sheath of fibroblasts that transform into contractile myofibroblasts which

also have an increased capacity to secrete matrix. Aberrant signalling between the epithelium and underlying myofibroblasts is thought to be the principal cause of airway wall remodelling, since the cells are prolific producers of a range of tissue growth factors such as epidermal growth factor (EGF), transforming growth factor (TGF) -α and -β, connective tissue-derived growth factor (CTGF), platelet-derived growth factor (PDGF), endothelin (ET), insulin-like growth factors (IGF), nerve growth factors and vascular endothelial growth factors (Fig. 14.34). The same interaction between epithelium and mesenchymal tissues is central to branching morphogenesis in the developing fetal lung. It has been suggested that these mechanisms are reactivated in asthma, but instead of causing airway growth and branching, they lead to thickening of the airway wall (remodelling, Fig. 14.34). Increased deposition of collagens, proteoglycans and matrix proteins creates a microenvironment conducive to ongoing inflammation since these complex molecules also possess cell-signalling functions, which aid cell movement, prolong inflammatory cell survival and prime them for mediator secretion.

Smooth muscle. A prominent feature of asthma is hyperplasia of the helical bands of airway smooth muscle. In addition to increasing in amount, the smooth muscle alters in function to contract more easily and stay contracted because of a change in actin–myosin cross-link cycling. These changes allow the asthmatic airways to contract too much and too easily at the least provocation. Asthmatic smooth muscle also secretes a wide range of cytokines, chemokines and growth factors that help sustain the chronic inflammatory response. *ADAM33*, the newly described asthma gene, may be involved in driving increased airway smooth muscle and other features of remodelling through increased availability of growth factors.

Nerves. Neural reflexes, both central and peripheral, contribute to the irritability of asthmatic airways. Central reflexes involve stimulation of nerve endings in the epithelium and submucosa with transmission of impulses via the spinal cord and brain back down to the airways where release of acetylcholine from nerve endings stimulates M_3 receptors on smooth muscle causing contraction. Local neural reflexes involve antidromic neurotransmission and the release of a variety of neuropeptides. Some of these are smooth muscle contractants (substance P, neurokinin A), some are vasoconstrictors (e.g. calcitonin gene-related peptide, CGRP) and some vasodilators (e.g. neuropeptide Y, vasoactive intestinal polypeptide). The polymorphism in the neuropeptide S receptor (GPR 154) is associated with asthma susceptibility. Bradykinin generated by tissue and serum proteolytic enzymes (including mast cell tryptase and tissue kallikrein) is also a potent stimulus of local neural reflexes involving (non-myelinated) nerve fibres.

Clinical features

The principal symptoms of asthma are wheezing attacks and episodic shortness of breath. Symptoms are usually worst during the night, this being a particularly good marker of uncontrolled disease. Cough is a frequent symptom that sometimes predominates, especially in children in whom nocturnal cough can be a presenting feature. There exists great variation in the frequency and duration of the attacks. Some patients have only one or two attacks a year that last for a few hours, whilst others have attacks lasting for weeks. Some patients have chronic symptoms that persist, on top of which there are fluctuations. Attacks may be precipitated by a wide range of triggers (Fig. 14.31). Asthma is a major cause of impaired quality of life with impact on work and recreational, as well as physical activities, and emotions.

Investigations

There is no single satisfactory diagnostic test for all asthmatic patients.

Lung function tests

Peak expiratory flow rate (PEFR) measurements on waking, prior to taking a bronchodilator and before bed after a bronchodilator, are particularly useful in demonstrating the variable airflow limitation that characterizes the disease. An example is shown in Figure 14.14 (p. 825). The diurnal variation in PEFR is a good measure of asthma activity and is of help in the longer-term assessment of the patient's disease and its response to treatment. To assess possible occupational asthma, peak flows need to be measured for at least 2 weeks at work and 2 weeks off work.

Spirometry is useful, especially in assessing reversibility. Asthma can be diagnosed by demonstrating a greater than 15% improvement in FEV_1 or PEFR following the inhalation of a bronchodilator. However, this degree of response may not be present if the asthma is in remission or in severe chronic asthma when little reversibility can be demonstrated or if the patient is already being treated with long-acting bronchodilators.

The carbon monoxide (CO) transfer test is normal in asthma.

Exercise tests

These have been widely used in the diagnosis of asthma in children. Ideally, the child should run for 6 minutes on a treadmill at a workload sufficient to increase the heart rate above 160 beats per minute. Alternative methods use cold air challenge, isocapnoeic hyperventilation (forced overbreathing with artificially maintained P_aCO_2) or aerosol challenge with hypertonic solutions. A negative test does not automatically rule out asthma.

Histamine or methacholine bronchial provocation test (see p. 848)

This test indicates the presence of airway hyperresponsiveness, a feature found in most asthmatics, and can be particularly useful in investigating those patients whose main symptom is cough. The test should not be performed on individuals who have poor lung function ($FEV_1 < 1.5$ L) or a history of 'brittle' asthma. In children, controlled exercise testing as a measure of BHR is often easier to perform.

Trial of corticosteroids

All patients who present with severe airflow limitation should undergo a formal trial of corticosteroids. Prednisolone 30 mg orally should be given daily for 2 weeks with lung function measured before and immediately after the course. A substantial improvement in FEV_1 (>15%) confirms the presence of a reversible element and indicates that the administration of inhaled steroids will prove beneficial to the patient. If the trial is for 2 weeks or less, the oral corticosteroid can be withdrawn without tailing off the dose,

and should be replaced by inhaled corticosteroids in those who have responded.

Exhaled nitric oxide (NO), a measure of airway inflammation and an index of corticosteroid response, is used in children as a test for the efficacy of corticosteroids.

Blood and sputum tests

Patients with asthma may have an increase in the number of eosinophils in peripheral blood ($>0.4 \times 10^9$/L). The presence of large numbers of eosinophils in the sputum is a more useful diagnostic tool.

Chest X-ray

There are no diagnostic features of asthma on the chest X-ray, although overinflation is characteristic during an acute episode or in chronic severe disease. A chest X-ray may be helpful in excluding a pneumothorax, which can occur as a complication, or in detecting the pulmonary shadows associated with allergic bronchopulmonary aspergillosis.

Skin tests

Skin-prick tests (SPT) should be performed in all cases of asthma to help identify allergic causes. Measurement of allergen-specific IgE in the serum is also helpful if SPT facilities are not available, if the patient is taking antihistamines or if a wide range of allergens are being investigated. Asthma frequently occurs in conjunction with other atopic disorders, especially rhinitis.

Allergen provocation tests

Allergen challenge is not required in the clinical investigation of patients, except in cases of suspected occupational asthma. Another controversial exception is the investigation of food allergy causing asthma. This diagnosis can be difficult, although many patients are concerned about the possibility. In the absence of any obvious allergy, e.g. peanut or milk, if the patient has asthma without any other systemic features, then food allergy is most unlikely to be the cause. Open food challenges are unreliable and if the diagnosis is seriously entertained, blind oral challenges with the food disguised in opaque gelatine capsules are necessary to confirm or refute a causative link (see p. 238). There is much speculation about food intolerance (as opposed to allergy) and asthma including the role of food additives, which occasionally can precipitate severe attacks.

Management

The aims of treatment are to:

- abolish symptoms
- restore normal or best possible lung function
- reduce the risk of severe attacks
- enable normal growth to occur in children
- minimize absence from school or employment.

This involves:

- patient and family education about asthma
- patient and family participation in treatment
- avoidance of identified causes where possible
- use of the lowest effective doses of convenient medications to minimize short-term and long-term side-effects.

Many asthmatics belong to self-help groups whose aim is to further their understanding of the disease and to foster self-confidence and fitness.

Control of extrinsic factors

Measures must be taken to avoid causative allergens such as pets, moulds and certain foodstuffs (see allergic rhinitis), particularly in childhood. Avoidance of the house-dust mite is very difficult. A recent Cochrane review confirms earlier data that there is no evidence for the effectiveness of physical or chemical measures to control house-dust mite levels. The use of covers for bedding and changes to living accommodation has no beneficial effect on outcomes Active and passive smoking should be avoided, as should beta-blockers in either tablet or eye drop form. Individuals intolerant to aspirin should avoid all NSAIDs. Other agents (e.g. preservatives and colouring materials such as tartrazine) should be avoided if shown to be a causative factor. Fifty per cent of individuals sensitized to occupational agents may be cured if they are kept permanently away from exposure. The remaining 50% continue to have symptoms that may be as severe as when exposed to materials at work, especially if they were symptomatic for a long time before the diagnosis was made.

This emphasizes that:

- The rapid identification of extrinsic causes of asthma and their removal is necessary wherever possible (e.g. occupational agents, family pets).
- Once extrinsic asthma is initiated, it may become self-perpetuating, possibly by non-immune mechanisms.

Drug treatment

The mainstay of asthma therapy is the use of therapeutic agents delivered as aerosols or powders directly into the lungs (Practical box 14.4). The advantages of this method of administration are that drugs are delivered direct to the airways and first-pass metabolism in the liver is avoided; thus lower doses are necessary and systemic unwanted effects are minimized.

Practical Box 14.4 Inhaled therapy

Patients should be taught how to use inhalers and their technique checked regularly.

Use of a metered-dose inhaler

1. The canister is shaken.
2. The patient exhales to functional residual capacity (not residual volume), i.e. normal expiration.
3. The aerosol nozzle is placed to the open mouth.
4. The patient simultaneously inhales rapidly and activates the aerosol.
5. Inhalation is completed.
6. The breath is held for 10 seconds if possible. Even with good technique only 15% of the contents is inhaled and 85% is deposited on the wall of the pharynx and ultimately swallowed.

   NB Chlorofluorocarbon (CFC) propellants have been/are being replaced by hydrofluoralkane (HFA) propellants. The new aerosols may feel and taste different and patients need reassurance of their efficacy.

Spacers

These are plastic conical spheres inserted between the patient's mouth and the inhaler. They are designed to reduce particle velocity so that less drug is deposited in the mouth. Spacers also diminish the need for coordination between aerosol activation and inhalation. They are useful in children and in the elderly and they reduce the risk of candidiasis.

Both national and international guidelines have been published on the stepwise treatment of asthma (Box 14.3) based on three principles:

- Asthma self-management with regular asthma monitoring using PEF meters and individual treatment plans discussed with each patient and written down.
- The appreciation that asthma is an inflammatory disease and that anti-inflammatory (controller) therapy should be started even in mild cases.
- Use of short-acting inhaled bronchodilators (e.g. salbutamol and terbutaline) only to relieve breakthrough symptoms. Increased use of bronchodilator treatment to relieve increasing symptoms is an indication of deteriorating disease.

A list of drugs used in asthma is shown in Box 14.4. These are given in a stepwise fashion (1 to 6) as indicated in Box 14.3.

Once the asthma is under control, for at least 2–3 months, the drug therapy should be re-assessed in order to use the minimal dosage of inhaled steroids, combined with an oral long-acting β_2 agonist (stepdown). These agents should *not* be stopped for at least 6 months, if at all.

β_2-adrenoceptor agonists

The most widely used bronchodilator preparations contain β_2-adrenoceptor agonists that are selective for the respiratory tract and do not stimulate the β_1 adrenoceptors of the myocardium. These drugs are potent bronchodilators because they relax the bronchial smooth muscle. Such treatment is very effective in relieving symptoms but does little for the underlying inflammatory nature of the disease.

Uses

- *Mildest asthmatics with intermittent attacks*. Only these people should rely on bronchodilator treatment alone.

Short-acting β agonists (SABAs) such as salbutamol (100 μg), known as albuterol in the USA, or terbutaline (250 μg) should be prescribed as 'two puffs as required'. Some patients use nebulizers at home for self-administration of salbutamol or terbutaline. Such treatment is effective, but patients must not rely on repeated home administration of nebulized β_2-adrenoceptor agonists for worsening asthma, and must be encouraged to seek medical advice urgently if their condition does not improve. The excessive use of β_2 agonists has been linked to the two epidemics of asthma mortality in the 1960s and 1980s.

Box 14.4 Drugs used in asthma

Short-acting relievers

Inhaled β_2 agonists (e.g. salbutamol (albuterol in USA), terbutaline)

Long-acting relief/disease controllers

Inhaled long-acting β_2 agonists (e.g. salmeterol, formoterol)
Inhaled corticosteroids (e.g. beclometasone, budesonide, fluticasone)
Compound inhaled salmeterol and fluticasone
Sodium cromoglicate
Leukotriene modifiers (e.g. montelukast, zafirlukast, zileuton)

Other agents with bronchodilator activity

Inhaled antimuscarinic agents (e.g. ipratropium, oxitropium)
Theophylline preparations
Oral corticosteroids (e.g. prednisolone 40 mg daily)

Steroid-sparing agents

Methotrexate
Ciclosporin
Intravenous immunoglobulin
Anti-IgE monoclonal antibody – omalizumab
Etanercept (p. 76)

Box 14.3 The stepwise management of asthma

Step	PEFR	Treatment
1 Occasional symptoms, less frequent than daily	100% predicted	**As-required short-acting β_2 agonists** If used more than once daily, move to step 2
2 Daily symptoms	≤80% predicted	**Regular inhaled preventer therapy**: Anti-inflammatory drugs: inhaled low-dose corticosteroids up to 800 μg daily. Leukotriene receptor agonists (LTRA), theophylline and sodium cromoglicate are less effective If not controlled, move to step 3
3 Severe symptoms	50–80% predicted	**Inhaled corticosteroids and long-acting inhaled β_2 agonist** Continue inhaled corticosteroid Add – regular inhaled long-acting β_2 agonist (LABA) Still not controlled, add either LTRA, modified release oral theophylline or β_2 agonist If not controlled, move to Step 4
4 Severe symptoms uncontrolled with high-dose inhaled corticosteroids	50–80% predicted	**High-dose inhaled corticosteroid and regular bronchodilators** Increase high-dose inhaled corticosteroids up to 2000 μg daily Plus regular long-acting β_2 agonists Plus either LTRA or modified release theophylline or β_2 agonist
5 Severe symptoms deteriorating	≤50% predicted	**Regular oral corticosteroids** Add prednisolone 40 mg daily to step 4
6 Severe symptoms deteriorating in spite of prednisolone	≤30% predicted	Hospital admission

Short-acting bronchodilator treatment taken at any step on an as-required basis.

- *SABAs* can be taken at any step, as and when required from **step 1 to step 6**.
- *Poorly controlled asthmatics* on standard doses of inhaled steroids. These patients require salmeterol or formoterol, which are highly selective and potent long-acting β_2-adrenoceptor agonists (LABA) effective by inhalation for up to 12 hours, thereby reducing the need for administration to once or twice daily. Long-acting β_2-adrenoceptor agonists improve symptoms and lung function and reduce exacerbations in patients. They should never be used alone but always in combination with an inhaled corticosteroid. Increasingly these drugs are administered as fixed-dose combinations with corticosteroids (salmeterol/fluticasone and formoterol/budesonide) in the same inhaler **(step 3)**.

β_2-adrenoceptor agonists are less effective when given by mouth than when the drug is inhaled, and to help those who cannot coordinate activation of the aerosol and inhalation, several breath-activated or dry powder devices have been developed.

Antimuscarinic bronchodilators

Muscarinic receptors are found in the respiratory tract; large airways contain mainly M_3 receptors whereas the peripheral lung tissue contains M_3 and M_1 receptors (see p. 814). Non-selective muscarinic antagonists – ipratropium bromide (20–40 μg three or four times daily) or oxitropium bromide (200 μg twice daily) – by aerosol inhalation can be useful during asthma exacerbations, but their overall role in asthma is limited.

Anti-inflammatory drugs

Sodium cromoglicate and nedocromil sodium prevent activation of many inflammatory cells, particularly mast cells, eosinophils and epithelial cells, but not lymphocytes, by blocking a specific chloride channel which in turn prevents calcium *influx*. These drugs are effective in patients with *milder asthma* (**step 2**). Sodium cromoglicate is taken regularly either in the form of a Spincap containing 20 mg or in aerosol form from a metered-dose inhaler delivering 5 mg per puff. The dose should be two puffs four times daily from an inhaler, or one Spincap three or four times daily. Nedocromil sodium is taken as an aerosol at a dose of 4 mg (two puffs) two to four times daily.

Asthma guidelines advise that inhaled corticosteroids are more efficacious than the chromones, but the latter are free of side-effects and, therefore, may offer some advantages in children, especially when there is strong evidence for an allergic component.

Inhaled corticosteroids

All patients who have regular persistent symptoms (even mild symptoms) need regular treatment with inhaled corticosteroids delivered in a stepwise fashion (**from step 2 upwards**) or as a high dose followed by a reduction to maintenance levels. Beclometasone dipropionate (BDP) is the most widely used inhaled steroid and is available in doses of 50, 100, 200 and 250 μg per puff. Other inhaled steroids include budesonide, fluticasone, mometasone and triamcinolone.

Much of the inhaled dose does not reach the lung but is either swallowed or exhaled. Deposition in the lung varies between 10% and 25% depending on inhaler technique and the technical characteristics of the aerosol device. Drug which is deposited in the airways reaches the systemic circulation directly, through the bronchial circulation, while any drug that is swallowed has to pass through the liver before it can reach the systemic circulation. Gram for gram, fluticasone and mometasone are more potent than beclometasone with considerably less systemic bioavailability, owing to their greater sensitivity to hepatic metabolism. The newer hydrofluoroalkane (HFA) aerosols of beclometasone deliver a higher proportion of useable drug than the old CFC-based aerosols, and the effective dosage of HFA-beclometasone is equivalent to the same dose of HFA-fluticasone, whereas previously the effective dose ratio was 2:1. Absorption of beclometasone and budesonide does not seem to present a risk at doses up to 800 μg/day, but when using high-dose inhaled steroids in patients who have not responded to standard doses, fluticasone or mometasone may be preferred because of their lower bioavailability. The dose–response curve for inhaled corticosteroids is flat beyond 800 μg beclometasone or equivalent, and in patients with moderate asthma who are taking this daily, addition of salmeterol or formoterol is more effective than doubling the dose of inhaled corticosteroid.

The unwanted effects of inhaled corticosteroids are oral candidiasis (5% of patients), and hoarseness due to the effect of corticosteroids on the laryngeal muscles. Subcapsular cataract formation is rare but can occur in the elderly. Osteoporosis is a complication when inhaled corticosteroids are taken in high doses (beclometasone or budesonide >800 μg daily). In children, inhaled corticosteroids at doses greater than 400 μg daily have been shown to retard short-term growth. Inhaled corticosteroid use should be stepped down once asthma comes under control (p. 854). Candidiasis and GI absorption can be reduced by using spacers, mouthwashing and teeth cleaning after use. More recently inhaled corticosteroids, e.g. ciclesonide 80 μg daily, that are esterified in the lung thereby reducing systemic effects, have become available with a higher therapeutic index.

Asthmatic patients who smoke are less responsive to inhaled corticosteroids, and additional therapy, e.g. with leukotriene receptor antagonists, is required.

Many patients with anything more than mild/moderate asthma benefit from combination LABA/corticosteroid therapy and there is some evidence that the two drugs interact therapeutically. Adding a LABA once the dose of inhaled corticosteroid has reached 800–1000 μg is of proven greater benefit than further increasing the steroid dose.

Oral corticosteroids and steroid-sparing agents

Oral corticosteroids may be necessary for those individuals not controlled on inhaled corticosteroids (**step 5**). The dose should be kept as low as possible to minimize side-effects. The effect of short-term treatment with prednisolone 30 mg daily is shown in Figure 14.14 (p. 825). Some patients require continuing treatment with oral corticosteroids. Several studies suggest that treatment with low doses of **methotrexate** (15 mg weekly) can significantly reduce the dose of prednisolone needed to control the disease in some patients, and **ciclosporin** also improves lung function in some steroid-dependent asthmatics. Several other steroid-sparing strategies including ciclosporin and immunoglobulin have also been tried, but with varying success.

Ask the authors
www.studentconsult.com

Cysteinyl leukotriene receptor antagonists (LTRAs)

This class of anti-asthma therapy targets one of the principal asthma mediators by inhibiting the cysteinyl LT_1 receptor. A second receptor (cyst LT_2) has been identified on inflammatory cells. Montelukast, pranlukast (only available in South East Asia) and zafirlukast are given orally and are effective in a subpopulation of patients. However, it is not possible to predict which individuals will benefit: a 4-week trial of LTRA therapy is recommended before a decision is made to continue or stop. LTRAs should be considered in any patient who is not controlled on low to medium doses of inhaled steroids (**step 2**). Their action is additive to that of long-acting β_2 agonists. LTRAs are particularly useful in patients with aspirin-intolerant asthma, in those patients requiring high-dose inhaled or oral corticosteroids and in asthmatic smokers. Because these drugs are orally active they are helpful in asthma combined with rhinitis and in young children with asthma and/or virus associated wheezing.

Monoclonal antibodies

Newer agents that modulate IgE-associated inflammation are being developed. The most promising of these is a recombinant humanized monoclonal antibody that complexes with free IgE – omalizumab – blocking its interaction with mast cells and basophils. Clinical trials in children and adults with severe asthma despite corticosteroids show good efficacy when omalizumab is given subcutaneously once or twice each month depending on total serum IgE level and body weight. Recent proof-of-concept trials have also suggested a place for anti-TNF therapy (blocking monoclonal antibody or soluble receptor fusion protein, etanercept) in corticosteroid refractory severe asthma. There is a need to examine other biologicals as potential new controller therapies for the 5–10% of patients with severe disease, which accounts for a high proportion of the health costs of asthma.

Antibiotics

Although wheezing frequently occurs in infective exacerbations of COPD, there is limited evidence that antibiotics are helpful in the management of patients with asthma. Yellow or green sputum containing eosinophils and bronchial epithelial cells may be coughed up in acute exacerbations of asthma. This is normally due to viral, not bacterial, infection and antibiotics are not required. Occasionally, mycoplasma and *Chlamydia* infections can cause chronic relapsing asthma and in such cases the use of appropriate antimicrobials is worthwhile if a bacterial diagnosis has been established by culture or serology.

Asthma attack

Although these may occur spontaneously, asthma exacerbations are most commonly caused by lack of treatment adherence, respiratory virus infections associated with the common cold, and exposure to allergen or triggering drug, e.g. an NSAID. Whenever possible patients should have a written personalized plan that they can implement in anticipation of or at the start of an exacerbation that includes the early use of a short course of oral corticosteroids. If the PEFR is greater than 150 L/min, patients may improve dramatically on nebulized therapy and may not require hospital admission. Their regular treatment should be increased, to include treatment for 2 weeks with 30–60 mg of prednisolone followed by substitution by an inhaled corticosteroid preparation. Short courses of oral prednisolone can be stopped abruptly without tailing down the dose.

Acute severe asthma

The term 'status asthmaticus' was defined as asthma that had failed to resolve with therapy in 24 hours. Although this term is still used occasionally, it has been mainly discarded and replaced by 'acute severe asthma', i.e. severe asthma that has not been controlled by the patient's use of medication.

Patients with *acute severe asthma* typically have:

- inability to complete a sentence in one breath
- respiratory rate ≥ 25 breaths per minute
- tachycardia ≥ 110 beats/min (pulsus paradoxus, p. 690, is not useful as it is only present in 45% of cases)
- PEFR < 50% of predicted normal or best.

Features of life-threatening attacks are:

- a silent chest, cyanosis or feeble respiratory effort
- exhaustion, confusion or coma
- bradycardia or hypotension
- PEFR < 30% of predicted normal or best (approximately 150 L/min in adults).

Arterial blood gases should always be measured in asthmatic patients requiring admission to hospital. Pulse oximetry is useful in monitoring oxygen saturation during the admission and reduces the need for repeated arterial puncture. Features suggesting very severe life-threatening attacks are:

- a high P_aCO_2 >6 kPa
- severe hypoxaemia P_aO_2 < 8 kPa despite treatment with oxygen
- a low and falling arterial pH.

Treatment (Emergency box 14.2) is commenced with 5 mg of nebulized salbutamol or 10 mg terbutaline with oxygen as the driving gas. Nebulized antimuscarinics (e.g. ipratropium bromide) are also helpful. A chest X-ray is taken to exclude a pneumothorax. If no improvement occurs with nebulized therapy, intravenous infusions of β_2-adrenoceptor agonist (salbutamol or terbutaline 250 μg over 10 min) and/or magnesium sulphate (1.2–2 g over 20 min, which also relaxes airways smooth muscle) should be used. Intravenous aminophylline is not given as trials show marginal benefit and side-effects such as nausea and vomiting, cardiac arrhythmia and CNS side-effects are problematic. Hydrocortisone 200 mg i.v. should be administered 4-hourly for 24 hours, and 60 mg of prednisolone should be given orally daily. In patients who do not respond to this regimen, ventilation is often required.

Ideally, patients should be kept in hospital for at least 5 days, since the majority of sudden deaths occur 2–5 days after admission. During this time oxygen saturation should be monitored by oximetry. Oral prednisolone can be reduced from 60 mg to 30 mg once improvement occurs. Further reduction should be gradual on an outpatient basis until an appropriate maintenance dose or substitution by inhaled corticosteroid aerosols can be achieved.

A recent approach for moderate to severe persistent asthma is bronchial thermoplasty. This bronchoscopic procedure reduces the mass of airway smooth muscle, reducing bronchoconstriction, and is being evaluated.

Emergency Box 14.2

Treatment of severe asthma

At home

1. The patient is assessed. Tachycardia, a high respiratory rate and inability to speak in sentences indicate a severe attack.
2. If the PEFR is less than 150 L/min (in adults), an ambulance should be called. (All doctors should carry peak flow meters.)
3. Nebulized salbutamol 5 mg or terbutaline 10 mg is administered.
4. Hydrocortisone sodium succinate 200 mg i.v. is given.
5. Oxygen 40–60% is given if available.
6. Prednisolone 60 mg is given orally.

At hospital

1. The patient is reassessed.
2. Oxygen 40–60% is given.
3. The PEFR is measured using a low-reading peak flow meter, as an ordinary meter measures only from 60 L/min upwards.
   Measure O_2 saturation with a pulse oximeter.
4. Nebulized salbutamol 5 mg or terbutaline 10 mg is repeated and administered 4-hourly.
5. Add nebulized ipratropium bromide 0.5 mg to nebulized salbutamol/terbutaline.
6. Hydrocortisone 200 mg i.v. is given 4-hourly for 24 hours.
7. Prednisolone is continued at 60 mg orally daily for 2 weeks.
8. Arterial blood gases are measured; if the P_aco_2 is greater than 7 kPa, ventilation should be considered.
9. A chest X-ray is performed to exclude pneumothorax.
10. One of the following intravenous infusions is given if no improvement is seen:
    salbutamol 3–20 μg/min, or
    terbutaline 1.5–5.0 μg/min, or
    magnesium sulphate 1.2–2 g over 20 min.

Management of catastrophic sudden severe (brittle) asthma

This is an unusual variant of asthma in which patients are at risk from sudden death in spite of the fact that their asthma may be well controlled between attacks. Severe life-threatening attacks may occur within hours or even minutes. Such patients require a carefully worked out management plan agreed by respiratory physician, primary care physician and patient, and require:

- emergency supplies of medications at home, in the car and at work
- oxygen and resuscitation equipment at home and at work
- nebulized β_2-adrenoceptor agonists at home and at work; inhaled long-acting β_2-agonists with a corticosteroid can be very effective
- self-injectable epinephrine (adrenaline): two Epipens of 0.3 mg epinephrine at home, at work and to be carried by the patient at all times
- prednisolone 60 mg
- Medic Alert bracelet.

On developing wheeze, the patient should attend the nearest hospital immediately. Direct admission to intensive care may be required.

Prognosis of asthma

Although asthma often improves in children as they reach their teens, the disease frequently returns in the second, third and fourth decades. In the past the data indicating a natural decrease in asthma through teenage years have led to childhood asthma being treated as an episodic disorder. However, airway inflammation is present continuously from an early age and usually persists even if the symptoms resolve. Moreover, airways remodelling accelerates the process of decline in lung function over time. This has led to a reappraisal of the treatment strategy for asthma, mandating the early use of controller drugs and environmental measures from the time asthma is first diagnosed.

PNEUMONIA

Pneumonia is defined as an inflammation of the substance of the lungs. It is usually caused by bacteria. Clinically it usually presents as an acute illness with cough, purulent sputum and fever together with physical signs or radiological changes compatible with consolidation of the lung. The advent of antibiotics has dramatically decreased the mortality from pneumonia among young people but it remains a dangerous condition and is a major cause of death over the age of 70 years. Bacterial pneumonia is more frequent in alcoholics and in HIV-infected individuals as compared to the general population. The causative agents are the same as found in community-acquired pneumonia in previously healthy people.

Classification

Pneumonia can be classified both anatomically and on the basis of the aetiology.

Classification by site

Pneumonias are either localized, with the whole of one or more lobes affected ('lobar pneumonia'), or diffuse, when they primarily affect the lobules of the lung, often in association with the bronchi and bronchioles ('bronchopneumonia').

Classification by aetiology

An aetiological factor can be discovered in approximately 75% of patients. The term 'atypical pneumonia' described pneumonia caused by agents such as *Mycoplasma, Legionella, Chlamydia* and *Coxiella burnetii*. While the clinical features of these pneumonias can differ from pneumococcal disease, there is a considerable overlap in clinical presentation and as these agents account for almost one-fifth of the cases of pneumonia (Table 14.13), the term 'atypical' has now been abandoned. Pneumonias may also result from:

- chemical causes, such as in the aspiration of vomit (see p. 861)
- radiotherapy (see p. 877)
- allergic mechanisms (see p. 875).

Mycobacterium tuberculosis causes pneumonia; it is considered separately (p. 863), since both its mode of presentation and its treatment are very different from the other infective agents.

Precipitating factors

- *Strep. pneumoniae* – often follows viral infection with influenza or parainfluenza.

FURTHER READING

British Thoracic Society. British guidelines on the management of asthma. *Thorax* 2008; **63**(SIV): 1–121.

Eder W, Ege MJ, von Mutius E. The asthma epidemic. *New England Journal of Medicine* 2006; **355**: 2226–2235.

Holgate ST, Polosa R. The mechanisms, diagnosis and management of severe asthma in adults. *Lancet* 2006; **368**: 780–793.

Kogevinas M, Zock JP, Jarvis D et al. Exposure to substances in the workplace and new onset asthma: an international prospective population based study (ECRHS-II). *Lancet* 2007; **370**: 336–341.

Lancet 2008; **272**. All articles devoted to asthma.

Table 14.13 The aetiology of pneumonia in the UK

Infecting agent	Frequency as a cause of pneumonia in the USA (%)	Clinical circumstances
Streptococcus pneumoniae	35–80	Community pneumonia patients usually previously fit
Mycoplasma pneumoniae	2–14	As above
Influenza A virus (usually with a bacterial component)	10–15	As above
Haemophilus influenzae	5–6	Pre-existing lung disease: COPD
Chlamydia pneumoniae	4–13	Community-acquired pneumonia
Chlamydia psittaci	4–13	Contact with birds (though not inevitable)
Staphylococcus aureus	3–14	Children, intravenous drug users, associated with influenza virus infections
Legionella pneumophila	2–15	Institutional outbreaks (hospitals and hotels), sporadic, endemic
Coxiella burnetii	1	Abattoir and animal-hide workers
Pseudomonas aeruginosa	<4–9	Cystic fibrosis
Enteric Gram-negative bacilli	6–12	
M. tuberculosis	<1–5	
Pneumocystis jiroveci *Actinomyces israelii* *Nocardia asteroides* Cytomegalovirus *Aspergillus fumigatus*	<1	AIDS, lymphomas, leukaemias, use of cytotoxic drugs and corticosteroids
None isolated	15–40	–

NB Causes vary in different countries.
Modified from Garau J, Calbo E. Community acquired pneumonia. *Lancet* 2008; **371**: 455–458 with permission from Elsevier.

- Hospitalized 'ill' patients – often infected with Gram-negative organisms.
- Cigarette smoking (the strongest independent risk factor for invasive pneumococcal disease).
- Alcohol excess.
- Bronchiectasis (e.g. in cystic fibrosis).
- Bronchial obstruction (e.g. carcinoma) – occasionally associated with infection with 'non-pathogenic' organisms.
- Immunosuppression (e.g. AIDS or treatment with cytotoxic agents) – organisms include *Pneumocystis jiroveci, Mycobacterium avium-intracellulare*, cytomegalovirus.
- Intravenous drug users – frequently associated with *Staph. aureus* infection.
- Inhalation from oesophageal obstruction – often associated with infection with anaerobes.

Clinical features

The clinical presentation varies according to the immune state of the patient and the infecting agent. In the most common type of pneumonia – caused by *Strep. pneumoniae* – there is often a preceding history of a viral infection.

With *Strep. pneumoniae* infection the patient rapidly becomes more ill with a high temperature (up to 39.5°C), pleuritic pain and a dry cough. A day or two later, rusty-coloured sputum is produced and at about the same time the patient may develop labial herpes simplex. The patient breathes rapidly and shallowly, the affected side of the chest moves less, and signs of consolidation may be present together with a pleural rub. See Box 14.5 for severe community-acquired pneumonia.

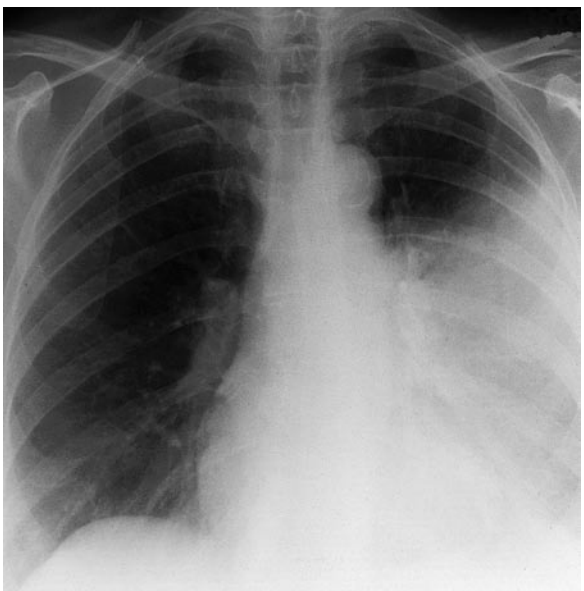


Fig. 14.35 **Chest X-ray showing lobar pneumonia.**

Investigations

Chest X-ray confirms the area of consolidation (Fig. 14.35), but radiological changes lag behind the clinical course so that X-ray changes may be minimal at the start of the illness. Conversely, consolidation may remain on the chest X-ray for several weeks after the patient is clinically cured. The chest X-ray usually returns to normal by 6 weeks, except in patients

with severe airflow limitation. Persistent changes on the chest X-ray after this time suggest a bronchial abnormality, usually a carcinoma, with persisting secondary pneumonia. Chest X-rays should rarely be repeated more frequently than at weekly intervals during the acute illness and then at 6 weeks after discharge from hospital.

In *Strep. pneumoniae* pneumonia, there is often a white blood cell count that is greater than 15×10^9/L (90% polymorphonuclear leucocytosis) and an erythrocyte sedimentation rate (ESR) greater than 100 mm/h with a CRP of >100 mg/L.

TYPES OF PNEUMONIA

The individual features of various pneumonias are given below. The overall investigation and management is shown in Figure 14.36 and discussed on page 862.

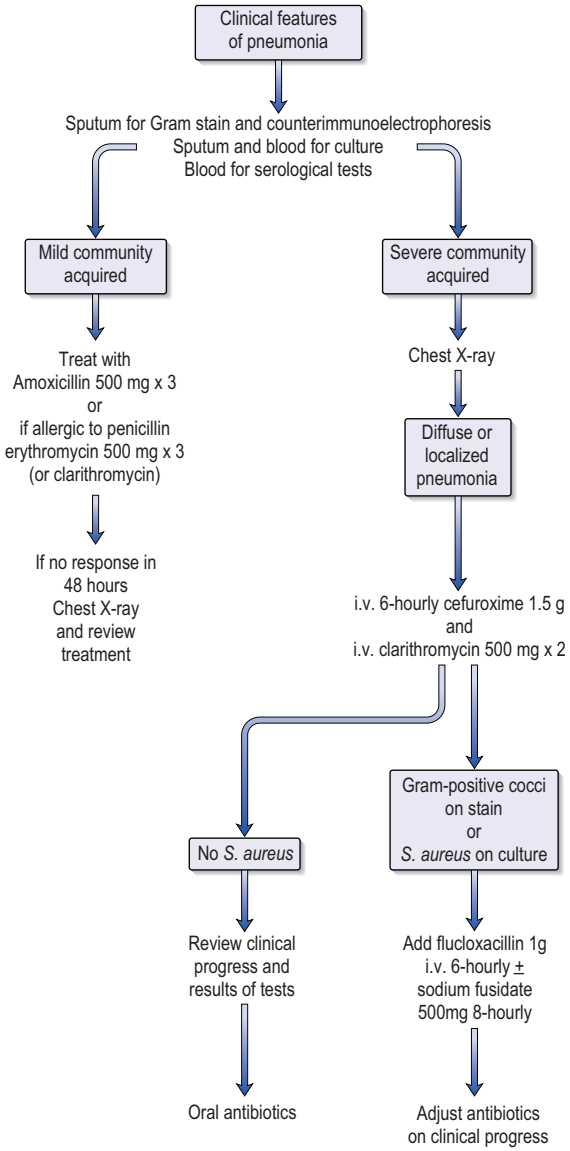


Fig. 14.36 **Algorithm for the management of community-acquired pneumonia** (see also p. 862 and Box 14.5). Review treatment after 48 hours.

Mycoplasma pneumonia

Mycoplasma pneumonia is relatively common and occurs in cycles of 3–4 years. It often occurs in patients in their teens and twenties, frequently amongst those living in boarding institutions. Generalized features such as headaches and malaise often precede the chest symptoms by 1–5 days. Cough may not be obvious initially and physical signs in the chest may be scanty.

On chest X-ray, usually only one lobe is involved but sometimes there may be florid bilateral disease. There is frequently no correlation between the X-ray appearances and the clinical state of the patient.

The white blood cell count is not raised. Cold agglutinins occur in half of the cases. The diagnosis is confirmed by rising titres of complement fixing antibodies. These have comparable high specificities and sensitivities (97–99%) to that of a PCR. A kit for PCR is becoming available. Treatment is with macrolides, e.g. erythromycin 500 mg four times daily (or clarithromycin or azithromycin) for 7–10 days. A tetracycline, e.g. doxycycline 100 mg twice daily, is also effective.

Although most patients recover in 10–14 days, the disease can be protracted, with cough and X-ray changes lasting for weeks and relapses occurring. Lung abscesses and pleural effusions are rare.

Extrapulmonary complications can occur at any time during the illness and occasionally dominate the clinical picture. Most are rare but they include:

- myocarditis and pericarditis
- rashes and erythema multiforme
- haemolytic anaemia and thrombocytopenia
- myalgia and arthralgia
- meningoencephalitis and other neurological abnormalities
- gastrointestinal symptoms (e.g. vomiting, diarrhoea).

Viral pneumonia

Primary viral pneumonia is uncommon in adults, influenza A virus or adenovirus infection being the commonest causes. More often, viral infection predisposes patients to bacterial pneumonia by damaging the respiratory epithelium and facilitating bacterial infection.

Cytomegalovirus pneumonia is seen in immunocompromised patients (p. 107).

Influenza A (H5N1) (p. 116) normally does not affect humans but can be transmitted from poultry (where it causes avian flu), when there has been unusually close contact. Patients present with fever, breathlessness, cough and diarrhoea. Lymphopenia and thrombocytopenia are present and pulmonary infiltrates are seen on chest X-ray. There is a high mortality.

Severe acute respiratory syndrome (SARS) is due to a novel coronavirus (see p. 119). The incubation period is approximately 5 days with spread between humans occurring mainly by droplet infection. The outbreak in 2003 affected many healthcare workers. Fever, malaise, headache and rigors were followed in the second week by cough, breathlessness and diarrhoea. Lymphopenia, thrombocytopenia and pulmo-

More online
www.studentconsult.com

nary infiltrates (mainly in the lower zones) occur. At the end of the second week 20% of patients deteriorate, developing ARDS (p. 917), and the mortality is high.

Other pneumonias

Haemophilus influenzae

H. influenzae is a frequent cause of exacerbation of chronic bronchitis and can cause pneumonia in COPD patients. The pneumonia can be diffuse or confined to one lobe. There are no special features to separate it from other bacterial pneumonias. It responds well to treatment with oral amoxicillin 500 mg × 3 daily.

Chlamydia psittaci (see also p. 127)

Typically the individual has been exposed to infected birds, especially parrots, but cases may occur without a history of contact. The incubation period is 1–2 weeks and the disease may pursue a very low-grade course over several months. Symptoms include malaise, high fever, cough and muscular pains. The liver and spleen are occasionally enlarged, and scanty 'rose spots' may be seen on the abdomen. The chest X-ray shows segmental or a diffuse pneumonia. Occasionally the illness presents with a high, swinging fever and dramatic prostration with photophobia and neck stiffness that can be confused with meningitis. The diagnosis is confirmed by the demonstration of a rising titre of complement-fixing antibody. Macrolides or tetracycline are the antibiotics of choice.

Chlamydia pneumoniae

Outbreaks of *C. pneumoniae* have been reported in institutions and within families, suggesting person-to-person spread without any avian or animal reservoir. Serological tests on patients admitted to hospital with community-acquired pneumonia suggest that 5–10% may be infected with *C. pneumoniae*. In general, the disease is mild with 50% of *C. pneumoniae* infections presenting as pneumonia, 28% as acute bronchitis, 10% with a flu-like illness and 12% with upper respiratory illnesses. Type-specific microimmunofluorescence tests are required to distinguish *C. pneumoniae* from *C. psittaci* and *C. trachomatis*. Treatment is with macrolides or tetracycline.

Staphylococcus aureus

Staph. aureus rarely causes pneumonia except after a preceding influenzal viral illness. The infection starts in the bronchi, leading to patchy areas of consolidation in one or more lobes, which break down to form abscesses. These may appear as cysts on the chest X-ray.

Pneumothorax, effusion and empyemas are frequent. Septicaemia develops with metastatic abscesses in other organs. All patients with staphylococcal pneumonia are very ill; intravenous antibiotics must be administered promptly, but are not always effective. Fulminating staphylococcal pneumonia can lead to death in hours.

Areas of pneumonia (septic infarcts) are also seen in staphylococcal septicaemia. This is frequently seen in intravenous drug abusers, and in patients with central catheters being used for parenteral nutrition. The infected puncture site is the source of the staphylococcus. Pulmonary symptoms are often few but breathlessness and cough occur and the chest X-ray reveals areas of consolidation. Abscess formation is frequent.

Diagnosis and treatment are shown in Figure 14.36.

Coxiella burnetii (Q fever) (see also p. 139)

The patient develops systemic symptoms of fever, malaise and headache, often associated with multiple lesions on the chest X-ray. The illness may run a chronic course and is occasionally associated with endocarditis. Diagnosis is made by an increase in the titre of complement-fixing antibody. Treatment is usually with macrolides or tetracycline. Severe cases may require rifampicin.

Legionella pneumophila

Three epidemiological patterns of Legionnaires' disease are recognized:

- outbreaks among previously fit individuals staying in hotels, institutions or hospitals where the shower facilities or cooling systems have been contaminated with the organism
- sporadic cases where the source of the infection is unknown; most cases involve middle-aged and elderly men who are smokers, but it is also seen in children
- outbreaks occurring in immunocompromised patients, e.g. on corticosteroid therapy.

Legionella grows well in water up to 40°C in temperature, and the infection is almost certainly spread by the aerosol route. Adequate chlorination and temperature control of the water supply are necessary to prevent the disease.

The incubation period is 2–10 days. Males are affected twice as commonly as females. The infection may be mild, but the characteristic picture is of malaise, myalgia, headache and a fever with rigors and a pyrexia of up to 40°C. Half of the patients have gastrointestinal symptoms, with nausea, vomiting, diarrhoea and abdominal pain. Patients may be acutely ill, with mental confusion and other neurological signs. Haematuria occurs and occasionally renal failure.

The patient is tachypnoeic with an initially dry cough that later may become productive and purulent. The chest X-ray usually shows lobar and then multilobar shadowing, sometimes with a small pleural effusion. Cavitation is rare.

A strong presumptive diagnosis of *L. pneumophila* infection is possible in the majority of patients if they have three of the four following features:

- a prodromal virus-like illness
- a dry cough, confusion or diarrhoea
- lymphopenia without marked leucocytosis
- hyponatraemia.

Hypoalbuminaemia and high serum levels of liver aminotransferases are also common in this disease.

Diagnosis is confirmed by direct immunofluorescent staining of the organism in the pleural fluid, sputum or bronchial washings. *Legionella* is not visible on Gram staining. Culture on special media is possible but takes up to 3 weeks. A urinary antigen test is commercially available and is highly specific. Retrospective confirmation can be made by demonstrating a four-fold increase in antibody titre in the blood.

Treatment is usually with one of the macrolides, azithromycin being the drug of choice. Quinolones are also effective and rifampicin can be added in very ill patients. Mortality can be up to 30% in elderly patients but most patients recover fully.

Gram-negative bacteria

These are the cause of many *hospital-acquired pneumonias* but are occasionally responsible for cases in the community.

Klebsiella pneumoniae

Pneumonia due to *Klebsiella* usually occurs in elderly people with a history of heart or lung disease, diabetes, alcohol excess or malignancy. The onset is often sudden, with severe systemic upset. The sputum is purulent, gelatinous or blood-stained. The upper lobes are more commonly affected and the consolidation is often extensive. There is often swelling of the infected lobe so that on the lateral chest X-ray there is bulging of the fissures. The organism can be found in the sputum or in the blood.

Treatment is dependent on the sensitivity of the organism, but a cephalosporin is usually optimal. The mortality is high, partly owing to the presence of the predisposing condition.

Pseudomonas aeruginosa

Pneumonia due to *Pseudomonas* is of considerable significance in patients with cystic fibrosis, since it correlates with a worsening clinical condition and mortality. It is also seen in patients with neutropenia following cytotoxic chemotherapy. The isolation of *P. aeruginosa* from sputum must be interpreted with care because the organism grows well on bacterial culture medium and may simply represent contamination from the upper airways.

Treatment. *Pseudomonas* and other Gram-negative organisms respond well to treatment with the 4-quinolone antibiotic ciprofloxacin (200–400 mg i.v. over 30–60 minutes twice daily) or ceftazidime (2 g bolus i.v. 8-hourly). Ticarcillin (15–20 g daily i.v. infusion) and piperacillin are active against these bacilli. These penicillins are usually given in combination with an aminoglycoside, e.g. gentamicin or netilmicin, for maximum benefit.

Treatment regimens may have to be modified in the light of sensitivity testing. Aminoglycosides are nephrotoxic and ototoxic, so blood levels should be monitored (p. 927).

Moraxella catarrhalis

This organism has been found to be associated with exacerbations of COPD and occasionally with fatal pneumonia. Some strains produce a β-lactamase capable of destroying amoxicillin. The organism is susceptible to co-trimoxazole and the quinolones.

Anaerobic bacteria

Infections with these organisms usually occur in patients with an underlying condition, such as diabetes, and are often associated with aspiration. *Bacteroides* is the most common organism and is sensitive to metronidazole. The prognosis depends largely on the precipitating cause.

Pneumonias due to opportunistic infections

Immunocompromised patients develop pneumonia with all the usual organisms and with a number of organisms which do not normally cause illness in healthy hosts. However, with HAART (p. 919) the incidence of these infections has fallen dramatically in AIDS patients.

Pneumocystis jiroveci

This is by far the most common opportunistic infection, accounting for 50% of the cases of pneumonia in patients with acquired immune deficiency syndrome (AIDS, see p. 200), particularly when the CD4 lymphocyte count is ≤200/mm^3. *Pneumocystis jiroveci* pneumonia is also seen in patients receiving immunosuppressive therapy and in malnourished children in the developing world. *Pneumocystis jiroveci* is found in the air, and pneumonia arises from reinfection rather than reactivation of persisting organisms acquired in childhood. Clinically the pneumonia is associated with a high fever, breathlessness and dry cough. The clinical features in patients with AIDS are described on page 187. The typical radiographic appearance is of a diffuse bilateral alveolar and interstitial shadowing beginning in the perihilar regions and spreading out in a butterfly pattern. Other chest X-ray appearances include localized infiltration, nodules, cavitation or a pneumothorax. In patients receiving aerosolized pentamidine for prophylaxis, in countries where HAART is not available, infiltrates may be localized to the upper zones. Empirical treatment is justified in very sick high-risk patients, but wherever possible a firm diagnosis should be obtained by stimulation of sputum with hypertonic saline or fibreoptic bronchoscopy with bronchoalveolar lavage; the diagnosis can be made in 90% of cases by staining sputum using indirect immunofluorescence with monoclonal antibodies.

Treatment of *Pneumocystis* pneumonia is with high-dose co-trimoxazole (see p. 200).

Other causes of shadowing on the chest X-ray in AIDS patients include:

- cytomegalovirus (p. 202)
- *M. avium-intracellulare* (p. 204)
- *M. tuberculosis* (p. 203)
- *L. pneumophila* (p. 860)
- *Cryptococcus* (p. 201)
- pyogenic bacteria
- Kaposi's sarcoma (p. 204)
- lymphoid interstitial pneumonia (p. 873)
- non-specific interstitial pneumonia (p. 873).

Diffuse pneumonia (bronchopneumonia)

Diffuse pneumonia is very common. It is differentiated from severe bronchitis by signs of bronchial breathing or patchy shadows on the chest X-ray. Widespread diffuse pneumonia is a common terminal event, occurring when patients dying from other conditions (e.g. cancer) are unable to cough up retained secretions, allowing infection to develop throughout the lungs. Decisions on therapy will vary according to the particular clinical circumstances but aggressive antibacterial treatment is rarely appropriate.

Aspiration pneumonia

The acute aspiration of gastric contents into the lungs can produce an extremely severe and sometimes fatal illness owing to the intense destructiveness of gastric acid. This can complicate anaesthesia, particularly during pregnancy, when it is termed Mendelson's syndrome.

In the absence of a tracheo-oesophageal fistula, aspiration occurs only during periods of impaired consciousness

(e.g. during sleep), in reflux oesophagitis with an oesophageal stricture, or in bulbar palsy. Because of the bronchial anatomy, the most usual sites for spillage are the apical and posterior segments of the right lower lobe. The persistent pneumonia is often due to anaerobes and may progress to lung abscess or even bronchiectasis. It is vital to identify any underlying problem, since without appropriate corrective measures aspiration will recur.

Treatment is discussed on page 863.

Rare causes of pneumonia

Pneumonia may occur as a minor feature during infection by *Bordetella pertussis*, typhoid and paratyphoid bacillus, brucellosis, leptospirosis and a number of viral infections including measles, chickenpox and glandular fever. Details of these infections are described in Chapter 2.

GENERAL MANAGEMENT OF PNEUMONIA

Refer to the algorithm given in Figure 14.36. Sputum and blood should always be sent for culture but antibiotic treatment should not be delayed. Severe cases need to be admitted to hospital and a chest X-ray performed. Other investigations, e.g. blood gases, are useful to detect respiratory failure and provide a baseline for comparison if the patient deteriorates. See Figure 14.36.

Further investigations may be necessary for the diagnosis of certain types of pneumonia:

- *Pneumococcal antigen* – counter-immunoelectrophoresis (CIE) of sputum, urine and serum (three to four times more sensitive than sputum or blood cultures).
- *Legionella antigen* – in urine, very specific. Immunofluorescence of organism on special media. Antibodies are less reliable.
- *Mycoplasma antibodies* (IgM and IgG) – in acute and convalescent samples – cold agglutinins present in 50%, complement fixing antibodies or PCR.
- *Chlamydia antibodies* – immunofluorescent complement fixation.

The choice of antibiotics is inevitably empirical, and is largely directed at *Strep. pneumoniae* infections. Apart from mycoplasma, the other pathogens are only responsible for a small minority of infections. Acquired antibiotic resistance of common respiratory bacterial pathogens is a recognized concern, but is still rare in the UK and rarely causes clinical failure. There is no convincing evidence that newer antibiotics provide any significant therapeutic advantage over established therapies.

Treatment of mild community-acquired pneumonia. Oral amoxicillin remains the preferred agent, but should be given at a dose of at least 500 mg 8-hourly. Oral erythromycin (or clarithromycin, which is better tolerated) is an alternative choice for those allergic to penicillin. For more severe cases treated in hospital, combined therapy with amoxicillin and a macrolide (erythromycin or clarithromycin) is recommended. When oral therapy is contraindicated, parenteral ampicillin or benzylpenicillin should be combined with clarithromycin. If *Staph. aureus* infection is suspected or is proven on culture, intravenous flucloxacillin ± sodium fusidate should be added. Fluoroquinolones are recommended for those intolerant of penicillins or macrolides. *For severe cases*, parenteral antibiotics should be given with the combination of a broad-spectrum lactamase-stable beta-lactam antibiotic (co-amoxiclav or cefuroxime) and clarithromycin. Parenteral antibiotics should be switched to oral once the temperature has settled for a period of 24 hours and provided there is no contraindication to oral therapy. The choice of antibiotics may be narrowed once microbiological results are available but do not forget that up to 10% of patients with pneumonia may have mixed infections.

The overall mortality for patients admitted to hospital with **community-acquired pneumonia** is currently 5%, except for *Staph. aureus* pneumonia where it exceeds 25%. Patients who die from pneumonia usually have not received the appropriate antibiotics in sufficient doses before or during the early stages of hospital admission.

Box 14.5 shows features of severe community-acquired pneumonia which indicate a poorer prognosis. These can be used to devise a scoring system (**CURB** – **C**onfusion, **U**rea > 7 mmol/L, **R**espiratory rate 30/min, **B**lood pressure (diastolic) < 60 mmHg). If all these factors are present and age is over 65 years, management in a high intensity care unit will be required. In spite of treatment in the intensive care unit, approximately 50% of such patients will die, partly as many have co-morbidities.

General measures

These include care of the mouth and skin. Fluids should be given to avoid dehydration. The patient is normally nursed sitting up or in the most comfortable position. Cough should normally be encouraged, but if it is unproductive and distressing, suppressants such as codeine linctus can be given. Physiotherapy is needed to help and encourage the patient to cough. Pleuritic pain may require analgesia, but powerful analgesics (e.g. opiates) should be used with care because they cause respiratory depression. In severe hypoxia, oxygen therapy should be given. However, since the hypoxia is often due to a physiological shunt, it makes little difference to the hypoxaemia.

Hospital-acquired (nosocomial) pneumonias

Patients with mild forms of hospital-acquired pneumonias

Box 14.5 CURB-65 Criteria for the diagnosis of severe community-acquired pneumonia

- **C**onfusion
- **U**rea >7 mmol/L
- **R**espiratory rate ≥30/min
- **B**lood pressure (systolic <90 or diastolic ≤60 mmHg
- Age >65 years of age

Score 0–1 – Treat as outpatient
Score 2 – Admit to hospital
Score 3+ – often require ICU care

Mortality rates increase with increasing score

Other markers of severe pneumonia

- Chest X-ray – more than one lobe involved
- P_aO_2 <8 kPa
- Low albumin (<35 g/L)
- White cell count (< 4×10^9/L or >20×10^9/L)
- Blood culture – positive

need to be reviewed regularly to make sure that there is nothing else responsible for their deterioration (e.g. heart failure, pulmonary embolism). Although very mild forms of hospital-acquired pneumonia may be treated with co-amoxiclav 500 mg three times daily, most patients will have co-morbidities which will dictate more aggressive antibiotic therapy. These should be managed in the same way as severe community-acquired pneumonias once appropriate samples for culture and sensitivities have been taken. Gram-negative bacteria are common and treatment should normally include a second-generation cephalosporin (e.g. cefuroxime) and aminoglycosides (e.g. gentamicin). Patients with chronic chest infection and others in whom *Pseudomonas* infection is suspected should receive i.v. ciprofloxacin or ceftazidime. Immunosuppressed patients may require very high-dose broad-spectrum antibiotics as well as antifungal and antiviral agents. Aspiration pneumonia (p. 861) is relatively common in hospital and usually involves infection with multiple bacteria, including anaerobes. The oral route is usually inappropriate in these patients. A combination of metronidazole (intravenous or rectal) and either co-amoxiclav or cefuroxime i.v. is recommended.

COMPLICATIONS OF PNEUMONIA

Lung abscess

This term is used to describe severe localized suppuration in the lung associated with cavity formation on the chest X-ray, often with the presence of a fluid level, and not due to tuberculosis.

There are many causes of lung abscess, but the most common is aspiration, particularly amongst heavy alcohol users following aspiration pneumonia. Lung abscesses also frequently follow the inhalation of a foreign body into a bronchus and occasionally occur when the bronchus is obstructed by a bronchial carcinoma. Chronic or subacute lung abscesses follow an inadequately treated pneumonia.

Abscesses also develop during the course of specific pneumonias, particularly when the infecting agent is *Staph. aureus* or *Klebsiella pneumoniae*. Septic emboli, usually staphylococci, result in multiple lung abscesses. Infarcted areas of lung occasionally cavitate and rarely become infected. Amoebic abscesses occasionally develop in the right lower lobe following transdiaphragmatic spread from an amoebic liver abscess.

The clinical features are persisting and worsening pneumonia associated with the production of large quantities of sputum, which is often foul-smelling owing to the growth of anaerobic organisms. There is usually a swinging fever; malaise and weight loss occur. The chest signs may be few but clubbing often develops if the condition is not rapidly cured. The patient is often anaemic with a high ESR and CRP.

Empyema

Empyema means the presence of pus within the pleural cavity. This usually arises from bacterial spread from a severe pneumonia or after the rupture of a lung abscess into the pleural space. Typically an empyema cavity becomes infected with anaerobic organisms and the patient is severely ill with a high fever and a neutrophil granulocytosis.

Investigations

Bacteriological investigation of lung abscess and empyema is best conducted on specimens obtained by transtracheal aspiration, bronchoscopy or percutaneous transthoracic aspiration with ultrasound or CT guidance. Bronchoscopy is helpful to exclude carcinomas and foreign bodies.

Treatment

Although anaerobic organisms are found in up to 70% of lung abscesses and empyemas, there is usually a mixed flora, often with aerobes, particularly *Strep. milleri*. Anaerobic cocci, black-pigmented bacteroids and fusobacteria are the anaerobes found most commonly.

Appropriate antibiotic treatment is given for up to 6 weeks. Antibiotics should be given to cover both aerobic and anaerobic organisms. An appropriate initial choice is cefuroxime 1 g i.v. 6-hourly and metronidazole 500 mg i.v. 8-hourly for 5 days, followed by oral cefaclor and metronidazole for a prolonged period depending on bacterial sensitivities. Abscesses occasionally require surgery.

In addition to antibiotics, empyemas require prompt insertion of an intercostal drain under ultrasound or CT guidance. If drainage is inadequate, surgery to clear the cavity and insertion of a large-bore tube is required.

TUBERCULOSIS (see also p. 145)

Following many decades of decline, tuberculosis is on the increase in developed countries, because of AIDS, the use of immunosuppressive drugs which depress the host defence mechanisms, decreased socio-economic conditions, as well as increased immigration of persons from areas of high endemicity. In developing countries it is 20–50 times more common. In 2004 there were 9 million new cases world-wide. The major reason for this increase was because of the rapid rise in cases from sub-Saharan Africa (due to AIDS) and Russia.

Epidemiology

Tuberculosis is the world's leading cause of death from a single infectious disease, with 1.7 million deaths (without HIV infection) in 2004. This is the result of:

- inadequate programmes for disease control with poorly supervised treatment
- multiple drug resistance (MDR), extensive drug resistant TB (XDR-TB)
- co-infection with HIV
- a rapid rise in the world's population of young adults – the age group with the highest mortality from tuberculosis
- overcrowding and poor nutrition.

Tuberculosis is a notifiable disease in the UK; the number of cases is now fairly stable, with 8037 notifications in 2005, compared to just over 5000 ten years ago. In India the estimated incidence in 2004 was approximately 1.8 million. The incidence of tuberculosis in the UK in immigrants from the Asian subcontinent and from the West Indies is respectively 40 and four times as common as in the native white population. This leads to great variation in the frequency of the disease in different areas of the UK.

Pathology

The first infection with *M. tuberculosis* is known as primary tuberculosis. It is usually subpleural, often in the mid to upper

Pneumonia
General management of pneumonia
Complications of pneumonia
Tuberculosis

FURTHER READING

Garau J, Calbo E. Community acquired pneumonia. *Lancet* 2008; **371**: 455–458.

Weinstein RA. Planning for epidemics – the lessons of SARS. *New England Journal of Medicine* 2004; **350**: 2332–2334.

zones (Ghon focus). Within an hour of reaching the lung, tubercle bacilli reach the draining lymph nodes at the hilum of the lung and a few escape into the bloodstream.

In the initial reaction the alveolar macrophages ingest the bacillus. The bacteria proliferate inside the cells and the macrophages release chemokines and cytokines which attract neutrophil granulocytes, monocytes and other inflammatory cells. The macrophages present the antigen to the T lymphocytes with the development of cellular immunity that can be demonstrated 3–8 weeks after the initial infection by a positive reaction in the skin to an intradermal injection of protein from tubercle bacilli (tuberculin/PPD). A delayed hypersensitivity-type reaction occurs, resulting in tissue necrosis, and at this stage the classical pathology of tuberculosis can be seen. Granulomatous lesions consist of a central area of necrotic material of a cheesy nature, called caseation, surrounded by epithelioid cells and Langhans' giant cells with multiple nuclei, both cells being derived from the macrophage. Lymphocytes are present and there is a varying degree of fibrosis. Subsequently the caseated areas heal completely and many become calcified. It is known that at least 20% of these calcified primary lesions contain tubercle bacilli, initially lying dormant but capable of being activated following depression of the host defence system (latent TB). Reactivation leads to typical post-primary pulmonary tuberculosis with cavitation, usually in the apex or upper zone of the lung. 'Post-primary tuberculosis' refers to all forms of tuberculosis that occur after the first few weeks of the primary infection when immunity to the mycobacterium has developed.

Clinical features and investigations

Primary tuberculosis is symptomless in the great majority of individuals. Occasionally there may be a vague illness, sometimes associated with cough and wheeze. A small transient pleural effusion or erythema nodosum may occur, both representing hypersensitivity manifestations of the infective process.

Enlargement of lymph nodes compressing the bronchi can give rise to collapse of segments or lobes of the lung. Apart from cough and a monophonic wheeze, the individual remains remarkably well and the collapse disappears as the primary complex heals. Occasionally, persistent collapse can give rise to subsequent bronchiectasis, often in the middle lobe (Brock's syndrome).

- The manifestations of *primary and post-primary tuberculosis* are shown in Figure 14.37, together with the times when they usually occur.
- Extrapulmonary manifestations are summarized on page 145.
- Miliary tuberculosis can occur within a year of the primary infection, or can occasionally occur much later as a manifestation of reactivation or, rarely, reinfection with tubercle bacillus.

Reactivation in the lung, or indeed in any extrapulmonary location, can occur as immunity wanes, usually with age or chronic ill-health. All manifestations are shown in Figure 14.37.

Miliary tuberculosis

This disease is the result of acute diffuse dissemination of tubercle bacilli via the bloodstream (haematogenous spread). It can be a difficult diagnosis to make, especially in older people, where it is particularly covert. It can lead on from a primary infection or reactivation of a latent focus due to immunosuppression for any reason, e.g. drugs, co-

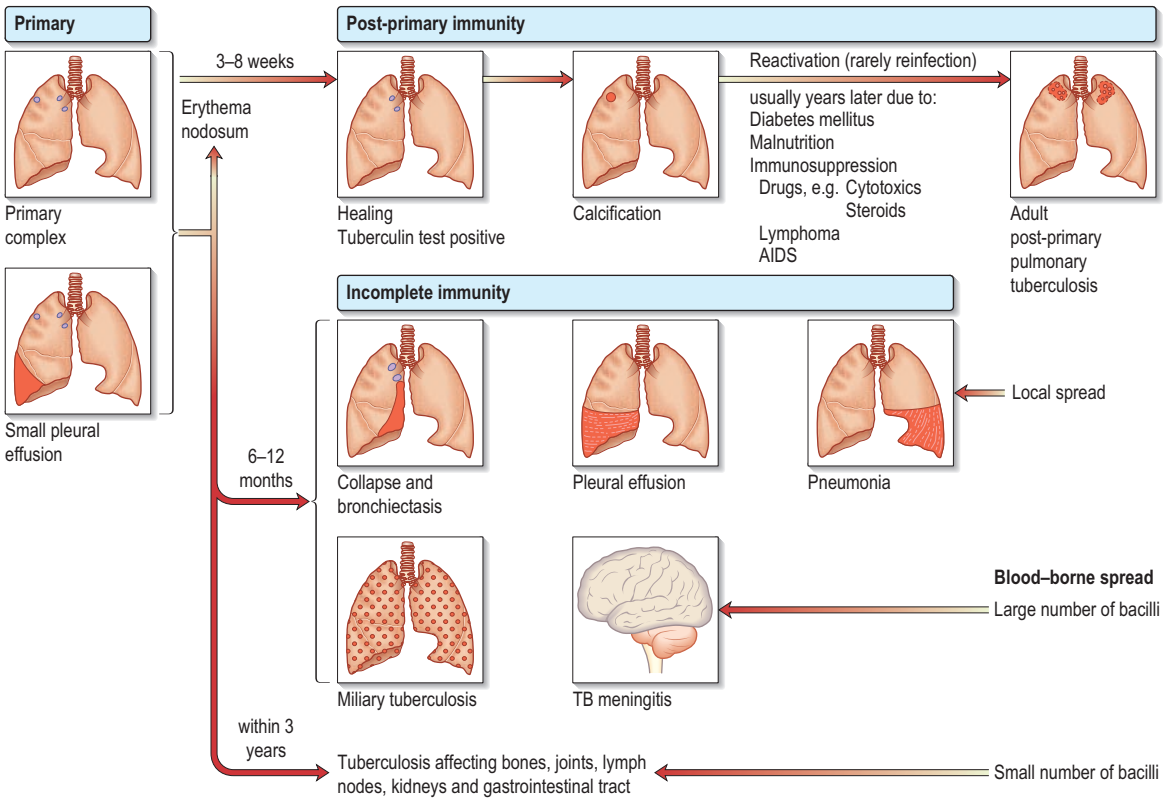


Fig. 14.37 **Manifestations of primary and post-primary tuberculosis.**

morbidity. This form of disseminated tuberculosis is universally fatal without treatment.

It may present in an entirely non-specific manner with the gradual onset of vague ill-health, loss of weight and then fever. Occasionally the disease presents as tuberculous meningitis. Usually there are no abnormal physical signs in the early stages, although eventually the spleen and liver become enlarged. Choroidal tubercles are seen in the eyes. These lesions are about one-quarter of the diameter of the optic disc and are yellowish and slightly shiny and raised in nature, later becoming white in the centre. There may be one or many in each eye.

Investigation

- *The chest X-ray* may be entirely normal in miliary tuberculosis as the tubercles are not visible until uniform miliary shadows 1–2 mm in diameter are seen throughout the lung; they have a hard outline. The lesions can increase in size up to 5–10 mm. Sarcoidosis and staphylococcal or mycoplasma pneumonia can mimic the chest X-ray appearance of miliary tuberculosis.
- *CT scanning* may reveal lung parenchymal abnormalities at an earlier stage.
- *The Mantoux test* is usually positive but may be negative in 30–50% of people with very severe disease.
- *Transbronchial biopsies* are frequently positive before any abnormality is visible on the chest X-ray.
- *Biopsy and culture of liver and bone marrow* may be necessary in patients presenting with a pyrexia of unknown origin (PUO). A trial of antituberculous therapy can be used in individuals with a PUO. The fever should settle within 2 weeks of starting chemotherapy if it is due to tuberculosis. This approach is used in susceptible individuals when a diagnosis cannot be confirmed by other means.

Adult post-primary pulmonary tuberculosis

Typically there is gradual onset of symptoms over weeks or months. Tiredness, malaise, anorexia and loss of weight together with a fever and cough remain the outstanding features of pulmonary tuberculosis. Drenching night sweats are rather uncommon and are more usually due to anxiety. Sputum in tuberculosis may be mucoid, purulent or blood-stained. Many patients suffer a dull ache in the chest and it is not uncommon for patients to complain of recurrent colds. A pleural effusion or pneumonia can be the presenting feature of tuberculosis.

Physical examination reveals little. Finger clubbing is only present if the disease is advanced and associated with considerable production of purulent sputum. There are often no physical signs in the chest even in the presence of extensive radiological changes, though occasionally persistent crackles may be heard. Physical signs of an associated effusion, pneumonia or fibrosis may be present.

Investigations

An abnormal chest X-ray is often found with no symptoms, but the reverse is extremely rare – pulmonary tuberculosis is unlikely in the absence of any radiographic abnormality.

The chest X-ray (Fig. 14.38) typically shows patchy or nodular shadows in the upper zones, loss of volume, and fibrosis with or without cavitation. Calcification may be present. The X-ray appearances alone can strongly suggest tuberculosis, but every effort must be made to obtain microbiological evidence. A single X-ray does not give an indication of the activity of the disease. Very similar chest X-ray appearances occur in histoplasmosis and other fungal infections of the lung, including cryptococcosis, coccidioidomycosis, blastomycosis and aspergillosis, as well as in bronchial carcinoma or cavitating pulmonary infarcts.

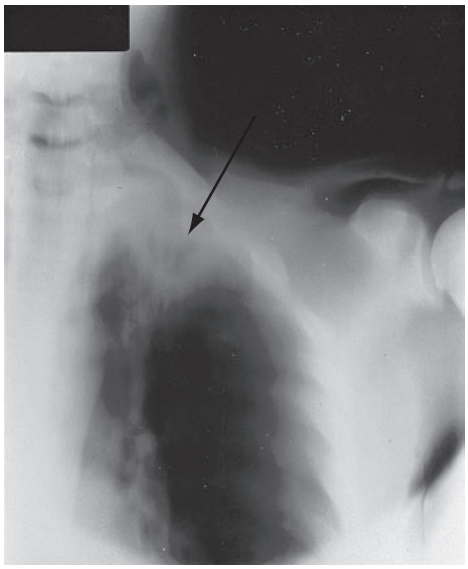


Fig. 14.38 **Chest X-ray showing tuberculosis of left upper lobe with cavitation.**

Lymph node presentation of tuberculosis

The patient presents with a tender lump or fluctuant mass, usually supraclavicular or in the anterior triangle of the neck.

This form of tuberculosis is discussed on page 145.

Tuberculosis in HIV-infected persons

(see also p. 203)

Tuberculosis in an HIV-infected person is an AIDS-defining illness. Disease may arise from rapid progression of primary infection, by reactivation and by reinfection. The clinical pattern of disease is described on page 203. Treatment is with conventional therapy, but with four rather than three drugs, and it should be supervised (directly observed therapy short course, DOTS). Adverse drug reactions are common and the prognosis may be poor, especially if treatment is not supervised. Multiple drug resistance (MDR) occurs worldwide in about 20%, with an incidence of 2% for extensive drug resistance (XDR) (from 37 countries). Initially all patients with XDR were thought to be HIV positive but now nosocomial transmission is thought to play a significant role (p. 866).

Diagnosis

The diagnosis of tuberculosis is made on the basis of the following investigations:

- **Imaging**. Chest X-ray (see above), and CT scan if necessary.
- **Staining**. The sputum is stained with an auramine–phenol fluorescent test. The Ziehl–Neelsen (ZN) stain for acid and alcohol-fast bacilli (AAFB) is less sensitive.

- **Culture**. The sputum is cultured on Lowenstein–Jensen or Middlebrook media for 4–8 weeks. Liquid culture (Bactec (Becton–Dickinson)) is used in many laboratories and has the advantage of shorter culture times. Cultures to determine the sensitivity of the bacillus to antibiotics take a further 3–4 weeks. More rapid culture and sensitivity tests are being developed, e.g. microscopic observation drug susceptibility (MODS).
- **Fibreoptic bronchoscopy** with washings from the affected lobes is useful if no sputum is available. This has replaced former techniques such as gastric washings. Transbronchial biopsies can also be obtained for histology and microbiological assessment.
- **Biopsies** of the pleura, lymph nodes and solid lesions within the lung (tuberculomas) may be required to confirm the diagnosis.
- **Whole blood interferon-γ assay or skin tests** (see p. 867).

The slow growth of *M. tuberculosis* in culture has hindered the ability to make a rapid definitive diagnosis. Radiolabelled DNA and RNA probes specific for various mycobacterial species can identify organisms in culture. The sensitivity of these methods has been enhanced by amplifying target DNA using the polymerase chain reaction. This allows direct testing of sputum and other fluids to provide a laboratory diagnosis within 48 hours and may provide rapid information on the presence or absence of rifampicin resistance. This is still not entirely reliable and should not be accepted as a final diagnosis, particularly in the difficult case when it is most likely to be used.

Treatment

Bed rest does not affect the outcome of the disease. Some patients will require hospitalization for a brief period; these include ill patients, smear-positive, highly infectious patients (particularly with multidrug-resistant TB), those in whom the diagnosis is uncertain and those individuals from whom it is essential to gain cooperation. The successful treatment of tuberculosis is the continuous self-administration of drugs for 6 months; lack of patient compliance is a major reason why 5% of patients do not respond to treatment. In vitro resistance to one or more of the antituberculous drugs used to occur in fewer than 1% of patients in the UK, but the rate of resistance is gradually increasing (isoniazid resistance 4–6%, multidrug resistance 1%).

Directly observed therapy short course (DOTS)

In order to improve compliance, special clinics are used to supervise treatment regimens directly. Incentives to attend (e.g. free meals, cash payments) may be helpful. The effectiveness of DOTS is, however, variable in different countries and compliance is still a problem. Long-stay hospital treatment is required only for persistently uncooperative patients, many of whom are homeless and abuse alcohol.

Six-month regimen

Six months of treatment with once-daily rifampicin 600 mg and isoniazid 300 mg is standard practice for patients with pulmonary and lymph node disease. (For those whose bodyweight is below 55 kg, rifampicin is reduced to 450 mg daily.) These are given as combination tablets and are taken 30 minutes before breakfast, since the absorption of rifampicin is influenced by food. This is supplemented for the first 2 months by pyrazinamide at a dose of 1.5 g (bodyweight <50 kg) or 2.0 mg daily. Pyrazinamide is of particular value in treating mycobacteria present within macrophages, and for this reason it may have a very valuable effect on preventing subsequent relapse. Pyridoxine 10 mg daily is given to reduce the risk of isoniazid-induced neuropathy. Current recommendations advise the addition of ethambutol (using 15 mg per day) if the risk of drug resistance is increased.

Longer regimens

Treatment of bone tuberculosis should be continued for a total of 9 months, and of tuberculous meningitis for 1 year. The drugs used are the same as for pulmonary tuberculosis, with pyrazinamide prescribed for the first 2 months only and for drug-resistant disease.

Drug-resistant organisms

The development of resistance after initial drug sensitivity (*secondary drug resistance*) occurs in patients who do not comply with the treatment regimens. This is seen world-wide and, in the UK, in the immigrant population. *Primary drug resistance* is seen in immigrants to the UK and those exposed to others infected with resistant organisms.

The WHO defines two categories of drug resistance:

- multidrug resistance (MDR) – isoniazid, rifampicin
- extensive drug resistance (XDR) – isoniazid, rifampicin, quinolone and at least one of the following second-line drugs: kanamycin, capteomycin, amikacin.

MDR and XDR is a world-wide major therapeutic problem with a high mortality and occurs mainly in HIV-infected patients. Nosocomial transmission of multidrug-resistant tuberculosis to healthcare workers and to other patients is recognized and poses a major public health problem. The drug treatment of suspected drug resistance in both HIV-positive and HIV-negative patients is as follows:

- with multiple drug resistance use at least three drugs to which the organism is sensitive
- with resistance to one of the four main drugs, use the other three.

Therapy should be continued for up to 2 years and in HIV-positive patients for at least 12 months after negative cultures. Second-line drugs available for treatment of resistant *M. tuberculosis* are capreomycin, cycloserine, clarithromycin, azithromycin, ciprofloxacin, ofloxacin, ethionamide, kanamycin, amikacin, moxifloxacin and rifabutin.

Unwanted effects of drug treatment

Rifampicin induces liver enzymes, which may be transiently elevated in the serum of many patients. The drug should be stopped only if the serum bilirubin becomes elevated or if transferases are >3× elevated, which is uncommon. Thrombocytopenia has been reported. Rifampicin stains body secretions pink and the patients should be warned of the change in colour of their urine, tears (contact lens) and sweat. Induction of liver enzymes means that concomitant drug treatment may be made less effective (see Ch. 16). Oral contraception will not be effective, so alternative birth-control methods should be used. Rifabutin, a new rifamycin, is similar and is used for prophylaxis against *M. avium-intracel-*

Iulare complex infection in HIV patients with CD4 counts $<200\ mm^3$.

Isoniazid has very few unwanted effects. At high doses it may produce a polyneuropathy but this is extremely rare when the normal dose of 200–300 mg is given daily. Nevertheless, it is customary to prescribe pyridoxine 10 mg daily to prevent this effect. Occasionally, isoniazid gives rise to allergic reactions in the form of a skin rash and fever, with hepatitis occurring in fewer than 1% of cases. The latter, however, may be fatal if the drug is continued.

Pyrazinamide may cause hepatic toxicity, though this is much rarer with present dosage schedules. Pyrazinamide reduces the renal excretion of urate and may precipitate hyperuricaemic gout.

Ethambutol can cause a dose-related optic retrobulbar neuritis that presents with colour blindness for green, reduction in visual acuity and a central scotoma (commoner at doses of 25 mg/kg). This usually reverses provided the drug is stopped when symptoms develop; patients should therefore be warned of its effects. All patients prescribed the drug should be seen by an ophthalmologist prior to treatment and doses of 15 mg/kg should be used.

Streptomycin can cause irreversible damage to the vestibular nerve. It is more likely to occur in the elderly and in those with renal impairment. Allergic reactions to streptomycin are more common than to rifampicin, isoniazid and pyrazinamide. This drug is used only if patients are very ill, have multidrug-resistant TB or are not responding adequately to therapy.

Follow-up

Patients should be seen regularly for the duration of chemotherapy and once more after 3 months, since relapse, though very unlikely, usually occurs within this period of time. Patients with multidrug-resistant TB should be followed up for at least 1 year after treatment is completed.

Chemoprophylaxis

Patients who have any chest X-ray changes compatible with previous tuberculosis and who are about to undergo long-term treatment that has an immunosuppressive effect, such as renal dialysis or treatment with corticosteroids, should receive chemoprophylaxis with isoniazid 300–450 mg daily.

Prevention

BCG vaccination

BCG is a live attenuated vaccine derived from *M. bovis* (a bovine strain of *Mycobacteria*) that lost its virulence after growth in the laboratory for many passages. Early trials showed that it decreases the risk of developing tuberculosis by about 70%. A recent study has shown that world-wide use of BCG is very cost-effective. However, in the UK (with a low incidence of TB) it is less cost-effective to administer this vaccine, and the procedure is being stopped in many areas of the UK. However, in other areas of the UK with a high immigrant population, the vaccine is being administered at birth rather than at the traditional age of 13 years. This is to prevent the disease from developing in young children, where it can progress extremely rapidly and in whom any delay in diagnosis can be fatal. BCG has been shown to be particularly effective in preventing miliary tuberculosis and tuberculous meningitis.

In meta-analysis the overall protective efficacy is around 50%. This efficacy varies throughout the world from zero to 94% protection and appears to depend on latitude, being most beneficial in Norway, Sweden and Denmark (80–94%) and least so in the southern states of the USA and in India (0–20%). This lack of efficacy is thought to be related to a number of local factors including the frequency of infection with environmental mycobacteria (e.g. *M. fortuitum, M. kansasii*), which may induce a degree of protection similar to but not enhanced by BCG.

Contact tracing

Tuberculosis is spread from person to person, and effective tracing of close contacts has helped to limit spread of the disease as well as to identify diseased individuals at an early stage. Screening procedures involve screening all close family members or other individuals who share the same kitchen and bathroom facilities. Close contacts at work or school may also be screened. Contacts who are ill should be thoroughly investigated for tuberculosis. If they are well, a chest X-ray is taken and a tuberculosis test is performed (Practical box 14.5). The Mantoux test is the standard tuberculin skin test but this requires the person to be seen again 72 hours after the intradermal injection. Whole blood IFN-γ assays are being increasingly used. These have the advantage of just a single visit as it is an in vitro test. Early secretory antigenic target-6 (ESAT-6), a specific myobacterium antigen,

Practical Box 14.5 Tuberculosis testing

Mainly used for:

- Contact tracing
- BCG vaccination programmes
- Diagnosis of tuberculosis in developing countries

The Mantoux test

Patients are tested with purified protein derivative (PPD) of *Mycobacterium tuberculosis*.

The test is based on cell-mediated immunity with the development of induration and inflammation at the site of infection due to infiltration with mainly T lymphocytes. The test can be falsely negative in patients with AIDS because of impairment of delayed hypersensitivity.

- 0.1 mL of a 1:1000 strength PPD (equivalent to 10 tuberculin units) is injected intradermally.
- The induration (not the erythema) is measured after 72 hours. The test is positive if the induration is 10 mm or more in diameter.

Whole blood interferon-gamma assay (Quantiferon-TB gold and T-Spot-TB)

Can be used in all circumstances.

- In vitro T cell based assay
- T cells of individuals who have previously been sensitized to TB will produce IFN-γ when they are incubated with TB antigens (e.g. PPD or early secretory antigen target-6 (ESAT-6))
- ELISA test or an enzyme-linked monospot assay
- Advantages
  No return visit to look at test result
  Negative in BCG immunized people
- Disadvantages
  Variable sensitivity and specificity

is not present in BCG and thus the test is negative in patients previously immunized with BCG – unlike the skin test which is positive.

In adults, even if a tuberculosis test is positive, provided the chest X-ray is negative nothing more need be done. Chemoprophylaxis should be given in patients whose recent test conversion has been documented and for young adults (16–34 years) who are positive on tuberculosis testing without BCG history and found at new immigrant screening. In patients with HIV infection, who have not had BCG, chemoprophylaxis with isoniazid is given, reducing the relative risk of developing active TB by 40% in highly endemic areas.

In children, a positive test is usually taken as evidence of infection, and treatment is instituted. If the test is negative in children and young adults (<35 years), it is repeated at 6 weeks, and if it remains negative then BCG is administered. Children under the age of 1 year who have a family member with tuberculosis are given chemoprophylaxis with a daily dose of isoniazid 5–10 mg/kg for 6 months together with immunization with a strain of BCG that is resistant to isoniazid.

In general, in the UK much greater emphasis is placed on contact tracing and investigation of those under the age of 35 years and in some immigrant groups (African, Asian and Eastern European) in whom the disease is more prevalent.

Other mycobacteria

M. kansasii occurs in water and milk, though not in soil. Disease caused by this mycobacterium has mainly been described in Europe and the USA. It rarely causes a relatively benign type of human pulmonary disease, more common in HIV-positive individuals, usually in middle-aged males. Men working in dusty jobs (e.g. miners) appear to be especially at risk, as are those who have underlying COPD.

M. avium-intracellulare is a cause of pulmonary infection in AIDS patients (see p. 204).

DIFFUSE DISEASES OF THE LUNG PARENCHYMA

Diffuse parenchymal lung disorders (DPLD, also referred to as interstitial lung disease) are a heterogeneous group of disorders accounting for about 15% of respiratory clinical practice. There is diffuse lung injury and inflammation that can progress to lung fibrosis. The classification is shown in Box 14.6.

GRANULOMATOUS LUNG DISEASE

A granuloma is a mass or nodule composed of chronically inflamed tissue formed by the response of the mononuclear phagocyte system (macrophage/histiocyte) to an insoluble or slowly soluble antigen or irritant. If the foreign substance is inert (e.g. an inhaled dust), the phagocytes turn over slowly; if the substance is toxic or reproducing, the cells turn over faster, producing a granuloma. A granuloma is characterized by epithelioid multinucleate giant cells, as seen in tuberculosis. Granulomas are also seen in other infections, including fungal and helminthic, in sarcoidosis, and in extrinsic allergic alveolitis, and can also be due to foreign bodies (e.g. talc).

Box 14.6 Diffuse parenchymal lung diseases (DPLDs)

- Granulomatous lung disease, e.g. sarcoid
- Granulomatous lung disease with vasculitis, e.g. Wegener's, Churg–Strauss, microscopic vasculitis
- Pulmonary autoimmune rheumatic diseases, e.g. rheumatoid arthritis, systemic lupus erythematosus
- Idiopathic interstitial pneumonias – see Box 14.7
- Drugs – see Table 14.17
- Other forms
  - Langerhans' cell histiocytosis
  - Goodpasture's syndrome
  - Diffuse alveolar haemorrhage
  - Idiopathic pulmonary haemosiderosis
  - Lymphangioleiomyomatosis
  - Pulmonary alveolar proteinosis

Granulomatous lung disease with pulmonary vasculitis is discussed on page 870.

Sarcoidosis

Sarcoidosis is a multisystem granulomatous disorder, commonly affecting young adults and usually presenting with bilateral hilar lymphadenopathy, pulmonary infiltration and skin or eye lesions. Beryllium poisoning can produce a clinical and histological picture identical to sarcoidosis, though contact with this element is now strictly controlled.

Epidemiology and aetiology

Sarcoidosis is a common disease of unknown aetiology that is often detected by routine chest X-ray. There is great geographical variation. The prevalence in the UK is approximately 19 in 100 000 of the population. Sarcoidosis is common in the USA but is uncommon in Japan. The course of the disease is much more severe in American blacks than in whites. There is no relation with any histocompatibility antigen, but first degree relatives (particularly in Caucasians) have an increased risk of developing sarcoidosis. Other aetiological factors suggested are an atypical mycobacterium or fungus, the Epstein–Barr virus, and occupational, genetic, social or other environmental factors (there is a higher incidence in rural than in urban populations). None of these has been substantiated.

Immunopathology

- Typical sarcoid granulomas consist of focal accumulations of epithelioid cells, macrophages and lymphocytes, mainly T cells.
- There is depressed cell-mediated reactivity to tuberculin and other antigens such as *Candida albicans*.
- There is overall lymphopenia: circulating T lymphocytes are low but B cells are slightly increased.
- Bronchoalveolar lavage shows a great increase in the number of cells; lymphocytes are greatly increased (particularly CD4 helper cells).
- The number of alveolar macrophages is increased but they represent a reduced percentage of the total number of bronchoalveolar lavage cells.
- Transbronchial biopsies show infiltration of the alveolar walls and interstitial spaces with leucocytes, mainly T cells, prior to granuloma formation.

FURTHER READING

Basu S, Andrews JR, Poolman EM et al. Prevention of nosocomial transmission of extensively drug-resistant tuberculosis in rural South African district hospitals. *Lancet* 2007; **370**: 1500–1507.

British Thoracic Society. Recommendations for assessing risk and managing M. tuberculosis infection and disease in patients due to start anti-TNF-alpha treatment. *Thorax* 2005; **60**: 800–805.

Lalvani A. Diagnosing tuberculosis in the 21st century: new tools to tackle an old enemy. *Chest* 2007; **131**: 1898–1906.

Lancet 2006; **367**: 875–958. A series of articles on tuberculosis.

Maartens G, Wilkinson RJ. Tuberculosis. *Lancet* 2007; **370**: 2030–2043.

United Nations Millennium Development Goals report, 2005. http://www.un.org/millenniumgoals/

It seems likely that the decrease in circulating T lymphocytes and changes in delayed hypersensitivity responses are the result of sequestration of lymphocytes within the lung. There is no evidence to suggest that patients with sarcoidosis suffer from an overall defect in immunity, since the frequency of fungal, viral and bacterial infections is not increased and there is no substantiated evidence of a greater risk of developing malignant neoplasms.

Clinical features

The peak incidence is in the third and fourth decades, with a female preponderance. Sarcoidosis can affect many different organs of the body. The most common presentation is with respiratory symptoms or abnormalities found on chest X-ray (50%). Fatigue or weight loss occurs in 5%, peripheral lymphadenopathy in 5% and a fever in 4%. Neurological presentations are rare but well recognized and can mimic a variety of conditions. A chest X-ray may be negative in up to 20% of non-respiratory cases, though pulmonary lesions may be detected later.

Bilateral hilar lymphadenopathy

This is a characteristic feature of sarcoidosis. It is often symptomless and simply detected on a routine chest X-ray. Occasionally, the bilateral hilar lymphadenopathy is associated with a dull ache in the chest, malaise and a mild fever.

Although there may be no evidence of infiltration in the lung fields on plain chest X-ray, evidence from CT scanning (Fig. 14.39), transbronchial biopsies and bronchoalveolar lavage indicates that the lung parenchyma is nearly always involved.

The differential diagnosis of bilateral hilar lymphadenopathy includes:

- lymphoma – though it is rare for this to affect only the hilar lymph nodes
- pulmonary tuberculosis – though it is rare for the hilar lymph nodes to be symmetrically enlarged
- carcinoma of the bronchus with malignant spread to the contralateral hilar lymph nodes – again it is rare for this to give rise to the typical symmetrical picture.

Pulmonary infiltration

This type of sarcoidosis may be progressive leading to increasing effort dyspnoea and eventually cor pulmonale and death. The chest X-ray shows mottling in the mid-zones proceeding over time to generalized fine nodular shadows. Eventually, widespread pulmonary line shadows develop, reflecting the underlying fibrosis. A honeycomb appearance can occasionally occur. Pulmonary function tests show a typical restrictive lung defect (see below).

The combination of pulmonary infiltration and normal lung function tests is highly suggestive of sarcoidosis. The principal differential diagnoses are tuberculosis, pneumoconiosis, idiopathic pulmonary fibrosis and alveolar cell carcinoma.

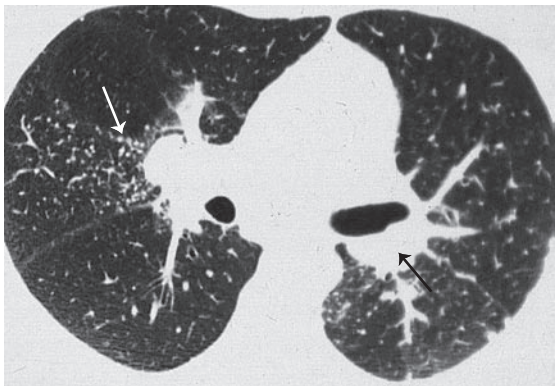


Fig. 14.39 **CT scan in sarcoidosis.** Note enlarged glands at the hilum (black arrow) and nodular shadowing, particularly in right middle lobe (white arrow).

Extrapulmonary manifestations

Skin and ocular sarcoidosis are the most common extrapulmonary presentations.

Skin lesions occur in 10% of cases. Sarcoidosis is the most common cause of erythema nodosum (see p. 1250). The association of bilateral symmetrical hilar lymphadenopathy with erythema nodosum occurs only in sarcoidosis. A chilblain-like lesion known as lupus pernio is also seen, as are skin nodules (see p. 1253).

Ocular and associated effects. Anterior uveitis is common and may present with misting of vision, pain and a red eye, but posterior uveitis may present simply as progressive loss of vision. Although ocular sarcoidosis accounts for about 5% of uveitis presenting to ophthalmologists, evidence of asymptomatic uveitis may be found in up to 25% of patients with sarcoidosis. Conjunctivitis may occur and retinal lesions have also been reported. *Uveoparotid fever* is a syndrome of bilateral uveitis and parotid gland enlargement together with occasional development of facial nerve palsy and is sometimes seen with sarcoidosis.

Keratoconjunctivitis sicca and lacrimal gland enlargement may also occur.

Metabolic manifestations. It is rare for sarcoidosis to present with problems of calcium metabolism, though hypercalcaemia is found in 10% of established cases. Hypercalcaemia and hypercalciuria can lead to the development of renal calculi and nephrocalcinosis. The cause of the hypercalcaemia is due to an increase in circulating 1,25-dihydroxyvitamin D_3, with 1α-hydroxylation occurring in sarcoid macrophages in the lung in addition to that taking place in the kidney.

The central nervous system. Involvement of the central nervous system (CNS) is rare (2%) but can lead to severe neurological disease (see p. 1159).

Bone and joint involvement. Arthralgia without erythema nodosum is seen in 5% of cases. Bone cysts are found, particularly in the digits, with associated swelling. In the absence of swelling, routine X-rays of the hands are unnecessary.

Hepatosplenomegaly. Sarcoidosis is a cause of hepatosplenomegaly, though it is rarely of any clinical consequence. A liver biopsy is occasionally performed when the diagnosis is in doubt and shows granulomas.

Cardiac involvement is rare (3%). Ventricular dysrhythmias, conduction defects and cardiomyopathy with congestive cardiac failure are seen.

Investigations

- **Imaging.** Chest X-ray (see above). High resolution CT is useful for assessment of diffuse lung parenchymal involvement, and multislice CT detects and characterizes small nodules.
- **Full blood count.** There may be a mild normochromic, normocytic anaemia with raised ESR or CRP.
- **Serum biochemistry.** Serum calcium is often raised and there is hypergammaglobulinaemia.
- **Transbronchial biopsy** is the most useful investigation, with positive results in 90% of cases of pulmonary sarcoidosis with or without X-ray evidence of lung parenchymal involvement. Pulmonary non-caseating granulomas are found in approximately 50% of patients with extrapulmonary sarcoidosis in whom the chest X-ray is normal.
- **Serum level of angiotensin-converting enzyme (ACE)** is raised by two standard deviations above the normal mean value in over 75% of patients with untreated sarcoidosis. Raised (but lower) levels are also seen in patients with lymphoma, pulmonary tuberculosis, asbestosis and silicosis, limiting the diagnostic value of the test. However, the test is useful in assessing the activity of the disease and therefore as a guide to treatment with corticosteroids. Reduction of serum ACE during treatment with corticosteroids does not, however, imply complete resolution of the disease.
- **Lung function tests** show a restrictive lung defect in patients with pulmonary infiltration. There is a decrease in TLC, a decrease in both FEV_1 and FVC, and a decrease in gas transfer. Lung function is usually normal in patients who present with extrapulmonary disease or who only have hilar adenopathy on chest X-ray.

Treatment

Both the need to treat and the value of corticosteroid therapy are contested in many aspects of this disease. Hilar lymphadenopathy on its own with no evidence on chest X-ray of involvement of the lungs or decrease in lung function tests does not require treatment. Persisting infiltration visible on the chest X-ray or abnormal lung function tests are unlikely to improve without corticosteroid treatment. If the disease is not improving spontaneously 6 months after diagnosis, treatment should be started with prednisolone 30 mg for 6 weeks, reducing to alternate-day treatment with prednisolone 15 mg for 6–12 months. Although there have been no controlled trials of corticosteroids, it is difficult to withhold them when there is continuing deterioration of lung function. Eye involvement or persistent hypercalcaemia are definite indications for systemic steroids.

If the erythema nodosum of sarcoidosis is severe or persistent it will respond rapidly to a 2-week course of prednisolone 5–15 mg daily, as will patients with uveoparotid fever. Myocardial sarcoidosis and neurological manifestations are also treated with prednisolone.

Prognosis

Sarcoidosis is a much more severe disease in certain racial groups, particularly American blacks, where death rates of up to 10% have been recorded. It is probable that the disease is fatal in fewer than 5% of cases in the UK, most often as a result of respiratory failure and cor pulmonale and rarely, from myocardial sarcoidosis and renal damage. The chest X-ray provides a guide to prognosis. The disease remits within 2 years in over two-thirds of patients with hilar lymphadenopathy alone, in approximately one-half with hilar lymphadenopathy plus chest X-ray evidence of pulmonary infiltration, but in only one-third of patients with X-ray evidence of infiltration without any demonstrable lymphadenopathy. Lung function tests are the most useful way to monitor progression.

GRANULOMATOUS LUNG DISEASE WITH VASCULITIS

The classification of pulmonary vasculitis and granulomatous disorders is unsatisfactory. In broad terms it is reasonable to consider two main groups: the respiratory manifestations of systemic diseases, and disorders associated with the presence of anti-neutrophil cytoplasmic antibodies (ANCAs).

Anti-neutrophil cytoplasmic antibodies (ANCAs) (see also pp. 504 and 596) are found in the acute phase of vasculitides, particularly Wegener's granulomatosis, Churg–Strauss syndrome and microscopic polyangiitis (polyarteritis) associated with neutrophil infiltration of the vessel wall.

Two major ANCA reactivities are recognized: proteinase-3 (PR3) ANCA and myeloperoxidase (MPO) ANCA.

Approximately 10% of all vasculitis patients are ANCA-negative: this is more common in Wegener's granulomatosis limited to the upper respiratory tract. Ten to fifteen per cent of cases of progressive glomerulonephritis with anti-glomerular basement membrane (GBM) antibodies are MPO ANCA-positive and these are the most likely to suffer pulmonary haemorrhage.

Wegener's granulomatosis (see p. 551)

This granulomatous disease of unknown aetiology is one of the primary systemic vasculitides in which the small arteries are predominantly affected (the other is Churg–Strauss syndrome, see below). It is characterized by lesions involving the upper respiratory tract, the lungs and the kidneys. Often the disease starts with severe rhinorrhoea with subsequent nasal mucosal ulceration followed by cough, haemoptysis and pleuritic pain. Occasionally there may be involvement of the skin and nervous system. A chest X-ray usually shows single or multiple nodular masses or pneumonic infiltrates with cavitation. The most remarkable radiographic feature is the migratory pattern, with large lesions clearing in one area and new lesions appearing in another. The typical histological changes are usually best seen in the kidneys, where there is a *necrotizing microvascular glomerulonephritis*. This disease responds well to treatment with cyclophosphamide 150–200 mg daily. Rituximab is also being used. A variant of Wegener's granulomatosis called 'midline granuloma' affects the nose and paranasal sinuses and is particularly mutilating; it has a poor prognosis.

Churg–Strauss syndrome

This condition occurs usually in males in their fourth decade, who have a triad of rhinitis and asthma, eosinophilia and systemic vasculitis. The aetiology is uncertain, with some believing that it represents an unusual progression of allergic disease in a subset of predisposed individuals. Others believe it is a primary vasculitis which presents like asthma because of the involvement of eosinophils.

Ask the authors
www.studentconsult.com

The pathology of this condition is dominated by an eosinophilic infiltration with a characteristic high blood eosinophil count, vasculitis of small arteries and veins, and extravascular granulomas. Typically it involves the lungs, peripheral nerves and skin, but kidney involvement is uncommon. Transient patchy pneumonia-like shadows may occur, but sometimes these can be massive and bilateral. Skin lesions include tender subcutaneous nodules as well as petechial or purpuric lesions. ANCA is usually positive. The disease responds well to corticosteroids. Occasionally Churg–Strauss syndrome may be revealed when oral steroids are withdrawn in patients being treated for asthma. There is no reason to think that other anti-asthma drugs precipitate the condition.

Microscopic vasculitis (polyangiitis)

This involves the kidneys and the lungs where it results in recurrent haemoptysis. ANCA is usually positive. In the early literature there was confusion between this condition, Churg–Strauss syndrome and polyarteritis nodosa. The latter, however, is ANCA-negative and rarely involves the lungs.

Pulmonary autoimmune rheumatic disease

Rheumatoid disease (see also p. 527)

The features of respiratory involvement in rheumatoid disease are illustrated in Figure 14.40.

- *Pleural adhesions, thickening and effusions* are the most common lesions. The effusion is often unilateral and tends to be chronic. It has a low glucose content but this can occur in any chronic pleural effusion.

  Several forms of parenchymal disease can occur in patients with rheumatoid arthritis. These include fibrosing alveolitis, rheumatoid nodules, cryptogenic organizing pneumonia, lymphoid interstitial pneumonia and bronchiectasis; other pulmonary problems include pulmonary hypertension, fibrosis and others shown in the figure. Moreover, some patients will have modified presentations because they are already on disease-modifying drugs such as prednisolone or methotrexate for their arthritis.
- *Fibrosing alveolitis* occurring in rheumatoid arthritis can be considered as a variant of the cryptogenic form of the disease (see p. 872). The clinical features and gross appearance are the same but the disease is often more chronic.
- *Rheumatoid nodules* appear on the chest X-ray as single or multiple nodules ranging in size from a few millimetres to a few centimetres. The nodules frequently cavitate. They usually produce no symptoms but can give rise to a pneumothorax or pleural effusion.
- *Obliterative disease of the small bronchioles* is rare. It is characterized by progressive breathlessness and irreversible airflow limitation. Corticosteroids may prevent progression.
- *Cricoarytenoid joint involvement* by rheumatoid arthritis gives rise to dyspnoea, stridor, hoarseness and occasionally severe obstruction necessitating tracheostomy.
- *Caplan's syndrome* is due to a combination of dust inhalation and the disturbed immunity of rheumatoid arthritis. It occurs particularly in coal-worker's pneumoconiosis but it can occur in individuals exposed to other dusts, such as silica and asbestos. Typically the lesions appear as rounded nodules 0.5–5.0 cm in diameter, though sometimes they become incorporated into large areas of fibrosis that are indistinguishable radiologically from progressive massive fibrosis. There may not be much evidence of simple pneumoconiosis prior to the development of the nodule. These lesions may precede the development of the arthritis. Rheumatoid factor is always present in the serum.

Drugs used in the treatment of rheumatoid arthritis cause many pulmonary problems: pneumonitis, e.g. methotrexate, gold, NSAIDs; fibrosis, e.g. methotrexate; bronchospasm, e.g. NSAIDs, aspirin; infection, e.g. corticosteroids, methotrexate. With the use of anti-TNF therapy, reactivation of tuberculosis can occur.

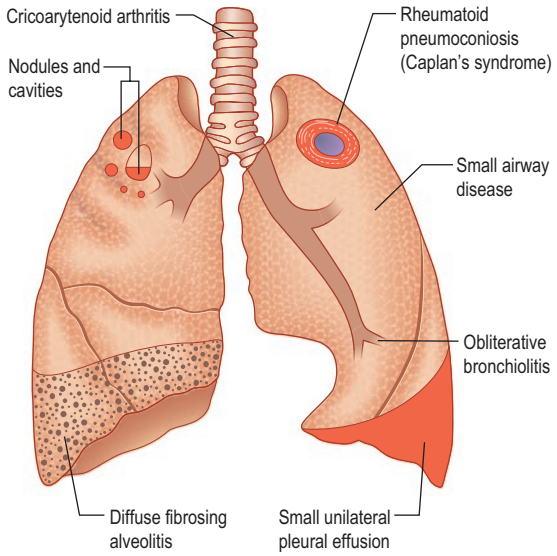


Fig. 14.40 **Respiratory manifestations of rheumatoid disease.** Many drugs affect the lungs, see text, page 877.

Systemic lupus erythematosus (see also p. 543)

Pleurisy is the most common respiratory manifestation of this disease, occurring in up to two-thirds of cases, with or without an effusion. Effusions are usually small and bilateral. Basal pneumonitis is often present, perhaps as a result of poor movement of the diaphragm, or restriction of chest movements because of pleural pain. Pneumonia also occurs, because of either infection or the disease process itself. In contrast to rheumatoid arthritis, diffuse pulmonary fibrosis is rare.

Systemic sclerosis (see pp. 545 and 1252)

Autopsy studies have indicated that there is almost always some diffuse fibrosis of alveolar walls and obliteration of capillaries and the alveolar space. Severe changes result in nodular then streaky shadowing on the chest X-ray, followed by cystic changes, ending up with a honeycomb lung. Lung function tests reveal a restrictive defect and poor gas transfer. Pneumonia may occur owing to aspiration from the dilated oesophagus (see p. 254). Breathlessness may be

worsened by restriction of chest wall movement owing to thickening and contraction of the skin and trunk.

IDIOPATHIC INTERSTITIAL PNEUMONIAS (IIP)

A recent international classification is shown in Box 14.7. IIP is characterized by diffuse inflammation and fibrosis in the lung parenchyma.

Idiopathic pulmonary fibrosis (IPF)

This is also known as *usual interstitial pneumonia* (UIP) and was previously known as cryptogenic fibrosing alveolitis (CFA).

It is relatively rare with a prevalence of up to 20 per 100 000 population, a mean onset in the late sixties and is more common in males. The cause is unknown but cigarette smoking, chronic aspiration, antidepressants, wood and metal dusts and infections, e.g. Epstein–Barr virus, may play a part.

Pathology

A patchy fibrosis of the interstitium (often with intervening normal lung), subpleural and paraseptal changes, minimal or absent inflammation, acute fibroblastic proliferation and collagen deposition (fibroblastic foci) and honeycombing are the key features.

Pathogenesis

There are many hypotheses but the epithelial/mesenchymal hypothesis (Fig. 14.41) suggests that injury results from repeated exogenous and endogenous unknown stimuli. Disruption of the alveolar epithelium and basement membrane activates inflammatory cells, which release proinflammatory cytokines and chemokines, pro-fibrogenic growth factors and endothelin-1 from endothelial cells. Fibroblasts differentiate into myofibroblasts, synthesize collagen and aggregate to form 'fibrotic foci'.

Box 14.7 Classification of idiopathic interstitial pneumonias

Clinical diagnosis	Pathological pattern
Idiopathic pulmonary fibrosis (IPF)	Usual interstitial pneumonia (UIP)
Desquamative interstitial pneumonia (DIP)	Desquamative interstitial pneumonia (DIP)
Respiratory bronchiolitis interstitial lung disease (RBILD)	Respiratory bronchiolitis interstitial lung disease (RBILD)
Acute interstitial pneumonia (AIP)	Diffuse alveolar damage (DAD)
Non-specific interstitial pneumonia (NSIP)	Non-specific interstitial pneumonia (NSIP)
Cryptogenic organizing pneumonia (COP)	Organizing pneumonia (OP)
Lymphoid interstitial pneumonia (LIP)	Lymphoid interstitial pneumonia (LIP)

ATS/ERS. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. *American Journal of Respiratory and Critical Care Medicine* 2002; **165**: 277.

Clinical features

The main features are progressive breathlessness, a non-productive cough and cyanosis, which eventually lead to respiratory failure, pulmonary hypertension and cor pulmonale. Gross finger clubbing occurs in two-thirds of cases and fine bilateral end-inspiratory crackles are heard on auscultation. An acute form known as the Hamman–Rich syndrome occurs in a small proportion of cases. A number of autoimmune diseases are seen in association with this condition (see also Box 14.6). IPF has also been reported in association with coeliac disease, ulcerative colitis and renal tubular acidosis.

Investigations

- **Chest X-ray** initially shows a ground glass appearance, followed by irregular reticulonodular shadowing (often maximal in the lower zones) and finally a honeycomb lung.
- **High-resolution CT scan (HRCT)** shows characteristic bilateral changes involving mainly the lower lobes (Fig. 14.42). There are subpleural reticular abnormalities with minimal or no ground glass changes, honeycombs, i.e. thick-walled cysts (Table 14.14) 0.5–2 cm in diameter in terminal and respiratory bronchioles, and traction bronchiectasis with a diagnostic accuracy of approximately 90%.
- **Respiratory function tests** show a restrictive ventilatory defect – the lung volumes are reduced, the FEV_1 to FVC ratio is normal to high (with both values being reduced), and carbon monoxide gas transfer is reduced. Peak flow rates may be normal.

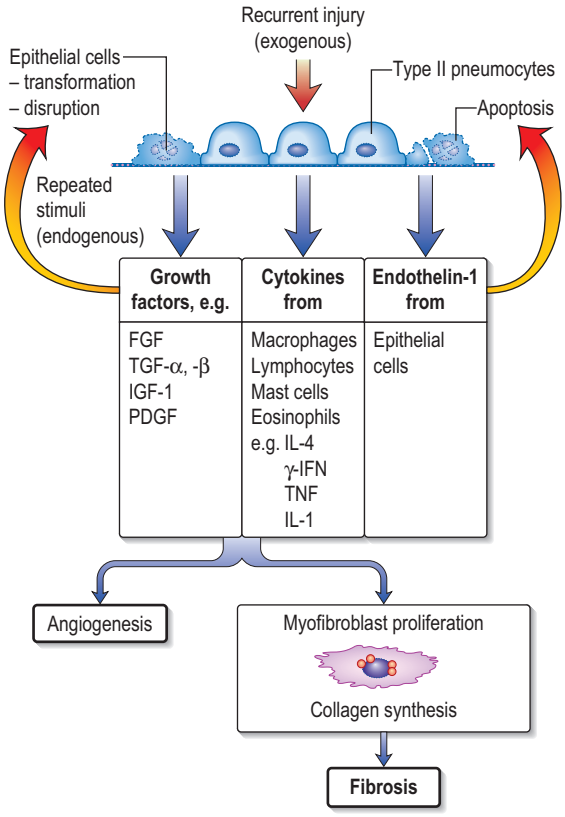


Fig. 14.41 **Pathogenesis of pulmonary fibrosis.**

Fig. 14.42 **CT scan showing idiopathic pulmonary fibrosis with reticular nodular shadowing and a honeycomb appearance.**

Table 14.14 The main causes of honeycomb lung

Localized	Diffuse
Systemic sclerosis	Idiopathic lung fibrosis
Sarcoidosis	Rheumatoid lung
Tuberculosis	Langerhans' cell histiocytosis
Asbestosis	Tuberous sclerosis
Berylliosis	Neurofibromatosis

- **Blood gases** show an arterial hypoxaemia caused by a combination of alveolar–capillary block and ventilation–perfusion mismatch with normal or low P_aCO_2 owing to hyperventilation.
- **Blood tests**. Anti-nuclear antibodies and rheumatoid factors are present in one-third of patients. The ESR and immunoglobulins are mildly elevated.
- **Bronchoalveolar lavage** shows increased numbers of cells, particularly neutrophils and macrophages.
- **Histological confirmation** is necessary in some patients. Transbronchial lung biopsy is rarely diagnostic, but can exclude other conditions which present similarly, e.g. sarcoidosis or lymphangitis carcinomatosa. A video-assisted thoracoscopic lung biopsy is performed to obtain a larger specimen, which will allow a clear histological diagnosis to be made.

Differential diagnosis

The diagnosis of IPF is usually made in a patient presenting with the above signs and characteristic HRCT changes. The differential diagnosis of the chest X-ray appearance includes extrinsic allergic alveolitis, bronchiectasis, chronic left heart failure, sarcoidosis, industrial lung disease and lymphangitis carcinomatosa.

Prognosis and treatment

The median survival time for patients with IPF is approximately 5 years, although mortality is very high in the more acute forms. Treatment with prednisolone (30 mg daily) is usually prescribed for disabling disease. Azathioprine or cyclophosphamide may be added if there is no response. A number of treatments have been tried, including *N*-acetylcysteine, warfarin, interferon, endothelin-1 antagonists (bosentan) and monoclonal antibodies, with variable results. New randomized controlled trials are awaited. Supportive treatment includes domiciliary oxygen therapy. In severe disease, single lung transplantation can be offered.

Desquamative interstitial pneumonia (DIP)

This occurs mainly in middle-aged male smokers and is less severe but similar to usual interstitial pneumonia (UIP). Pathologically there are more mononuclear cells than in UIP, due to smoking (pigmented macrophages). The prognosis is good with corticosteroid therapy.

Respiratory bronchiolitis interstitial lung disease (RBILD)

Clinically, RBILD is like DIP. Pathologically, pigmented macrophages are seen in the lumen of respiratory bronchioles. It has a better prognosis than UIP.

Acute interstitial pneumonia (AIP)

There is a very acute onset of pneumonia, often preceded by an upper respiratory tract infection, and progressive respiratory failure. Diffuse alveolar damage (DAD) is seen on lung biopsy. The prognosis is poor.

Non-specific interstitial pneumonia (NSIP)

The onset is subacute with a fever in 30%, a lack of finger clubbing and no male preponderance. It is unlike the other interstitial pneumonias histologically as there is a chronic interstitial pneumonia with mononuclear inflammatory cells and some fibrosis. Prognosis is variable (depending on amount of fibrosis).

Cryptogenic organizing pneumonia (COP)

This condition, previously called bronchiolitis obliterans organizing pneumonia (BOOP), is an organizing pneumonia of unknown aetiology. No infective agent has been described. Typically, patients present with single or recurrent episodes of malaise associated with cough, breathlessness and fever. Pleuritic chest pain is sometimes present but finger clubbing is very rare. Chest X-rays show confluent bilateral parenchymal shadowing. Lung function tests may show a restrictive defect. The white blood count is normal, but the ESR may be raised. Open lung biopsy will reveal characteristic buds of connective tissue (Masson's bodies) in respiratory bronchioles and in alveolar ducts. These are diagnostic, but the diagnosis is usually made on history and X-ray appearances. The disease responds rapidly to corticosteroid treatment but may recur episodically, especially in older women.

Lymphoid interstitial pneumonia

This condition is more common in children than in adults and is characterized by infiltration with lymphocytes, plasma cells

and immunoblasts. It is thought to be a viral pneumonia and causes diffuse reticulonodular infiltrates on the chest X-ray. Corticosteroid therapy appears to be of benefit, as is zidovudine.

Drugs

Drugs causing diffuse parenchymal lung disease are shown on page 877 and in Table 14.17.

OTHER TYPES OF DIFFUSE LUNG DISEASE

Langerhans' cell histiocytosis (LCH)

This rare disease (a prevalence of 1 per 50 000) is characterized histologically by proliferation of Langerhans' cells identified by the presence of Birbeck granules on electron microscopy or the CD1a antigen on the surface of the cells. There is a wide variation in clinical presentation, from unifocal bone lesions in older children (which may regress spontaneously), to more disseminated disease in younger children (with a high mortality). Pulmonary LCH occurs almost exclusively in smokers. Chest X-rays (and HRCT) show multiple small cysts (honeycomb lung), fibrosis or widespread nodular shadows. *Treatment* is with stopping smoking, corticosteroids and immunosuppressive therapy. For advanced progressive disease, lung transplantation is the only option. Five-year and 10-year survivals are 75% and 65% respectively.

Goodpasture's syndrome (see also p. 595)

The disease often starts with symptoms of an upper respiratory tract infection followed by cough and intermittent haemoptysis, tiredness and eventually anaemia, although massive bleeding may occur. The chest X-ray shows transient blotchy shadows that are due to intrapulmonary haemorrhage. These features usually precede the development of an acute glomerulonephritis by several weeks or months. The course of the disease is variable: some patients spontaneously improve while others proceed to renal failure.

The disease usually occurs in individuals over 16 years of age. It is due to a type II cytotoxic reaction driven by antibodies directed against the basement membrane of both kidney and lung. It has been proposed that there may be a shared antigen. ANCA may be positive. An association with influenza A2 virus has been reported.

Treatment is with corticosteroids, but in some cases dramatic improvement has been seen with plasmapheresis to remove the autoantibodies.

Diffuse alveolar haemorrhage

This is clinically similar to Goodpasture's syndrome, but there are no anti-basement-membrane antibodies and the kidneys are less frequently involved. Most cases occur in children under 7 years of age. The child develops a chronic cough and anaemia. The chest X-ray shows diffuse shadows that are due to intrapulmonary bleeding, and eventually miliary nodulation. Characteristically, haemosiderin-containing macrophages are found in the sputum. There is an association with a sensitivity to cow's milk, and an appropriate diet is usually tried.

The prognosis in general is poor and treatment with corticosteroids or azathioprine is usually given.

Lymphangioleiomyomatosis

This is a rare disorder of young women with hamartomatous smooth muscle infiltration of the lungs. Patients present with dyspnoea, chylous pleural effusions, haemoptysis and pneumothorax. Treatment is with hormones/oophorectomy. Prognosis is variable.

Pulmonary alveolar proteinosis

This is a rare disease in which there is accumulation of lipoproteinaceous material within the alveoli. It can be congenital, but most cases are acquired and appear to have an autoimmune basis, with antibodies directed against the cytokine GM-CSF. The disease mostly affects men and presents with progressive exertional dyspnoea and cough. Inspiratory crackles are present in only about 50%. Diagnosis is made by bronchial lavage, which reveals a milky appearance and many large, foamy macrophages but few other inflammatory cells.

FURTHER READING

Danoff SK, Terry PB, Horton MR. A clinician's guide to the diagnosis and treatment of interstitial lung disease. *Southern Medical Journal* 2007; **100**: 579–587.

Iannuzzi MC, Rybicki BA, Teirstein AS. Sarcoidosis. *New England Journal of Medicine* 2007; **357**: 2153–2165.

King TE. Clinical advances in the diagnosis and therapy of interstitial lung disease. *American Journal of Respiratory and Critical Care Medicine* 2005; **172**: 268–279.

Noth I, Strek ME, Leff AR. Churg–Strauss syndrome. *Lancet* 2003; **361**: 587–594.

Tazi A. Adult pulmonary Langerhans' cell histiocytosis. *European Respiratory Journal* 2006; **27**: 1272–1285.

PULMONARY INFILTRATION WITH EOSINOPHILIA

The common types and characteristics of these diseases are shown in Table 14.15. They range from very mild, simple, pulmonary eosinophilias to the often fatal hypereosinophilic syndrome.

Simple and prolonged pulmonary eosinophilia

Simple pulmonary eosinophilia is a relatively mild illness with a slight fever and cough and usually lasts for less than 2 weeks. Occasionally, the disease becomes more prolonged, with a high fever lasting for over a month. There is usually an eosinophilia in the blood and this condition is then called prolonged pulmonary eosinophilia. In both conditions the chest X-ray shows either localized or diffuse opacities. The simple form is probably due to a transient allergic reaction in the alveoli. Many allergens have been implicated, including *Ascaris lumbricoides, Ankylostoma, Trichuris, Trichinella,*

Table 14.15 Common types and characteristics of pulmonary infiltration with eosinophilia

Disease	Symptoms	Blood eosinophils (%)	Multisystem involvement	Duration	Outcome
Simple pulmonary eosinophilia	Mild	10	None	<1 month	Good
Prolonged pulmonary eosinophilia	Mild/moderate	>20	None	>1 month	Good
Asthmatic bronchopulmonary eosinophilia	Moderate/severe	5–20	None	Years	Fair
Tropical pulmonary eosinophilia	Moderate/severe	>20	None	Years	Fair
Hypereosinophilic syndrome	Severe	>20	Always	Months/years	Poor

Taenia and *Strongyloides*. Drugs such as aspirin, penicillin, nitrofurantoin and sulphonamides have been implicated. Often, no allergen is identified. The disease is self-limiting and no treatment is required, apart from treating the cause. In the more chronic form all unnecessary treatment should be withdrawn and, where appropriate, worms are treated. Corticosteroid therapy is indicated, with resolution of the disease over the ensuing weeks.

Asthmatic bronchopulmonary eosinophilia

This is characterized by the presence of asthma, transient fleeting shadows on the chest X-ray, and blood or sputum eosinophilia. By far the most common cause world-wide is allergy to *A. fumigatus* (see below), although *Candida albicans* and other mycoses may be the allergen in a small number of patients. In many, no allergen can be identified. Whether these cases are intrinsic or driven by an unidentified extrinsic factor is uncertain.

Diseases caused by *Aspergillus fumigatus*

The various types of lung disease caused by *A. fumigatus* are illustrated in Figure 14.43.

The spores of *A. fumigatus* (diameter 5 mm) are readily inhaled and are present in the atmosphere throughout the year, though they are at their highest concentration in the late autumn. They can be grown from the sputum in up to 15% of patients with chronic lung disease in whom they do not produce disease. They are a cause of extrinsic asthma in atopic individuals.

Allergic bronchopulmonary aspergillosis

In this rare disease, *Aspergillus* actually grows in the walls of the bronchi and eventually produces proximal bronchiectasis. There are episodes of eosinophilic pneumonia throughout the year, particularly in late autumn and winter. The episodes present with a wheeze, cough, fever and malaise. They are associated with expectoration of firm sputum plugs containing the fungal mycelium, which results in the clearing of the pulmonary infiltrates on the chest X-ray. Occasionally the large mucus plugs obliterate the bronchial lumen, causing collapse of the lung.

Left untreated, repeated episodes of eosinophilic pneumonia can result in progressive pulmonary fibrosis that usually affects the upper zones and can give rise to a chest X-ray appearance similar to that produced by tuberculosis.

The peripheral blood eosinophil count is usually raised, and total levels of IgE are usually extremely high (both that specific to *Aspergillus* and non-specific). Skin-prick testing to protein allergens from *A. fumigatus* gives rise to positive immediate skin tests. Sputum may show eosinophils and mycelia, and precipitating antibodies are usually, but not always, found in the serum.

Lung function tests show a decrease in lung volumes and gas transfer in more chronic cases, but there is evidence of reversible airflow limitation in all cases.

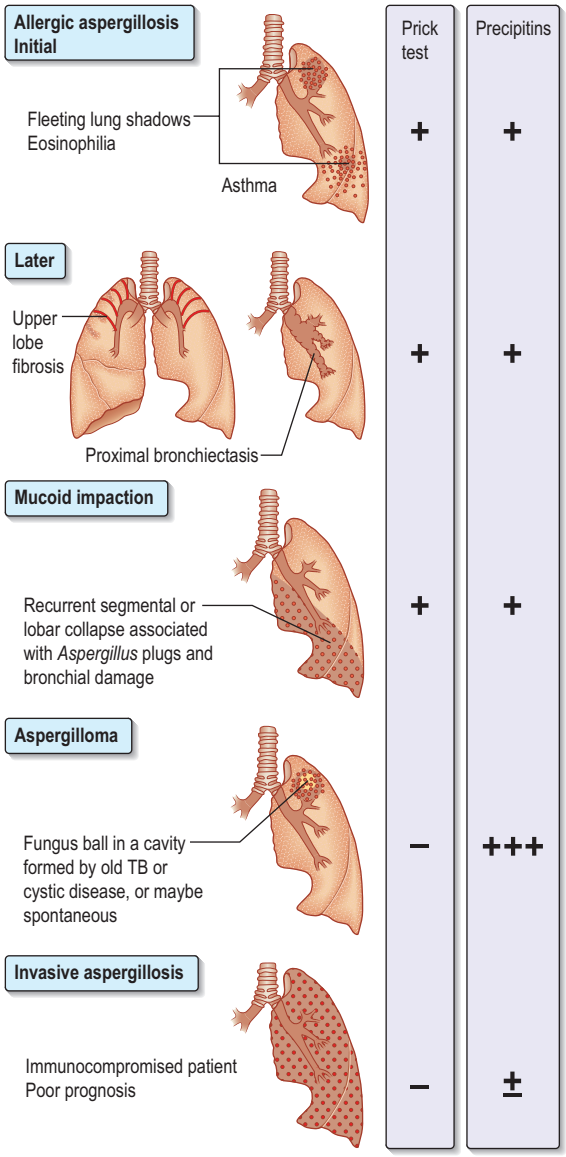


Fig. 14.43 **Diseases caused by *Aspergillus fumigatus*.** Allergic aspergillosis (initial and later) and three other forms.

FURTHER READING

Zmeili OS, Soubani AO. Pulmonary aspergillosis – update. *Quarterly Journal of Medicine* 2007; **100**: 317–334.

Treatment is with prednisolone 30 mg daily, which causes rapid clearing of the pulmonary infiltrates. Frequent episodes of the disease can be prevented by long-term treatment with prednisolone, but doses of 10–15 mg daily are usually required. Several antifungal agents have been tried in the past without success, but there is evidence that treatment with itraconazole improves pulmonary function. The asthma component responds to inhaled corticosteroids, although they do not influence the occurrence of pulmonary infiltrates.

Aspergilloma and invasive aspergillosis

Aspergilloma is the growth of *A. fumigatus* within previously damaged lung tissue where it forms a ball of mycelium within lung cavities. Typically the chest X-ray shows a round lesion with an air 'halo' above it. Continuing antigenic stimulation gives rise to large quantities of precipitating antibody in the serum. The aspergilloma itself causes little trouble, though occasionally massive haemoptysis may occur, requiring resection of the area of damaged lung containing the aspergilloma. Antifungal agents, such as amphotericin, have been tested in aspergilloma, but with little success. Invasive aspergillosis is a well-recognized complication of immunosuppression and requires aggressive antifungal therapy usually with i.v. amphotericin (250 μg/kg daily). Itraconazole or voriconazole is used if amphotericin is ineffective, and caspofungin may also work.

Tropical pulmonary eosinophilia

This term is reserved for an allergic reaction to microfilaria from *Wuchereria bancrofti*. The condition is seen in the Asian subcontinent and presents with cough and wheeze together with fever, lassitude and weight loss. The typical appearance of the chest X-ray is of bilateral hazy mottling that is often uniformly distributed in both lung fields. Individual shadows may be as large as 5 mm or may become more confluent, giving the appearance of pneumonia.

The disease is characterized by a very high eosinophil count in peripheral blood. The filarial complement fixation test is positive in almost every case, although the microfilaria are seldom found. The treatment of choice is diethylcarbamazine (see p. 163 for details).

The hypereosinophilic syndrome

This disease is characterized by eosinophilic infiltration in various organs, sometimes associated with an eosinophilic arteritis. The heart muscle is particularly involved, but pulmonary involvement in the form of a pleural effusion or interstitial lung disease occurs in about 40% of cases. Typical features are fever, weight loss, recurrent abdominal pain, persistent non-productive cough and congestive cardiac failure. Corticosteroid treatment may be of value in some cases.

EXTRINSIC ALLERGIC ALVEOLITIS

In this disease there is a widespread diffuse inflammatory reaction in both the small airways of the lung and the alveoli. It is due to the inhalation of a number of different antigens, the most common being microbial spores contaminating vegetable matter (e.g. straw, hay, mushroom compost). Some examples are illustrated in Table 14.16. By far the most common of these diseases world-wide is farmer's lung, which affects up to 1 in 10 of the farming community in disadvantaged, wet communities around the world. Cigarette smokers have a lower risk of developing the disease due to decreased antibody reaction to the antigen.

Table 14.16 Extrinsic allergic (bronchiolar) alveolitis – some causes

Disease	Situation	Antigens
Farmer's lung	Forking mouldy hay or any other mouldy vegetable material	Thermophilic actinomycetes, e.g. *Micropolyspora faeni* Fungi, e.g. *Aspergillus umbrasus*
Bird fancier's lung	Handling pigeons, cleaning lofts or budgerigar cages	Proteins present in the 'bloom' on the feathers and in excreta
Maltworker's lung	Turning germinating barley	*Aspergillus clavatus*
Humidifier fever	Contaminated humidifying systems in air conditioners or humidifiers in factories (especially in printing works)	Possibly a variety of bacteria or amoeba (e.g. *Naegleria gruberi*) Thermoactinomyces
Mushroom workers	Turning mushroom compost	Thermophilic actinomycetes
Cheese washer's lung	Mouldy cheese	*Penicillin casei* *Aspergillus clavatus*
Winemaker's lung	Mould on grapes	Botrytis

Pathogenesis

Histologically there is an initial infiltration of the small airways and alveolar walls with neutrophils followed by T lymphocytes and macrophages, leading to the development of small non-caseating granulomas. These comprise multinucleated giant cells, occasionally containing the inhaled antigenic material. The allergic response to the inhaled antigens involves both cellular immunity and the deposition of immune complexes causing foci of inflammation through the activation of complement via the classical pathway. Some of the inhalant materials may also lead to inflammation by directly activating the alternative complement pathway. These mechanisms attract and activate alveolar and interstitial macrophages, so that continued antigenic exposure results in the progressive development of pulmonary fibrosis.

Clinical features

Typically fever, malaise, cough and shortness of breath come on several hours after exposure to the causative antigen. Thus, a farmer forking hay in the morning may notice symptoms during the late afternoon and evening with resolution by the following morning. On examination, the patient may have a fever, tachypnoea, and coarse end-inspiratory crackles and wheezes throughout the chest. Cyanosis caused by ventilation–perfusion mismatch may be severe even at rest. Continued exposure leads to a chronic illness characterized by severe weight loss, effort dyspnoea and cough as well as the features of idiopathic pulmonary fibrosis (see p. 872).

Investigations

- **Chest X-ray** shows fluffy nodular shadowing with the subsequent development of streaky shadows, particularly in the upper zones. In very advanced cases, honeycomb lung occurs.
- **High-resolution CT** shows reticular and nodular changes with ground glass opacity, which can be further categorized with multislice CT.
- **Lung function tests** show a restrictive ventilatory defect with a decrease in carbon monoxide gas transfer.
- **Polymorphonuclear leucocyte count** is raised in acute cases. Eosinophilia is not a feature.
- **Precipitating antibodies** are present in the serum. One-quarter of pigeon fanciers have precipitating IgG antibodies against pigeon protein and droppings in their serum, but only a small proportion have lung disease. Precipitating antibodies are evidence of exposure, not disease.
- **Bronchoalveolar lavage** shows increased T lymphocytes and granulocytes.

Differential diagnosis

Although extrinsic allergic alveolitis due to inhalation of the spores of *Micropolyspora faeni* is common among farmers, it is probably more common for these individuals to suffer from asthma related to inhalation of antigens from a variety of mites that infest stored grain and other vegetable material. These include *Lepidoglyphus domesticus, L. destructor and Acarus siro*. Symptoms of asthma resulting from inhalation of these allergens are often mistaken for farmer's lung. Lung function tests will effectively discriminate between the disorders. Pigeon fancier's lung is quite common, but alveolitis from budgerigars, parrots and parakeets is very rare.

Management

Prevention is the aim. This can be achieved by changes in work practice, with the use of silage for animal fodder and the drier storage of hay and grain. Pigeon fancier's lung is more difficult to control since affected individuals remain strongly attached to their hobby. Prednisolone, initially in large doses of 30–60 mg daily, may achieve regression during the early stages of the disease. Established fibrosis will not resolve and in some patients the disease may progress inexorably to respiratory failure in spite of intensive therapy. Farmer's lung is a recognized occupational disease in the UK and sufferers are entitled to compensation, depending upon their degree of disability.

Humidifier fever (Table 14.16)

Humidifier fever, one cause of building-related illnesses (p. 961), may present with the typical features of extrinsic allergic alveolitis without any radiographic changes. This disease has occurred in outbreaks in factories in the UK, particularly in printing works. In North America it is more commonly found in office blocks with contaminated air-conditioning systems. Humidifier fever may be effectively prevented by sterilization of the re-circulating water used in large humidifying plants.

Drug and radiation-induced respiratory reactions (Table 14.17)

Drugs may produce a wide variety of disorders of the respiratory tract. The mechanisms are varied and include direct toxicity (e.g. bleomycin), immune complex formation with arteritis, hypersensitivity (involving both T cell and IgE mechanisms) and autoimmunity. Tuberculosis reactivation is seen with immunosuppressive drugs.

Pulmonary infiltrates with fibrosis may result from the use of a number of cytotoxic drugs used in the treatment of cancer. The most common cause of these reactions is bleomycin. The pulmonary damage is dose-related, occurring when the total dosage is greater than 450 mg, but will regress in some cases if the drug is stopped. The most sensitive test is a decrease in carbon monoxide gas transfer, and therefore gas transfer should be measured repeatedly during treatment with the drug. The use of corticosteroids may help resolution. Drugs affecting the respiratory system are shown in Table 14.17, together with the types of reaction they produce. Anaphylaxis with bronchospasm can occur with many drugs. The list is not exhaustive; for example, over 20 different drugs are known to produce a systemic lupus erythematosus-like

Table 14.17 Some drug-induced respiratory reactions

Disease	Drugs
Bronchospasm	Penicillins, cephalosporins
	Sulphonamides
	Aspirin/NSAIDs
	Monoclonal antibodies, e.g. infliximab
	Iodine-containing contrast media
	β-Adrenoceptor-blocking drugs, (e.g. propranolol)
	Non-depolarizing muscle relaxants
	Intravenous thiamine
	Adenosine
Diffuse parenchymal lung disease and/or fibrosis	Amiodarone
	Anakinra (IL-1 receptor antagonist)
	Nitrofurantoin
	Paraquat
	Continuous oxygen
	Cytotoxic agents (many, particularly busulfan, CCNU, bleomycin, methotrexate)
Pulmonary eosinophilia	Antibiotics:
	Penicillin
	Tetracycline
	Sulphonamides, e.g. sulfasalazine
	NSAIDs
	Cytotoxic agents
Acute lung injury	Paraquat
Pulmonary hypertension	Fenfluramine, dexfenfluramine, phentermine
SLE-like syndrome including pulmonary infiltrates, effusions and fibrosis	Hydralazine
	Procainamide
	Isoniazid
	Phenytoin
	ACE inhibitors
	Monoclonal antibodies
Reactivation of tuberculosis	Immunosuppressant drugs, e.g. steroids
	Biological agents, e.g. tumour necrosis factor blockers

CCNU, chloroethyl-cyclohexyl-nitrosourea (lomustine); NSAIDs, non-steroidal anti-inflammatory drugs; SLE, systemic lupus erythematosus.

syndrome, sometimes complicated by pulmonary infiltrates and fibrosis. Paraquat ingestion causes severe pulmonary oedema and death, and pulmonary fibrosis develops in many of the few who survive.

Irradiation of the lung during radiotherapy can cause a radiation pneumonitis. Patients complain of breathlessness and a dry cough. Radiation pneumonitis results in a restrictive lung defect. Corticosteroids should be given in the acute stage.

OCCUPATIONAL LUNG DISEASE

Exposure to dusts, gases, vapours and fumes at work can cause several different types of lung disease:

- acute bronchitis and even pulmonary oedema from irritants such as sulphur dioxide, chlorine, ammonia or the oxides of nitrogen
- pulmonary fibrosis due to mineral dust
- occupational asthma (see Table 14.12) – this is now the commonest industrial lung disease in the developed world
- extrinsic allergic alveolitis (see Table 14.16)
- bronchial carcinoma due to industrial agents (e.g. asbestos, polycyclic hydrocarbons, radon in mines).

The degree of fibrosis that follows inhalation of mineral dust varies. While iron (siderosis), barium (baritosis) and tin (stannosis) lead to dramatic dense nodular shadowing on the chest X-ray, their effect on lung function and symptoms is minimal. Exposure to silica or asbestos, on the other hand, leads to extensive fibrosis and disability. Coal dust has an intermediate fibrogenic effect and used to account for 90% of all compensated industrial lung diseases in the UK. The term 'pneumoconiosis' means the accumulation of dust in the lungs and the reaction of the tissue to its presence. The term is not wide enough to encompass all occupational lung disease and is now generally used only in relation to coal dust and its effects on the lung.

Coal-worker's pneumoconiosis

The disease is caused by dust particles approximately 2–5 μm in diameter that are retained in the small airways and alveoli of the lung. The incidence of the disease is related to total dust exposure, which is highest at the coal face, particularly if ventilation and dust suppression are poor. Improved ventilation and working conditions have reduced the risk of this disease.

Two very different syndromes result from the inhalation of coal.

Simple pneumoconiosis

This simply reflects the deposition of coal dust in the lung. It produces fine micronodular shadowing on the chest X-ray and is by far the most common type of pneumoconiosis. It is graded on the chest X-ray appearance according to standard categories set by the International Labour Office (see below). Considerable dispute remains about the effects of simple pneumoconiosis on respiratory function and symptoms. In many cases the symptoms are due to COPD related to cigarette smoking, but this is not always the case. Changes to UK workers' compensation legislation means that coal miners who develop COPD may be compensated for their disability regardless of their chest X-ray appearance.

Categories of simple pneumoconiosis are as follows:

1. small round opacities definitely present but few in number
2. small round opacities numerous but normal lung markings still visible
3. small round opacities very numerous and normal lung markings partly or totally obscured.

Simple pneumoconiosis may lead to the development of progressive massive fibrosis (PMF) (see below). PMF virtually never occurs on a background of category 1 simple pneumoconiosis but occurs in about 7% of those with category 2 and in 30% of those with category 3. Miners with category 1 pneumoconiosis are unlikely to receive compensation unless they also have evidence of COPD. Those with more extensive radiographic changes may be compensated solely on the basis of their X-ray appearances.

Progressive massive fibrosis

In PMF, patients develop round fibrotic masses several centimetres in diameter, almost invariably in the upper lobes and sometimes having necrotic central cavities. The pathogenesis of PMF is still not understood, though it seems clear that some fibrogenic promoting factor is present in individuals developing the disease, leading to the formation of immune complexes, analogous to the development of large fibrotic nodules in coal miners with rheumatoid arthritis (Caplan's syndrome). Rheumatoid factor and anti-nuclear antibodies are both often present in the serum of patients with PMF, and also in those suffering from asbestosis or silicosis. Pathologically there is apical destruction and disruption of the lung, resulting in emphysema and airway damage. Lung function tests show a mixed restrictive and obstructive ventilatory defect with loss of lung volume, irreversible airflow limitation and reduced gas transfer.

The patient with PMF suffers considerable effort dyspnoea, usually with a cough. The sputum may be black. The disease can progress (or even develop) after exposure to coal dust has ceased and may lead to respiratory failure.

Silicosis

This disease is uncommon though it may still be encountered in stonemasons, sand-blasters, pottery and ceramic workers and foundry workers involved in fettling (removing sand from metal castings made in sand-filled moulds). Silicosis is caused by the inhalation of silica (silicon dioxide). This dust is highly fibrogenic. For example, a coal miner can remain healthy with 30 g of coal dust in his lungs but 3 g of silica is sufficient to kill. Silica seems particularly toxic to alveolar macrophages and readily initiates fibrogenesis (see Fig. 14.41). The chest X-ray appearances and clinical features of silicosis are similar to those of PMF, but distinctive thin streaks of calcification may be seen around the hilar lymph nodes ('eggshell' calcification).

Asbestosis

Asbestos is a mixture of silicates of iron, magnesium, nickel, cadmium and aluminium, and has the unique property of occurring naturally as a fibre. It is remarkably resistant to

heat, acid and alkali, and has been widely used for roofing, insulation and fireproofing. Asbestos has been mined in southern Africa, Canada, Australia and eastern Europe. Several different types of asbestos are recognized: about 90% of asbestos is chrysotile, 6% crocidolite and 4% amosite. Chrysotile or white asbestos is the softest asbestos fibre. Each fibre is often as long as 2 cm but only a few microns thick. It is less fibrogenic than crocidolite.

Crocidolite (blue asbestos) is particularly resistant to chemical destruction and exists in straight fibres up to 50 mm in length and 1–2 μm in width. Crocidolite is the type of asbestos most likely to produce asbestosis and mesothelioma. This may be due to the fact that it is readily trapped in the lung. Its long, thin shape means that it can be inhaled, but subsequent rotation against the long axis of the smaller airways, particularly in turbulent airflow during expiration, causes the fibres to impact. Crocidolite is also particularly resistant to macrophage and neutrophil enzymatic destruction.

Exposure to asbestos occurred particularly in shipbuilding yards and in power stations, but it was used so widely that low levels of exposure were very common. Up to 50% of city dwellers have asbestos bodies (asbestos fibres covered in protein secretions) in their lungs at post mortem. Regulations in the UK prohibit the use of crocidolite and severely restrict the use of chrysotile. Careful dust control measures are enforced, which should eventually abolish the problem. Workers continue to be exposed to blue asbestos in the course of demolition or in the replacement of insulation, and it should be remembered that there is a considerable time lag between exposure and development of disease, particularly mesothelioma (20–40 years).

The risk of primary lung cancer (usually adenocarcinoma) is increased in people exposed to asbestos, even in non-smokers. This risk is about 5–7-fold greater in those who have parenchymal asbestosis and about 1.5-fold in those with pleural plaques without parenchymal fibrosis. A synergistic relationship exists between asbestosis and cigarette smoking with the risk of bronchial carcinoma multiplied 5-fold above the risk attributable to smoking alone.

Diseases caused by asbestos are summarized in Table 14.18. Bilateral diffuse pleural thickening, asbestosis, mesothelioma and asbestos-related carcinoma of the bronchus are all eligible for industrial injuries benefit in the UK.

Asbestosis

Asbestosis is defined as fibrosis of the lungs caused by asbestos dust, which may or may not be associated with fibrosis of the parietal or visceral layers of the pleura. It is a progressive disease characterized by breathlessness and accompanied by finger clubbing and bilateral basal end-inspiratory crackles. Fibrosis, not detectable on chest X-ray, is often revealed on high-resolution CT scan. No treatment is known to alter the progress of the disease, though corticosteroids are often prescribed.

Mesothelioma

The number of cases of mesothelioma has increased progressively since the mid-1980s and in 2001 there were 1848 deaths in the UK. Rates of mesothelioma are continuing to increase in the UK, and are expected to peak at about 2300 per year, between 2010 and 2020. Pleural effusions are the most common presentation of mesothelioma, typically with persistent chest wall pain, which should raise the index of

FURTHER READING

Lin RT, Takahashi K, Karjalainen A et al. Ecological association between asbestos-related diseases and historical asbestos consumption. *Lancet* 2007; **369**: 844–849.

Table 14.18 The effects of asbestos on the lung

	Exposure	Chest X-ray	Lung function	Symptoms	Outcome
Asbestos bodies	Light	Normal	Normal	None	Evidence of asbestos exposure only
Pleural plaques	Light	Pleural thickening (parietal pleura) and calcification (also in diaphragmatic pleura)	Mild restrictive ventilatory defect	Rare, occasional mild effort dyspnoea	No other sequelae
Effusion	First two decades following exposure	Effusion	Restrictive	Pleuritic pain, dyspnoea	Often recurrent
Bilateral diffuse pleural thickening	Light/moderate	Bilateral diffuse thickening (of both parietal and visceral pleura) more than 5 mm thick and extending over more than one-quarter of the chest wall	Restrictive ventilatory defect	Effort dyspnoea	May progress in absence of further exposure
Mesothelioma	Light (interval of 20–40 years from exposure to disease)	Pleural effusion, usually unilateral	Restrictive ventilatory defect	Pleuritic pain, increasing dyspnoea	Median survival 2 years
Asbestosis	Heavy (interval of 5–10 years from exposure to disease)	Diffuse bilateral streaky shadows, honeycomb lung	Severe restrictive ventilatory defect and reduced gas transfer	Progressive dyspnoea	Poor, progression in some cases after exposure
Asbestos-related carcinoma of the bronchus	The features of asbestosis, bilateral diffuse pleural thickening or bilateral pleural plaques plus those of bronchial carcinoma				Fatal

suspicion even if the initial pleural fluid or biopsy samples are non-diagnostic. Often a video-assisted thoracoscopic lung biopsy is needed to obtain sufficient tissue for diagnosis. In the event of a positive pleural biopsy diagnosis, local radiotherapy should be given to prevent the seeding of mesothelioma cells down the needle track. Pemetrexed and vinorelbine are being tested in trials but no treatment influences the universally fatal outcome.

Byssinosis

This disease occurs world-wide but is declining rapidly in areas where the numbers of people employed in cotton mills are falling. The symptoms start on the first day back at work after a break (Monday sickness) with improvement as the week progresses. Tightness in the chest, cough and breathlessness occur within the first hour in dusty areas of the mill, particularly in the blowing and carding rooms where raw cotton is cleaned and the fibres are straightened.

The exact nature of the disease and its aetiology remain disputed. Two features are that pure cotton does not cause the disease, and that cotton dust has some effect on airflow limitation in all those exposed. Individuals with asthma are particularly badly affected by exposure to cotton dust. The most likely aetiology is endotoxins from bacteria present in the raw cotton causing constriction of the airways of the lung. There are no changes on the chest X-ray and there is considerable dispute as to whether the progressive airflow limitation seen in some patients with the disease is due to the cotton dust or to other factors such as cigarette smoking or coexistent asthma.

Berylliosis

Beryllium–copper alloy has a high tensile strength and is resistant to metal fatigue, high temperature and corrosion. It is used in the aerospace industry, in atomic reactors and in many electrical devices.

When beryllium is inhaled, it can cause a systemic illness with a clinical picture similar to sarcoidosis. The major chronic problem is that of progressive dyspnoea with pulmonary fibrosis. However, strict control of levels in the working atmosphere has made the disease a rarity.

LUNG CYSTS

These may be congenital, bronchogenic cysts or may result from a sequestrated pulmonary segment. Hydatid disease causes fluid-filled cysts. Thin-walled cysts are due to lung abscesses, which are particularly found in staphylococcal pneumonia, tuberculous cavities, septic pulmonary infarction, primary bronchogenic carcinoma, cavitating metastatic neoplasm, or paragonimiasis caused by the lung fluke *Paragonimus westermani*.

TUMOURS OF THE RESPIRATORY TRACT

Bronchial carcinoma accounts for 95% of all primary tumours of the lung. Alveolar cell carcinoma accounts for 2% of lung tumours, and other less malignant or benign tumours account for the remaining 3%.

MALIGNANT TUMOURS

Bronchial carcinoma

This is the most common malignant tumour world-wide with 1.3 million deaths annually. It is the third most common cause of death in the UK after heart disease and pneumonia. Mortality rates world-wide are highest in Scotland, closely followed by England and Wales. In the UK, 32 000 people die each year from bronchial carcinoma, with a male-to-female ratio of 3:1. Although the mortality rate from this disease has levelled off in men, it continues to rise in women and now causes more deaths from malignant disease in women than any other tumour, including breast cancer.

The strength of the association between cigarette smoking and bronchial carcinoma overshadows any other aetiological factors (Table 14.19), but there is a higher incidence of bronchial carcinoma in urban compared with rural areas, even when allowance is made for cigarette smoking. Passive smoking (the frequent inhalation of other people's smoke by non-smokers) increases the risk of bronchial carcinoma by a factor of 1.5. Occupational factors include exposure to asbestos, and an association is also claimed for workers in contact with arsenic, chromium, iron oxide, petroleum products and oils, coal tar, products of coal combustion, and radiation. Tumours associated with occupational factors are mostly adenocarcinomas and appear to be less related to cigarette smoking.

Cell types

Based on the characteristics of the disease and its response to treatment bronchial carcinoma is divided into small-cell carcinoma and non-small-cell carcinoma (NSCC). Studies of mean doubling times of carcinomas indicate that development from the initial malignant change to presentation takes about 15 years for adenocarcinoma, 8 years for squamous carcinoma and 3 years for small-cell carcinoma.

Non-small-cell carcinoma

Squamous or epidermoid carcinoma is the commonest type, accounting for approximately 40% of all carcinomas. Most present as obstructive lesions of the bronchus leading to infection. It occasionally cavitates (10%) at presentation. The cells are usually well differentiated but occasionally anaplastic. Local spread is common but widespread metastases occur relatively late.

Adenocarcinoma arises from mucous cells in the bronchial epithelium. Invasion of the pleura and the mediastinal lymph nodes is common, as are metastases to the brain, adrenal glands and bones. Adenocarcinoma accounts for

Table 14.19 Death rates from lung cancer (age standardized) per 100 000 according to smoking habits in male British doctors

	Death rate	Number of cigarettes per day	Death rate
Non-smokers	10	1–14	78
Ex-smokers	43	15–24	127
Continuing smokers		25 or more	251
Any tobacco	104		
Pipe/cigar	58		
Cigarettes	140		

approximately 10% of all bronchial carcinomas. It is the most common bronchial carcinoma associated with asbestos and is proportionally more common in non-smokers, in women, in the elderly, and in the Far East.

Large cell carcinomas are less-differentiated forms of squamous cell and adenocarcinomas. These account for about 25% of all lung cancers and metastasize early.

Bronchoalveolar cell carcinoma (also termed bronchiolar carcinoma) accounts for only 1–2% of lung tumours and occurs either as a peripheral solitary nodule or as diffuse nodular lesions of multicentric origin. Occasionally this tumour is associated with expectoration of very large volumes of mucoid sputum.

Small-cell carcinoma

This tumour accounts for 20–30% of all lung cancers. It arises from endocrine cells (Kulchitsky cells). These cells are members of the APUD system, which explains why many polypeptide hormones are secreted by these tumours. Some of these polypeptides act in an autocrine fashion: they feed back on the cells and cause cell growth. Small-cell carcinoma spreads early and is almost always inoperable at presentation. The tumour is rapidly growing and highly malignant. It responds to chemotherapy and radiotherapy but the overall prognosis remains poor.

Clinical features

The frequencies of the common symptoms of lung cancer on presentation are shown in Table 14.20. Chest pain and discomfort are often described as fullness and pressure in the chest. Sometimes the pain may be pleuritic owing to invasion of the pleura or ribs.

Often there are no abnormal physical signs. Enlarged supraclavicular lymph nodes may be present. There may be signs of a pleural effusion or of lobar collapse. Signs of an unresolved pneumonia or of associated underlying disease (e.g. diffuse pulmonary fibrosis in asbestosis) may be present.

Direct spread

The tumour may directly involve the pleura and ribs. Carcinoma in the apex of the lung can erode the ribs and involve the lower part of the brachial plexus (C8, T1 and T2), causing severe pain in the shoulder and down the inner surface of the arm (Pancoast's tumour). The sympathetic ganglion can also be involved, producing Horner's syndrome. Hilar tumours may involve the left recurrent laryngeal nerve, causing unilateral vocal cord paresis with hoarseness and a bovine cough.

Bronchial carcinoma can also directly invade the phrenic nerve, causing paralysis of the ipsilateral hemidiaphragm. It can involve the oesophagus, producing progressive dysphagia, and the pericardium, producing pericardial effusion and malignant dysrhythmias. *Superior vena caval obstruction* (p. 458) causes early morning headache, facial congestion and oedema involving the upper limbs; the jugular veins are distended, as are the veins on the chest that form a collateral circulation with veins arising from the abdomen.

Table 14.20 The frequency of the common presenting symptoms of bronchial carcinoma

Symptom	Frequency (%)
Cough	41
Chest pain	22
Cough and pain	15
Coughing blood	7
Chest infection	<5
Malaise	<5
Weight loss	<5
Shortness of breath	<5
Hoarseness	<5
Distant spread	<5
No symptoms	<5

Metastatic complications

Bony metastases are common, giving rise to severe pain and pathological fractures. There is frequent involvement of the liver. Secondary deposits in the brain present as a change in personality, epilepsy or as a focal neurological lesion. Spinal cord compression is not uncommon and requires urgent treatment (p. 458). Secondary deposits in the adrenal gland are frequently found at post mortem but are often asymptomatic.

Non-metastatic extrapulmonary manifestations

Clinical extrapulmonary manifestations are relatively rare apart from finger clubbing (Table 14.21); approximately 10% of small-cell tumours produce ectopic hormones, giving rise

Table 14.21 Non-metastatic extrapulmonary manifestations of bronchial carcinoma (percentage of all cases)

Metabolic (universal at some stage)
Loss of weight
Lassitude
Anorexia
Endocrine (10%) (usually small-cell carcinoma)
Ectopic adrenocorticotrophin syndrome
Syndrome of inappropriate secretion of antidiuretic hormone (SIADH)
Hypercalcaemia (usually squamous cell carcinoma)
Rarer: hypoglycaemia, thyrotoxicosis, gynaecomastia
Neurological (2–16%)
Encephalopathies – including subacute cerebellar degeneration
Myelopathies – motor neurone disease
Neuropathies – peripheral sensorimotor neuropathy
Muscular disorders – polymyopathy, myasthenic syndrome (Eaton–Lambert syndrome)
Vascular and haematological (rare)
Thrombophlebitis migrans
Non-bacterial thrombotic endocarditis
Microcytic and normocytic anaemia
Disseminated intravascular coagulopathy
Thrombotic thrombocytopenic purpura
Haemolytic anaemia
Skeletal
Clubbing (30%)
Hypertrophic osteoarthropathy (±gynaecomastia) (3%)
Cutaneous (rare)
Dermatomyositis
Acanthosis nigricans
Herpes zoster

to paraneoplastic syndromes (see Box 9.1). Some antitumour markers are shown in Box 9.1 on page 455.

Hypertrophic pulmonary osteoarthropathy (HPOA) (see p. 555) occurs in approximately 3% of all bronchial carcinomas, particularly squamous cell carcinomas and adenocarcinomas. Symptoms include joint stiffness and severe pain in the wrists and ankles, sometimes associated with gynaecomastia. X-rays show a characteristic proliferative periostitis at the distal ends of long bones, which have an onion-skin appearance. HPOA is invariably associated with clubbing of the fingers. It may regress after resection of the lung tumour or as a result of vagotomy at thoracotomy.

Investigations

Chest X-ray

By the time the lung cancer is causing symptoms, it will almost always be visible on chest X-ray. Asymptomatic tumours may be seen on chest X-ray if they are more than 1 cm in diameter. Lateral views are useful to assess the hilum and masses behind the heart. A minority of tumours are confined to the central airways and mediastinum without obvious change on the plain chest X-ray. These will be readily seen at bronchoscopy or on CT scanning. Although investigation of isolated haemoptysis with a normal chest X-ray is often negative, a normal chest X-ray should not deter from further investigation, especially in smokers over the age of 40. About 70% of all primary lung cancers present with a mass, including virtually all small-cell lung cancers and most squamous cell carcinomas. Adenocarcinoma occurs more often in the periphery than the other cell types.

Carcinomas causing partial obstruction of a bronchus interrupt the mucociliary escalator, and bacteria are retained within the affected lobe. This gives rise to the so-called secondary pneumonia that is commonly seen on the chest X-ray.

Bronchial carcinoma can also appear as round shadows on a chest X-ray (see p. 822). Characteristically the edge of the tumour has a fluffy or spiked appearance, though sometimes it may be entirely smooth with cavitation.

The hilar lymph nodes on the side of the tumour are frequently involved. Bronchial carcinoma is also a common cause of large pleural effusions. Carcinoma can spread through the lymphatic channels of the lung to give rise to lymphangitis carcinomatosa; in bronchial carcinoma this is usually unilateral and associated with striking dyspnoea. The chest X-ray shows streaky shadowing throughout the lung. Bilateral lymphangitis carcinomatosa is more often due to metastatic spread, usually from tumours below the diaphragm (the stomach and colon) or from breast cancers.

Computed tomography

CT is useful for detection of small tumours and for identifying disease in the mediastinum, such as enlarged lymph nodes (see Fig. 14.12, p. 823) or local spread of the tumour, and for identifying secondary spread of carcinoma to the opposite lung by detecting masses too small to be seen on the chest X-ray. CT is a poor guide to whether nodes are involved by tumour but a normal CT scan prior to surgery excludes the need for mediastinoscopy and node biopsy. CT scanning should include the liver, adrenal glands and the brain since these are common sites for metastases.

Other imaging modalities

MRI is not useful for the diagnosis of primary lung tumours. PET or PET-CT scanning is now the investigation of choice for assessment of the mediastinum and for possible metastases (see p. 823).

Fibreoptic bronchoscopy (see also p. 828, Fig. 14.44)

This technique is used to define the bronchial anatomy and to obtain biopsy and cytological specimens. If the carcinoma involves the first 2 cm of either main bronchus, the tumour is inoperable as there would be insufficient resection margins for pneumonectomy. Widening and loss of the sharp angle of the carina indicates the presence of enlarged mediastinal lymph nodes, either malignant or reactive. These can be biopsied by passage of a needle through the bronchial wall. Vocal cord paresis on the left indicates involvement of the recurrent laryngeal nerve and inoperability.

Percutaneous aspiration and biopsy

Peripheral lung lesions cannot be seen by fibreoptic bronchoscopy. Samples may be obtained by aspiration or biopsy through the chest wall under CT guidance. The commonest complication is pneumothorax (up to 25% patients), especially if the mass is deep in the lung, as opposed to lesions next to the parietal pleura. Intercostal drainage is occasionally required. Mild haemoptysis occurs in 5%. Implantation metastases do not occur.

Other investigations

These include full blood count for the detection of anaemia, and biochemistry for liver involvement, hypercalcaemia and hyponatraemia.

Treatment (see also p. 486)

Treatment of lung cancer involves several different modalities and is best planned by a multidisciplinary team. Treatment decisions need to reflect the poor overall survival rates: only 20% of patients are alive 1 year after diagnosis and only 6–8% after 5 years (cf. 50% for breast or cervix). Patients are staged according to the TNM classification (p. 456) for non-small-cell cancer but small-cell cancer is treated according to whether it is limited or extensive (p. 486).

Fig. 14.44 Bronchoscopic view of a bronchial carcinoma obstructing a large bronchus.

Surgery

Surgery can be curative in non-small-cell lung cancer (T1, N0, M0) but only 5–10% of all cases are suitable for resection; about 70% of these survive for 5 years. Surgery is rarely appropriate in patients over 65 years as the operative mortality rate exceeds the 5-year survival rate. Trial data suggest that neo-adjuvant chemotherapy may downstage tumours to render them operable and may also improve 5-year survival in patients whose tumours are operable at presentation. Adjuvant chemotherapy following resection of adenocarcinomas may also improve survival.

Preoperative assessment includes exclusion of metastatic disease by blood tests and imaging as described above. Lung function tests, including walking oximetry, are used to predict postoperative potential. An active life after pneumonectomy is unlikely if the gas transfer is reduced below 50%.

Radiation therapy for cure

In patients who are fit and who have a slowly growing squamous carcinoma, high-dose radiotherapy (65 Gy or 6500 rads) can produce good results. It is the treatment of choice if surgery is declined. Poor lung function is a relative contraindication for radiotherapy. Radiation pneumonitis (defined as an acute infiltrate precisely confined to the radiation area and occurring within 3 months of radiotherapy) develops in 10–15% of cases. Radiation fibrosis, a fibrotic change occurring within a year or so of radiotherapy and not precisely confined to the radiation area, occurs to some degree in all cases. These complications usually cause no problems.

Radiation treatment for symptoms

Bone pain, haemoptysis and superior vena cava obstruction respond favourably to irradiation in the short term.

Chemotherapy

This is discussed on page 486. Adjuvant chemotherapy with radiotherapy improves response rate and extends median survival in non-small-cell cancer.

Laser therapy, endobronchial irradiation and tracheobronchial stents

These techniques are used in the palliation of inoperable lung cancer in selected patients with tracheobronchial narrowing from intraluminal tumour or extrinsic compression causing disabling breathlessness, intractable cough and complications, including infection, haemoptysis and respiratory failure.

A neodymium-Yag (Nd-Yag) laser passed through a fibreoptic bronchoscope can be used to vaporize inoperable fungating intraluminal carcinoma involving short segments of trachea or main bronchus. Benign tumours, strictures and vascular lesions can also be treated effectively with immediate relief of symptoms.

Endobronchial irradiation (brachytherapy) is useful for the treatment of both intraluminal tumour and malignant extrinsic compression. A radioactive source is afterloaded into a catheter placed adjacent to the carcinoma under fibreoptic bronchoscope control. Radiation dose falls rapidly with distance from the source, minimizing damage to adjacent normal tissue. Reduction in endoscopically assessed tumour size occurs in 70–95% of cases.

Tracheobronchial stents made of silicone or as expandable metal springs are available for insertion into strictures caused by tumour or from external compression or when there is weakening and collapse of the tracheobronchial wall.

Palliative care (see p. 494)

Patients dying of cancer of the lung need attention to their overall well-being. Palliative care must not be ignored simply because the patient cannot be cured. Much can be done to make the patient's remaining life symptom-free and as active as possible. As compared to patients with fatal cancers at other sites, patients with lung cancer tend to remain relatively independent and pain-free, but die more rapidly once they reach the terminal phase.

Daily treatment with prednisolone (up to 15 mg daily) may improve appetite. Morphine or diamorphine is given regularly for pain, either in the form of sustained-release morphine sulphate tablets twice daily or else as regular elixirs or injections. Many patients benefit from a continuous subcutaneous injection of opiates given by a pump. Candidiasis and other infections in the mouth are common and must be looked for and treated. Patients taking opiates are frequently constipated, so regular laxatives should be prescribed. Short courses of palliative radiotherapy are helpful for bone pain, severe cough or haemoptysis.

Both the patient and the relatives may require counselling, a task that should be shared between the respiratory teams, the primary care team and the nurses, social workers, hospital chaplains and doctors, who make up the palliative care team.

FURTHER READING

Spira A, Ettinger DS. Multidisciplinary management of lung cancer. *New England Journal of Medicine* 2004; **350**: 379–392.

Tracheal carcinoma

Primary tumours of the trachea are rare – their incidence relative to laryngeal and bronchial tumours is 1 : 75 and 1 : 180 respectively. The majority are malignant and cause severe and rapidly progressive dyspnoea and stridor. Flow–volume curves show typical and dramatic reductions in inspiratory flow (extrathoracic tracheal tumours) (see p. 816). Diagnosis is confirmed by bronchoscopy. Rapid and effective destruction of tumour by laser provides temporary relief of symptoms. Radiotherapy is often given and occasionally surgery may be possible but the prognosis is very poor.

Secondary tumours

Metastases in the lung are very common and usually present as round shadows (1.5–3.0 cm diameter). They are usually detected on chest X-ray in patients already diagnosed as having carcinoma, but may be the first presentation. *Typical sites for the primary tumour include the kidney, prostate, breast, bone, gastrointestinal tract, cervix or ovary.*

Metastases nearly always develop in the parenchyma and are often relatively asymptomatic even when the chest X-ray shows extensive pulmonary metastases. Rarely metastases may develop within the bronchi, when they may present with haemoptysis.

Carcinoma, particularly of the stomach, pancreas and breast, can involve mediastinal glands and spread along the lymphatics of both lungs (lymphangitis carcinomatosa), leading to progressive and severe breathlessness. On the chest X-ray, bilateral lymphadenopathy is seen together with streaky basal shadowing fanning out over both lung fields.

Occasionally a pulmonary metastasis may be detected as a *solitary round shadow* on chest X-ray in an asymptomatic

patient. The most common primary tumour to do this is a renal cell carcinoma.

The differential diagnosis includes:

- primary bronchial carcinoma
- tuberculoma
- benign tumour of the lung
- hydatid cyst.

Single pulmonary metastases can be removed surgically but, as CT scans usually show the presence of small metastases undetected on chest X-ray, detailed imaging including PET scanning and assessment is essential before undertaking surgery.

SCREENING FOR LUNG CANCER

Screening programmes (yearly chest X-ray, 4-monthly sputum cytology and CT scanning) have been tried in high-risk groups but the success rate is limited, underlining the need for prevention.

BENIGN TUMOURS

Pulmonary hamartoma

This is the most common benign tumour of the lung and is usually seen on the X-ray as a very well-defined round lesion 1–2 cm in diameter in the periphery of the lung. Growth is extremely slow, but the tumour may reach several centimetres in diameter. Rarely it arises from a major bronchus and causes obstruction.

Bronchial carcinoid

This rare tumour resembles an intestinal carcinoid tumour and is locally invasive, eventually spreading to mediastinal lymph nodes and finally to distant organs. It is a highly vascular tumour that projects into the lumen of a major bronchus causing recurrent haemoptysis. It grows slowly and eventually blocks the bronchus, leading to lobar collapse. As foregut derivatives, bronchial carcinoids may produce ACTH but do not usually produce the 5-hydroxytryptamine that is seen in midgut or hindgut carcinoid tumours.

Cylindroma, chondroma and lipoma

These are extremely rare tumours that may grow in the bronchus or trachea, causing obstruction.

Tracheal tumours

Benign tumours include squamous papilloma, leiomyoma, haemangiomas and tumours of neurogenic origin.

DISORDERS OF THE CHEST WALL AND PLEURA

Trauma

Trauma to the thoracic wall can cause penetrating wounds and lead to pneumothorax or haemothorax.

Rib fractures

Rib fractures are caused by trauma or coughing (particularly in the elderly), and can occur in patients with osteoporosis. Pathological rib fractures are due to metastatic spread from carcinoma of the bronchus, breast, kidney, prostate or thyroid. Ribs can also become involved by a mesothelioma. Fractures may not be readily visible on a PA chest X-ray, so lateral X-rays and oblique views may be necessary.

Pain prevents adequate chest expansion and coughing, and this can lead to pneumonia.

Treatment is with adequate oral analgesia, by local infiltration or an intercostal nerve block.

Two fractures in one rib can lead to a flail segment with paradoxical movement, i.e. part of the chest wall moves inwards during inspiration. This can produce inefficient ventilation and may require intermittent positive-pressure ventilation, especially if several ribs are similarly affected.

Rupture of the trachea or a major bronchus

Rupture of the trachea or even a major bronchus can occur during deceleration injuries, leading to pneumothorax, surgical emphysema, pneumomediastinum and haemoptysis. Surgical emphysema is caused by air leaking into the subcutaneous connective tissue; this can also occur after the insertion of an intercostal drainage tube. A pneumomediastinum occurs when air leaks from the lung inside the parietal pleura and extends along the bronchial walls.

Rupture of the oesophagus (p. 256)

Rupture of the oesophagus leads to mediastinitis usually with mixed bacterial infections. This is a serious complication of external injury, endoscopic procedures, bougienage or necrotic carcinoma, and requires vigorous antibacterial chemotherapy.

Lung contusion

This causes widespread fluffy shadows on the chest X-ray owing to intrapulmonary haemorrhage. This may give rise to acute respiratory distress syndrome (see p. 917).

Kyphoscoliosis

Kyphoscoliosis may be congenital, owing to disease of the vertebrae such as tuberculosis or osteomalacia, or due to neuromuscular disease such as Friedreich's ataxia or poliomyelitis. The respiratory effects of severe kyphoscoliosis are often more pronounced than might be expected and respiratory failure and death often occur in the fourth or fifth decade. The abnormality should be corrected at an early stage if possible. Positive airway pressure ventilation delivered through a tightly fitting nasal mask is the treatment of choice for respiratory failure (see p. 916).

Ankylosing spondylitis (see also p. 532)

Limitation of chest wall movement is often well compensated by diaphragmatic movement, and so the respiratory effects of this disease are relatively mild. It is occasionally associated with upper lobe fibrosis.

Pectus excavatum and carinatum

Pectus excavatum causes few problems other than embarrassment about the deep vertical furrow in the chest, which can be corrected surgically. The heart is seen to lie well to the left on the chest X-ray. Pectus carinatum (pigeon chest)

is often the result of rickets but is rarely seen in the West. No treatment is required.

Pleurisy

This is the term used to describe pain arising from any disease of the pleura. The localized inflammation produces sharp localized pain, made worse on deep inspiration, coughing and occasionally on twisting and bending movements. Pleurisy occurs with pneumonia, pulmonary infarct and carcinoma. Rarer causes include rheumatoid arthritis and systemic lupus erythematosus.

Epidemic myalgia (Bornholm disease) is due to infection by Coxsackie B virus. This illness is common in young adults in the late summer and autumn and is characterized by an upper respiratory tract illness followed by pleuritic pain in the chest and upper abdomen with tender muscles. The chest X-ray remains normal and the illness clears within a week.

Mesothelioma (see p. 879)

This is usually associated with asbestos exposure. It is described with other pleural diseases caused by asbestosis in Table 14.18.

Pleural effusion

A pleural effusion is an excessive accumulation of fluid in the pleural space. It can be detected on X-ray when 300 mL or more of fluid is present and clinically when 500 mL or more is present. The chest X-ray appearances (Fig. 14.45) range from the obliteration of the costophrenic angle to dense homogeneous shadows occupying part or all of the hemithorax. Fluid below the lung (a subpulmonary effusion) can simulate a raised hemidiaphragm. Fluid in the fissures may resemble an intrapulmonary mass. The physical signs are shown in Table 14.1 (p. 821).

Diagnosis

This is by pleural aspiration (see p. 827) with ultrasound guidance if the effusion is small. The fluid that accumulates may be a transudate or an exudate.

Transudates

Effusions that are transudates can be bilateral, but are often larger on the right side. The protein content is less than 30 g/L and the lactic dehydrogenase is less than 200 IU/L and/or the fluid to serum ratio is <0.6. Causes include:

- heart failure
- hypoproteinaemia (e.g. nephrotic syndrome)
- constrictive pericarditis
- hypothyroidism
- ovarian tumours producing right-sided pleural effusion – Meigs' syndrome.

Exudates

The protein content of exudates is > 30 g/L and the lactic dehydrogenase is > 200 IU/L. Causes include:

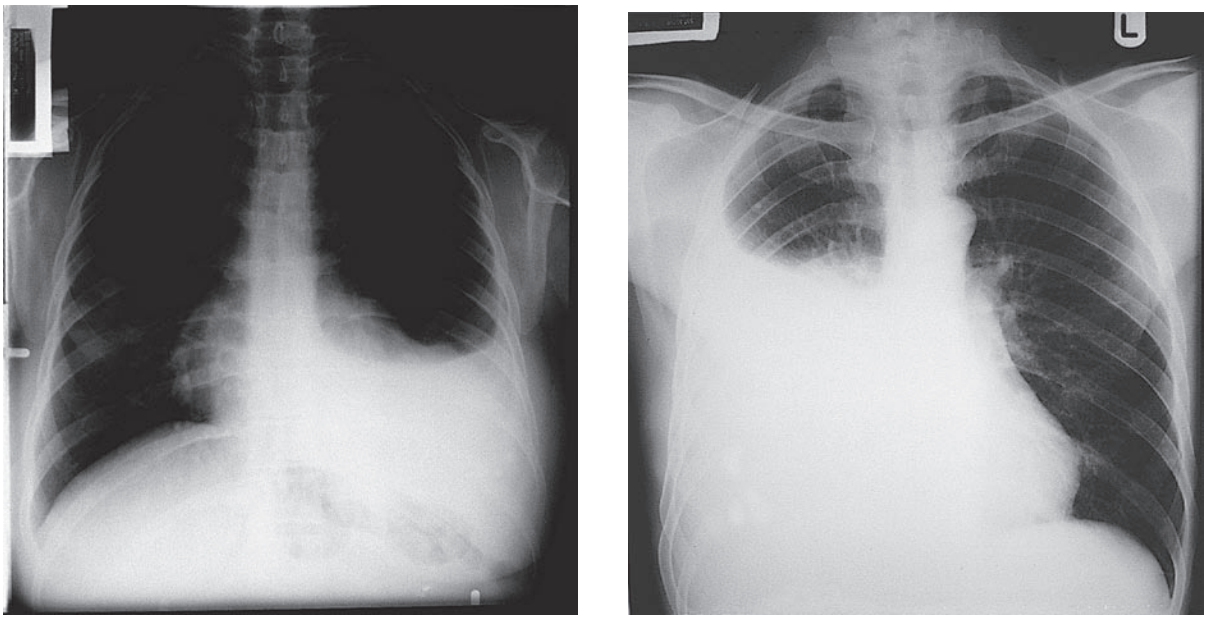
- bacterial pneumonia (common)
- carcinoma of the bronchus and pulmonary infarction – fluid may be blood-stained (common)
- tuberculosis
- connective tissue disease
- post-myocardial infarction syndrome (rare)
- acute pancreatitis (high amylase content) (rare)
- mesothelioma (rare)
- sarcoidosis (very rare)
- yellow-nail syndrome (effusion due to lymphoedema) (very rare)
- familial Mediterranean fever (rare).

Pleural biopsy (see p. 827) may be necessary if the diagnosis has not been established by simple aspiration.

FURTHER READING

British Thoracic Society. BTS guidelines for the management of pleural effusions. *Thorax* 2003; **58** (Suppl 2): ii8–ii38.

British Thoracic Society. BTS guidelines for the management of pleural disease. *Thorax* 2003; **58** (Suppl 2): ii39–ii52.



(a) (b)

Fig. 14.45 **Radiographs showing (a) small and (b) large pleural effusions.**

Treatment is of the underlying condition unless the fluid is purulent (empyema) in which case drainage is mandatory.

Management of malignant pleural effusions

Malignant pleural effusions that re-accumulate and are symptomatic can be aspirated to dryness followed by the instillation of a sclerosing agent such as tetracycline or bleomycin. Effusions should be drained slowly since rapid shift of the mediastinum causes severe pain and occasionally shock. This treatment produces only temporary relief.

Chylothorax

This is due to the accumulation of lymph in the pleural space, usually resulting from leakage from the thoracic duct following trauma or infiltration by carcinoma.

Empyema

This is the presence of pus in the pleural space and can be a complication of pneumonia (see p. 863).

Pneumothorax

'Pneumothorax' means air in the pleural space. It may be spontaneous or occur as a result of trauma to the chest. Spontaneous pneumothorax is commonest in young males, the male-to-female ratio being 6:1. It is caused by the rupture of a pleural bleb, usually apical, and is thought to be due to congenital defects in the connective tissue of the alveolar walls. Both lungs are affected with equal frequency. Often these patients are tall and thin. In patients over 40 years of age, the usual cause is underlying COPD. Rarer causes include bronchial asthma, carcinoma, a lung abscess breaking down and leading to bronchopleural fistula, and severe pulmonary fibrosis with cyst formation.

Pneumothorax may be localized if the visceral pleura has previously become adherent to the parietal pleura, or generalized if there are no pleural adhesions. Normally the pressure in the pleural space is negative but this is lost once a communication is made with atmospheric pressure; the elastic recoil pressure of the lung then causes it to partially deflate. If the communication between the airways and the pleural space remains open, a bronchopleural fistula is created. Once the communication between the lung and the pleural space is obliterated, air will be reabsorbed at a rate of 1.25% of the total radiographic volume of the hemithorax per day. Thus, a 50% collapse of the lung will take about 40 days to reabsorb completely once the air leak is closed.

It has been postulated that a valvular mechanism may develop through which air can be sucked into the pleural space during inspiration but not expelled during expiration. The intrapleural pressure remains positive throughout breathing, the lung deflates further, the mediastinum shifts, and venous return to the heart decreases, with increasing respiratory and cardiac embarrassment. This is called tension pneumothorax and is very rare unless the patient is on positive ventilation.

The usual presenting features are sudden onset of unilateral pleuritic pain or progressively increasing breathlessness. If the pneumothorax enlarges, the patient becomes more breathless and may develop pallor and tachycardia. There may be few physical signs if the pneumothorax is small.

Practical Box 14.6 Simple aspiration of pneumothorax

1. Explain the nature of the procedure and obtain consent.
2. Infiltrate 2% lidocaine down to the pleura in the second intercostal space in the mid-clavicular line.
3. Push a 3–4 cm 16 French gauge cannula through the pleura.
4. Connect the cannula to a three-way tap and 50 mL syringe.
5. Aspirate up to 2.5 L of air. Stop if resistance to suction is felt or the patient coughs excessively.
6. Repeat chest X-ray (in expiration) in the X-ray department.

The characteristic features and management are shown in Figure 14.46. The main aim is to get the patient back to active life as soon as possible. The procedure for simple aspiration is shown in Practical box 14.6.

Recurrence. A third of patients will have a recurrence. Chemical pleurodesis with talc is used for patients with contraindication for surgery. Bleb resection and pleurodesis are achieved using a video-assisted thoracoscopic (VATS) approach or by open thoracotomy which has no recurrence.

DISORDERS OF THE DIAPHRAGM

Diaphragmatic fatigue

The diaphragm can become fatigued if the force of contraction during inspiration exceeds 40% of the force it can develop in a maximal static effort. When this occurs acutely in patients with exacerbations of COPD or cystic fibrosis or in quadriplegics, positive-pressure ventilation is required. Further rehabilitation requires exercises to increase the strength and endurance of the diaphragm by breathing against a resistance for 30 minutes a day.

Unilateral diaphragmatic paralysis

This is common and symptomless. The affected diaphragm is usually elevated and moves paradoxically on inspiration. It can be diagnosed when a sniff causes the paralysed diaphragm to rise, the unaffected diaphragm to descend. Causes include:

- surgery
- carcinoma of the bronchus with involvement of the phrenic nerve
- neurological, including poliomyelitis, herpes zoster
- trauma to cervical spine, birth injury, subclavian vein puncture
- infection: tuberculosis, syphilis, pneumonia.

Bilateral diaphragmatic weakness or paralysis

This causes breathlessness in the supine position and is a cause of sleep apnoea leading to daytime headaches and somnolence. Tidal volume is decreased and respiratory rate increased. Vital capacity is substantially reduced when lying down, and sniffing causes a paradoxical inward movement of the abdominal wall best seen in the supine position. Causes include viral infections, multiple sclerosis, motor neurone disease, poliomyelitis, Guillain–Barré syndrome, quadriplegia after trauma, and rare muscle diseases. Treat-

www.studentconsult.com
Ask the authors

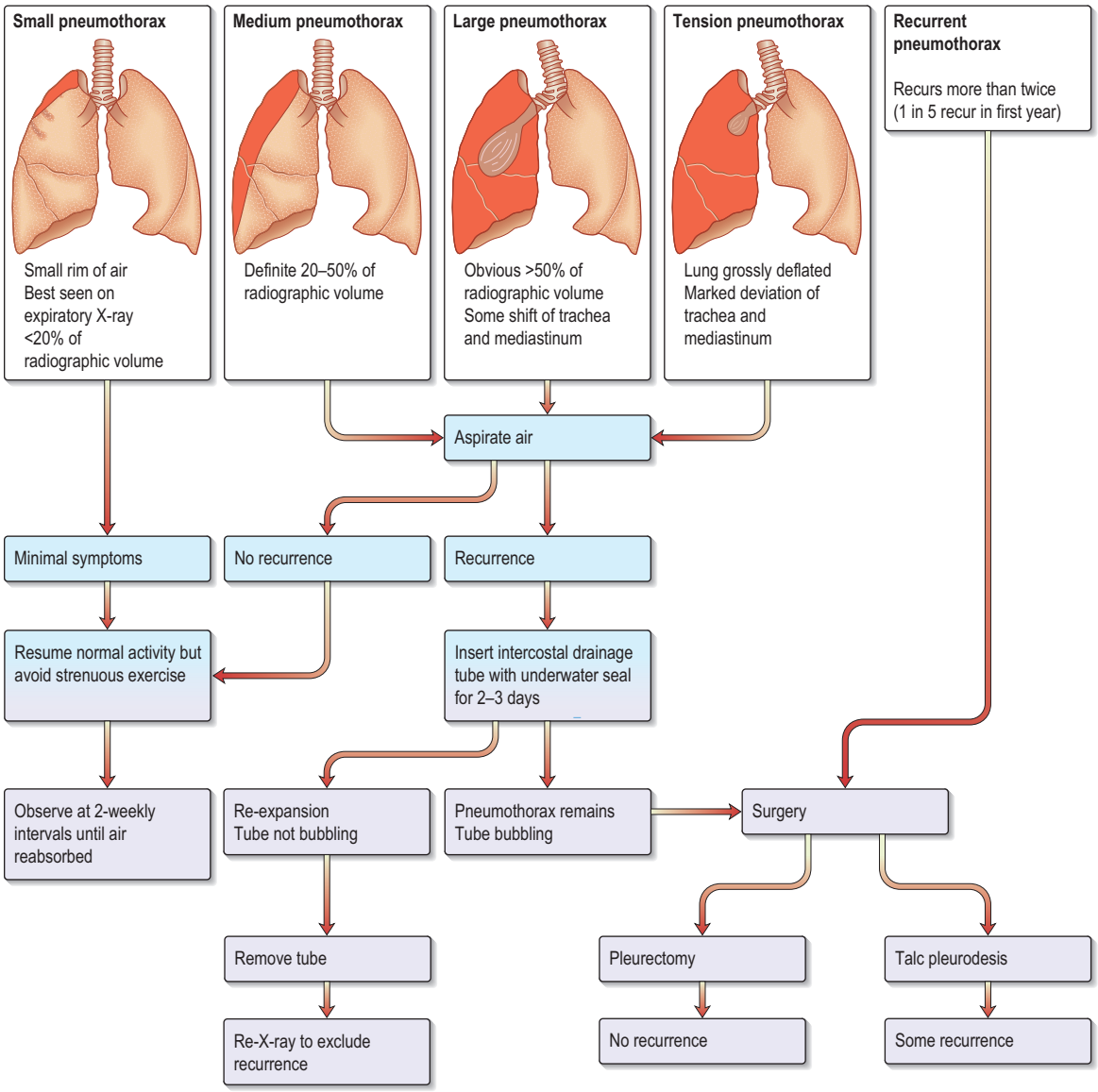


Fig. 14.46 **Pneumothorax: an algorithm for management.** VATS, video-assisted thoracoscopic approach.

ment is either diaphragmatic pacing or night-time assisted ventilation.

Complete eventration of the diaphragm

This is a congenital condition (invariably left-sided) in which muscle is replaced by fibrous tissue. It presents as marked elevation of the left hemidiaphragm, sometimes associated with gastrointestinal symptoms. Partial eventration, usually on the right, causes a hump (often anteriorly) on the diaphragmatic shadow on X-ray.

Diaphragmatic hernias

These are most commonly through the oesophageal hiatus, but occasionally occur anteriorly, through the foramen of Morgagni, posterolaterally through the foramen of Bochdalek, or at any site following traumatic tears.

Hiccups

Hiccups are due to involuntary diaphragmatic contractions with closure of the glottis and are extremely common. Occasionally patients present with persistent hiccups. This can be as a result of diaphragmatic irritation (e.g. subphrenic abscess) or a metabolic cause (e.g. uraemia). Treatment for persistent hiccups is with gabapentin 300 mg or pregabalin 50 mg three times daily. The underlying cause should be treated, if known.

MEDIASTINAL LESIONS

The mediastinum is defined as the region between the pleural sacs. It is additionally divided as shown in Figure 14.47. Tumours affecting the mediastinum are rare. Masses are detected very accurately on CT, as well as MR scan (Fig. 14.48).

Retrosternal or intrathoracic thyroid

The most common mediastinal tumour is a retrosternal or intrathoracic thyroid, which is nearly always an extension of

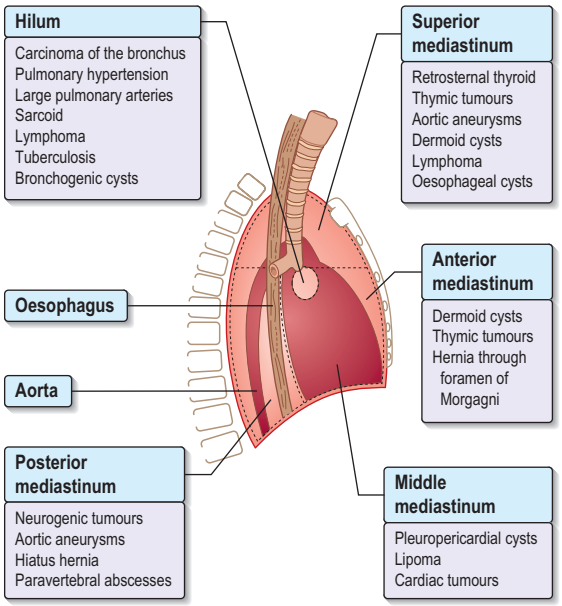


Fig. 14.47 **Subdivisions of the mediastinum and mass lesions.**

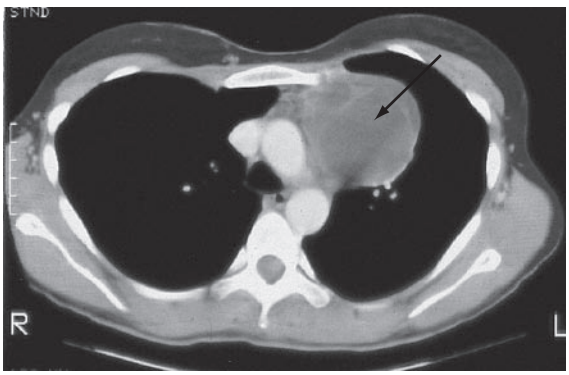


Fig. 14.48 **CT scan of a dermoid cyst in the mediastinum.**

the thyroid present in the neck. Enlargement of the thyroid by a colloid goitre, malignant disease or rarely, in thyrotoxicosis, can cause displacement of the trachea and oesophagus to the opposite side. Symptoms of compression develop insidiously before producing the cardinal feature of dyspnoea. Very occasionally an intrathoracic thyroid may be the cause of dysphagia and, rarely, of hoarseness of the voice and vocal cord paralysis from stretching of the recurrent laryngeal nerve. The treatment is surgical removal.

Thymic tumours (thymomas)

The thymus is large in childhood and occupies the superior and anterior mediastinum. It involutes with age but may be enlarged by cysts, which are rarely symptomatic, or by tumours, which may cause myasthenia gravis or lead to compression of the trachea or, rarely, the oesophagus. Surgery is the treatment of choice. Approximately half of the patients presenting with a thymic tumour have myasthenia gravis. Good's syndrome, a combined defect in humoral and cellular immunity, is seen in 10% of thymomas.

Pleuropericardial cysts

These cysts, which may be up to 10 cm in diameter, are filled with clear fluid and are usually situated anteriorly in the cardiophrenic angle on the right in 70% of cases. Infection only rarely occurs; malignant change does not occur. The diagnosis is usually made by needle aspiration. No treatment is required, but these patients should be followed up as an increase in cyst size suggests an alternative pathology; surgical excision is then advisable.

SIGNIFICANT WEBSITES

http://www.brit-thoracic.org.uk
British Thoracic Society

http://www.thoracic.org
American Thoracic Society

http://www.asthma.org.uk
UK National Asthma Campaign

http://www.quitsmoking.uk.com
Good site for those wanting to quit or to help patients to quit.

15 Intensive care medicine

General aspects of intensive care management 890

Applied cardiorespiratory physiology 890
Oxygen delivery and consumption 890
Disturbances of acid–base balance 893

Shock and acute disturbances of haemodynamic function 894
Pathophysiology 894
Clinical features of shock 898
Monitoring critically ill patients 899
Management of shock 904

Renal failure (acute kidney injury) 910

Respiratory failure 910
Monitoring of respiratory failure 911
Management of respiratory failure 911

Acute lung injury/acute respiratory distress syndrome 917
Pathogenesis and pathophysiology of ALI/ARDS 917
Clinical presentation of ALI/ARDS 918
Management of ALI/ARDS 918

Patient selection – withholding and withdrawing treatment 919

Brain death 920

Intensive care medicine (or 'critical care medicine') is concerned predominantly with the management of patients with acute life-threatening conditions ('the critically ill') in specialized units. As well as emergency cases, such units admit high-risk patients electively after major surgery (Table 15.1). Intensive care medicine also encompasses the resuscitation and transport of those who become acutely ill, or are injured in the community. Management of seriously ill patients throughout the hospital (e.g. in coronary care units, acute admissions wards, postoperative recovery areas, or emergency units), including critically ill patients who have been discharged to the ward ('outreach care') is also undertaken. Teamwork and a multidisciplinary approach are central to the provision of intensive care and are most effective when directed and coordinated by committed specialists. An example of guidelines for involving the critical care team in the management of an acutely ill patient is shown in Box 15.1.

Intensive care units (ICUs) are usually reserved for patients with established or potential organ failure and provide facilities for the diagnosis, prevention and treatment of multiple organ dysfunction. They are fully equipped with monitoring and technical facilities, including an adjacent laboratory (or 'near patient testing' devices) for the rapid determination of blood gases and simple biochemical data such as serum potassium, blood glucose and blood lactate levels. Patients receive continuous expert nursing care and the constant attention of appropriately trained medical staff. High dependency units (HDUs) offer a level of care intermediate between that available on the general ward and that provided in an ICU. They provide monitoring and support for patients with acute (or acute-on-chronic) single organ failure and for those who are at risk of developing multiple organ failure. These units are a comfortable environment for less severely ill patients who are often conscious and alert. They can also

Table 15.1 Some common indications for admission to intensive care

Surgical emergencies

Acute intra-abdominal catastrophe

Perforated viscus, especially with faecal soiling of peritoneum (often complicated by sepsis/septic shock)

Ruptured/leaking abdominal aortic aneurysm

Trauma (often complicated by hypovolaemic and later sepsis/septic shock)

- Multiple injuries
- Massive blood loss
- Severe head injury

Medical emergencies

Respiratory failure

Exacerbation of chronic obstructive pulmonary disease (COPD)

Acute severe asthma

Severe pneumonia (may be complicated by sepsis/septic shock)

Meningococcal infection

Status epilepticus

Severe diabetic ketoacidosis

Coma

Elective surgical

Extensive/prolonged procedure (e.g. oesophagogastrectomy)

Cardiothoracic surgery

Major head and neck surgery

Coexisting cardiovascular or respiratory disease

Obstetric emergencies

Severe pre-eclampsia/eclampsia

Haemorrhage

Amniotic fluid embolism

Box 15.1 Guidance for involving a Medical Emergency or 'Patient at Risk' team

This system should be activated if the following criteria are fulfilled and the patient is being actively treated:

- Heart rate below 40 b.p.m.
- Heart rate above 120 b.p.m.
- Systolic blood pressure above 200 mmHg
- Systolic blood pressure below 80 mmHg
- Urine output less than 0.5 mL/kg/h for 2 consecutive hours
- Respiratory rate above 30 breaths per minute
- Respiratory rate below 8 breaths per minute
- Oxygen saturation less than 90% whilst receiving supplemental oxygen
- Glasgow coma scale less than 8
- Core temperature greater than 39°C
- Core temperature less than 35°C

Modified from Lee A et al. The medical emergency team. *Anaesthesia and Intensive Care* 1995; **23**: 183–186.

provide a 'step-down' facility for patients being discharged from intensive care.

The provision of staff and the level of technical support must match the needs of the individual patient and resources are used more efficiently when they are combined in a single critical care facility rather than being divided between physically and managerially separate units.

In the UK only around 2.6% of hospital beds are designated for intensive care, but elsewhere in the developed world the proportion is often much higher.

GENERAL ASPECTS OF INTENSIVE CARE MANAGEMENT

Critically ill patients require multidisciplinary care with:

- Intensive skilled nursing care (usually 1:1 or 1:2 nurse/patient ratio in the UK).
- Regular physiotherapy.
- Careful management of pain and distress with analgesics and sedation as necessary.
- Constant reassurance and support (critically ill patients easily become disorientated and psychologically disturbed).
- Nutritional support (enteral nutrition should always be used if possible). Studies have shown pharmacological doses of various nutrients to have positive immunomodulatory effects (immunonutrition) including glutamine, polyunsaturated fatty acids and nucleotides. The efficacy and safety of this approach in critically ill patients remains in doubt.
- The use of insulin infusions (if necessary in high doses) to maintain blood glucose levels <8.3 mmol/L is recommended. Hypoglycaemia must be avoided.
- H_2-receptor antagonists or proton pump inhibitors in selected cases to prevent stress-induced ulceration.
- Compression stockings (full-length and graduated) and subcutaneous low-molecular-weight heparin to prevent venous thrombosis.
- Care of the mouth, prevention of constipation and of pressure sores.

In many critically ill patients the precise underlying diagnosis is initially unclear, but in all cases the *immediate objective* is to preserve life and prevent, reverse or minimize damage to vital organs such as the lungs, brain, kidneys and liver. This involves a rapid assessment of the physiological derangement followed by prompt institution of measures to support cardiovascular and respiratory function in order to restore perfusion of vital organs, improve delivery of oxygen to the tissues and encourage the removal of carbon dioxide and other waste products of metabolism. The patient's condition and response to treatment should be closely monitored throughout. The underlying diagnosis can then be established as the results of investigations become available, a more detailed history is obtained and a more thorough physical examination is performed.

Discharge of patients from intensive care should normally be planned in advance and should ideally take place during normal working hours. Planned discharge may involve a period in a 'step-down' intermediate care area. Premature or unplanned discharge, especially during the night, has been associated with higher hospital mortality rates. A summary including 'points to review' should be included in the clinical notes and there should be a detailed handover to the receiving team (medical and nursing). The intensive care team should continue to review the patient, who may deteriorate following discharge, on the ward and should be available at all times for advice on further management (e.g. tracheostomy care, nutritional support). In this way deterioration and readmission to intensive care (which is associated with a particularly poor outcome) or even cardiorespiratory arrest may be avoided.

This chapter concentrates on cardiovascular and respiratory problems. Many patients also have failure of other organs such as the kidney and liver; treatment of these is dealt with in more detail in the appropriate chapters.

FURTHER READING

Griffiths RD. Specialized nutrition support in critically ill patients. *Current Opinion in Critical Care* 2003; **9:** 249–259.

APPLIED CARDIORESPIRATORY PHYSIOLOGY

OXYGEN DELIVERY AND CONSUMPTION
(Fig. 15.1)

Oxygen delivery (DO_2, oxygen dispatch) is defined as the total amount of oxygen delivered to the tissues per unit time. It is dependent on the volume of blood flowing through the microcirculation per minute (i.e. the total cardiac output, $\dot{Q}_t$) and the amount of oxygen contained in that blood (i.e. the arterial oxygen content, C_aO_2). Oxygen is transported both in combination with haemoglobin and dissolved in plasma. The amount combined with haemoglobin is determined by the oxygen capacity of the haemoglobin (usually taken as 1.34 mL of oxygen per gram of haemoglobin) and its percentage saturation with oxygen (SO_2), while the volume dissolved in plasma depends on the partial pressure of oxygen (PO_2). Except when hyperbaric oxygen is administered, the amount of dissolved oxygen in plasma is insignificant.

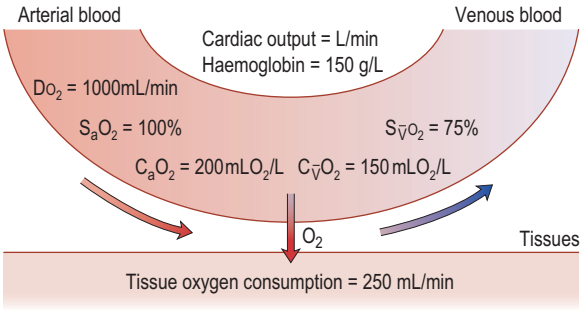


Fig. 15.1 **Tissue oxygen delivery and consumption in a normal 70 kg person breathing air.** Oxygen delivery (Do_2) = cardiac output × (haemoglobin concentration × oxygen saturation (S_aO_2) × 1.34). In normal adults oxygen delivery is roughly 1000 mL/min, of which 250 mL is taken up by tissues. Mixed venous blood is thus 75% saturated with oxygen. $C_{\bar{v}}o_2$, mixed venous oxygen content; $S_{\bar{v}}o_2$, mixed venous oxygen saturation; CaO_2, arterial oxygen content. From Singer M, Grant I (eds). ABC of Intensive Care. London: BMJ books, 1999, with permission.

Clinically, however, the utility of this global concept of oxygen dispatch is limited because it fails to account for changes in the relative flow to individual organs and its distribution through the microcirculation (i.e. the efficiency with which oxygen delivery is matched to the metabolic requirements of individual tissues or cells). Furthermore, some organs (such as the heart) have high oxygen requirements relative to their blood flow and may receive insufficient oxygen, even if the overall oxygen delivery is apparently adequate. Lastly microcirculatory flow is influenced by blood viscosity.

Cardiac output

Cardiac output is the product of heart rate and stroke volume, and is affected by changes in either (Fig. 15.2).

Heart rate

When heart rate increases, the duration of systole remains essentially unchanged, whereas diastole, and thus the time available for ventricular filling, becomes progressively shorter, and the stroke volume eventually falls. In the normal heart this occurs at rates greater than about 160 beats per minute, but in those with cardiac pathology, especially when this restricts ventricular filling (e.g. mitral stenosis), stroke volume may fall at much lower heart rates. Furthermore, tachycardias cause a marked increase in myocardial oxygen consumption ($\dot{V}_mo_2$) and this may precipitate ischaemia in areas of the myocardium with restricted coronary perfusion. When the heart rate falls, a point is reached at which the increase in stroke volume is insufficient to compensate for bradycardia and again cardiac output falls.

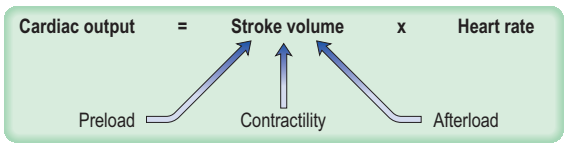


Fig. 15.2 **The determinants of cardiac output.**

Alterations in heart rate are often caused by disturbances of rhythm (e.g. atrial fibrillation, complete heart block) in which ventricular filling is not augmented by atrial contraction and stroke volume therefore falls.

Stroke volume

The volume of blood ejected by the ventricle in a single contraction is the difference between the ventricular end-diastolic volume (VEDV) and end-systolic volume (VESV) (i.e. stroke volume = VEDV − VESV). The ejection fraction describes the stroke volume as a percentage of VEDV (i.e. ejection fraction = (VEDV − VESV)/VEDV × 100%) and is an indicator of myocardial performance.

Three interdependent factors determine the stroke volume (see p. 683).

Preload

This is defined as the tension of the myocardial fibres at the end of diastole, just before the onset of ventricular contraction, and is therefore related to the degree of stretch of the fibres. As the end-diastolic volume of the ventricle increases, tension in the myocardial fibres is increased and stroke volume rises (Fig. 15.3). Myocardial oxygen consumption ($\dot{V}_mo_2$) increases only slightly with an increase in preload and this is therefore the most efficient way of improving cardiac output.

Myocardial contractility

This refers to the ability of the heart to perform work, independent of changes in preload and afterload. The state of myocardial contractility determines the response of the ventricles to changes in preload and afterload. Contractility is often reduced in intensive care patients, as a result of either pre-existing myocardial damage (e.g. ischaemic heart disease), or the acute disease process itself (e.g. sepsis). Changes in myocardial contractility alter the slope and position of the Starling curve; worsening ventricular performance is manifested as a depressed, flattened curve for stroke volume (Fig. 15.3 and Fig. 13.5).

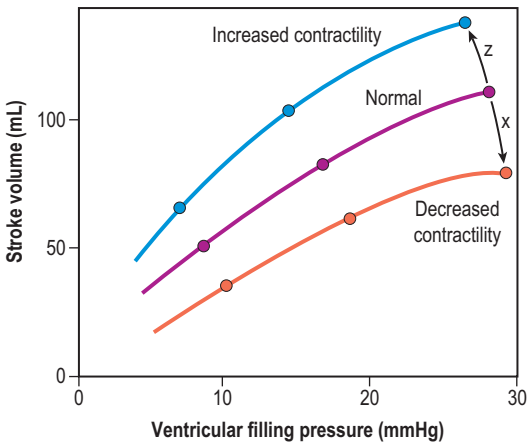


Fig. 15.3 **The Frank–Starling relationship**: as preload is increased, stroke volume rises. If the ventricle is overstretched, stroke volume will fall (x). In myocardial failure, the curve is depressed and flattened. Increasing contractility, e.g. due to sympathetic stimulation, shifts the curve upwards and to the left (z).

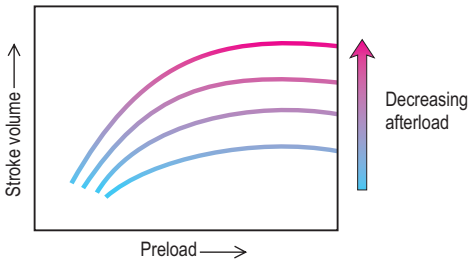


Fig. 15.4 **The effect of changes in afterload on the ventricular function curve.** At any given preload, decreasing afterload increases the stroke volume.

Afterload

This is defined as the myocardial wall tension developed during systolic ejection. In the case of the left ventricle, the resistance imposed by the aortic valve, the peripheral vascular resistance and the elasticity of the major blood vessels are the major determinants of afterload. Ventricular wall tension will also be increased by ventricular dilatation, an increase in intraventricular pressure or a reduction in ventricular wall thickness.

Decreasing the afterload can increase the stroke volume achieved at a given preload (Fig. 15.4), whilst reducing $\dot{V}_{m}o_2$. The reduction in wall tension may also lead to an increase in coronary blood flow, thereby improving the myocardial oxygen supply/demand ratio. Excessive reductions in afterload will cause hypotension.

Increasing the afterload, on the other hand, can cause a fall in stroke volume and is a potent cause of increased $\dot{V}_{m}o_2$. Right ventricular afterload is normally negligible because the resistance of the pulmonary circulation is very low.

Oxygenation of the blood

Oxyhaemoglobin dissociation curve

The saturation of haemoglobin with oxygen is determined by the partial pressure of oxygen (PO_2) in the blood, the relationship between the two being described by the oxyhaemoglobin dissociation curve (Fig. 15.5). The sigmoid shape of this curve is significant for a number of reasons:

- Modest falls in the partial pressure of oxygen in the arterial blood (P_aO_2) may be tolerated (since oxygen content is relatively unaffected) provided that the percentage saturation remains above 90%.
- Increasing the P_aO_2 to above normal has only a minimal effect on oxygen content unless hyperbaric oxygen is administered (when the amount of oxygen in solution in plasma becomes significant).
- Once on the steep 'slippery slope' of the curve (percentage saturation below about 90%), a small decrease in P_aO_2 can cause large falls in oxygen content, while increasing P_aO_2 only slightly, e.g. by administering 28% oxygen to a patient with chronic obstructive pulmonary disease (COPD), can lead to a useful increase in oxygen saturation and content.

The P_aO_2 is in turn influenced by the alveolar oxygen tension (P_AO_2), the efficiency of pulmonary gas exchange, and the partial pressure of oxygen in mixed venous blood ($P_{\bar{v}}o_2$).

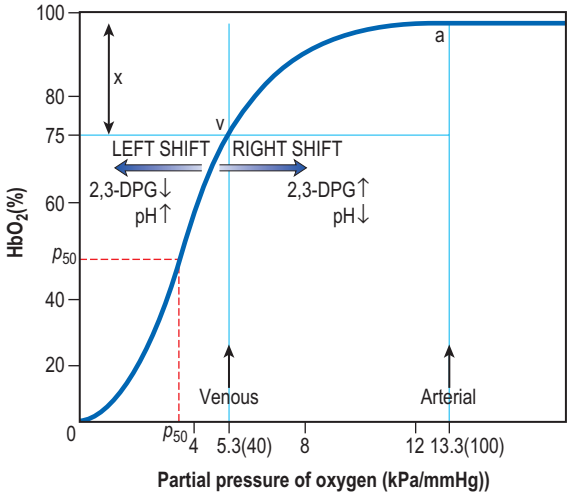


Fig. 15.5 **The oxyhaemoglobin dissociation curve.** a, arterial point; v, venous point; x, arteriovenous oxygen content difference. HbO_2 (%) is the percentage saturation of haemoglobin with oxygen. The curve will move to the right in the presence of acidosis (metabolic or respiratory), pyrexia or an increased red cell 2,3-DPG concentration. For a given arteriovenous oxygen content difference, the mixed venous PO_2 will then be higher. Furthermore, if the mixed venous PO_2 is unchanged, the arteriovenous oxygen content difference increases and more oxygen is offloaded to the tissues (see p. 391).

P_{50} (the PO_2 at which haemoglobin is half saturated with O_2) is a useful index of these shifts – the higher the P_{50} (i.e. shift to the right), the lower the affinity of haemoglobin for O_2.

Alveolar oxygen tension (P_AO_2)

The partial pressures of inspired gases are shown in Figure 15.6. By the time the inspired gases reach the alveoli they are fully saturated with water vapour at body temperature (37°C), which has a partial pressure of 6.3 kPa (47 mmHg) which contains CO_2 at a partial pressure of approximately 5.3 kPa (40 mmHg); the P_AO_2 is thereby reduced to approximately 13.4 kPa (100 mmHg).

The clinician can influence P_AO_2 by administering oxygen or by increasing the barometric pressure. Because of the reciprocal relationship between the partial pressures of oxygen and carbon dioxide in the alveoli, a small increase in P_AO_2 can be produced by lowering the P_ACO_2 (e.g. by using mechanical ventilation).

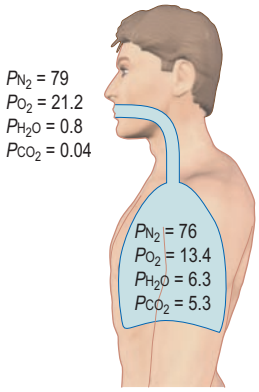


Fig. 15.6 **The composition of inspired and alveolar gas** (partial pressures in kPa).

Pulmonary gas exchange

In *normal* subjects there is a small alveolar–arterial oxygen difference ($P_{A-a}O_2$). This is due to:

- a small (0.133 kPa, 1 mmHg) pressure gradient across the alveolar membrane
- a small amount of blood (2% of total cardiac output) bypassing the lungs via the bronchial and thebesian veins
- a small degree of ventilation/perfusion mismatch.

Pathologically there are three possible causes of an increased $P_{A-a}O_2$ difference, as follows:

- **Diffusion defect**. This is not a major cause of hypoxaemia even in conditions such as lung fibrosis, in which the alveolar–capillary membrane is considerably thickened. Carbon dioxide is also not affected, as it is more soluble than oxygen.
- **Right-to-left shunts**. In certain congenital cardiac lesions, such as Fallot's tetralogy and when a segment of lung is completely unventilated, a significant amount of blood bypasses the lungs and causes arterial hypoxaemia. This hypoxaemia cannot be corrected by administering oxygen to increase the P_AO_2, because blood leaving normal alveoli is already fully saturated and further increases in PO_2 will not significantly affect its oxygen content. On the other hand, because of the shape of the carbon dioxide dissociation curve (Fig. 15.7), the high PCO_2 of the shunted blood can be compensated for by overventilating patent alveoli, thus lowering the CO_2 content of the effluent blood. Indeed, many patients with acute right-to-left shunts hyperventilate in response to the hypoxia and/or to stimulation of mechanoreceptors in the lung, so that the P_aCO_2 is normal or low.
- **Ventilation/perfusion mismatch** (see p. 817). Diseases of the lung parenchyma result in $\dot{V}/\dot{Q}$ mismatch, producing an increase in alveolar deadspace and hypoxaemia. The increased deadspace can be compensated by increasing overall ventilation. In contrast to the hypoxia resulting from a true right-to-left shunt, that due to areas of low $\dot{V}/\dot{Q}$ can be partially corrected by administering oxygen and thereby increasing the P_AO_2 even in poorly ventilated areas of lung.

Mixed venous oxygen tension ($P_{\bar{v}}O_2$) and saturation ($S_{\bar{v}}O_2$)

The $P_{\bar{v}}O_2$ is the partial pressure of oxygen in pulmonary arterial blood that has been thoroughly mixed during its passage through the right heart. Assuming P_aO_2 remains constant, $P_{\bar{v}}O_2$ and $S_{\bar{v}}O_2$ will fall if more oxygen has to be extracted from each unit volume of blood arriving at the tissues. A low $P_{\bar{v}}O_2$ therefore indicates either that oxygen delivery has fallen or that tissue oxygen requirements have increased without a compensatory rise in cardiac output. If $P_{\bar{v}}O_2$ falls, the effect of a given degree of pulmonary shunting on arterial oxygenation will be exacerbated. Thus, worsening arterial hypoxaemia does not necessarily indicate a deterioration in pulmonary function but may instead reflect a fall in cardiac output and/or a rise in oxygen consumption.

Conversely, a rise in $P_{\bar{v}}O_2$ and $S_{\bar{v}}O_2$ may reflect impaired tissue oxygen extraction (due to microcirculatory dysfunction) and/or a reduced oxygen utilization (due, for example, to mitochondrial dysfunction) as may be seen in severe sepsis (see below).

Monitoring the oxygen saturation in central venous ($S_{c\bar{v}}O_2$), rather than pulmonary artery blood is less invasive and has been shown to be a valuable guide to the resuscitation of critically ill patients (see p. 899).

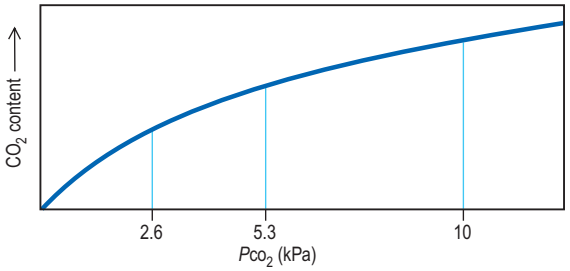


Fig. 15.7 **The carbon dioxide dissociation curve.** Note that in the physiological range the curve is essentially linear.

DISTURBANCES OF ACID–BASE BALANCE

The physiology of acid–base control is discussed on page 672. Acid–base disturbances can be described in relation to the diagram illustrated in Figure 12.14 (p. 675) (which shows P_aCO_2 plotted against arterial $[H^+]$).

Both acidosis and alkalosis can occur, each of which may be either metabolic (primarily affecting the bicarbonate component of the system) or respiratory (primarily affecting P_aCO_2). Compensatory changes may also be apparent. In clinical practice, arterial $[H^+]$ values outside the range 18–126 nmol/L (pH 6.9–7.7) are rarely encountered.

Blood gas and acid–base values (normal ranges) are shown in Table 15.2. For blood gas analysis, see page 911.

Respiratory acidosis. This is caused by retention of carbon dioxide. The P_aCO_2 and $[H^+]$ rise. A chronically raised P_aCO_2 is compensated by renal retention of bicarbonate, and the $[H^+]$ returns towards normal. A constant arterial bicarbonate concentration is then usually established within 2–5 days. This represents a primary respiratory acidosis with a compensatory metabolic alkalosis (see p. 678). Common causes of respiratory acidosis include ventilatory failure and COPD (type II respiratory failure where there is a high P_aCO_2 and a low P_aO_2 – see p. 837).

Respiratory alkalosis. In this case the reverse occurs and there is a fall in P_aCO_2 and $[H^+]$, often with a small reduction

Table 15.2 Arterial blood gas and acid–base values (normal ranges)

H^+	35–45 nmol/L	pH 7.35–7.45
PO_2 (breathing room air)	10.6–13.3 kPa	(80–100 mmHg)
PCO_2	4.8–6.1 kPa	(36–46 mmHg)
Base deficit	±2.5	
Plasma HCO_3^-	22–26 mmol/L	
O_2 saturation	95–100%	

in bicarbonate concentration. If hypocarbia persists, some degree of renal compensation may occur, producing a metabolic acidosis, although in practice this is unusual. A respiratory alkalosis is often produced, intentionally or unintentionally, when patients are mechanically ventilated; it may also be seen with hypoxaemic (type I) respiratory failure (see Ch. 14), spontaneous hyperventilation and in those living at high altitudes.

Metabolic acidosis (p. 675). This may be due to excessive acid production, most commonly lactate and H^+ (lactic acidosis) as a consequence of anaerobic metabolism during an episode of shock or following cardiac arrest. A metabolic acidosis may also develop in chronic renal failure or in diabetic ketoacidosis. It can also follow the loss of bicarbonate from the gut, for example, or from the kidney in renal tubular acidosis. Respiratory compensation for a metabolic acidosis is usually slightly delayed because the blood–brain barrier initially prevents the respiratory centre from sensing the increased blood [H^+]. Following this short delay, however, the patient hyperventilates and 'blows off' carbon dioxide to produce a compensatory respiratory alkalosis. There is a limit to this respiratory compensation, since in practice values for P_aCO_2 less than about 1.4 kPa (11 mmHg) are rarely achieved. Spontaneous respiratory compensation cannot occur if the patient's ventilation is controlled or if the respiratory centre is depressed, for example by drugs or head injury.

Metabolic alkalosis. This can be caused by loss of acid, for example from the stomach with nasogastric suction, or in high intestinal obstruction, or excessive administration of absorbable alkali. Overzealous treatment with intravenous sodium bicarbonate is sometimes implicated. Respiratory compensation for a metabolic alkalosis is often slight, and it is rare to encounter a P_aCO_2 above 6.5 kPa (50 mmHg), even with severe alkalosis.

SHOCK AND ACUTE DISTURBANCES OF HAEMODYNAMIC FUNCTION

Shock is the term used to describe acute circulatory failure with inadequate or inappropriately distributed tissue perfusion resulting in generalized cellular hypoxia and/or an inability of the cells to utilize oxygen.

Causes of shock

Abnormalities of tissue perfusion may result from:

- failure of the heart to act as an effective pump
- mechanical impediments to forward flow
- loss of circulatory volume
- abnormalities of the peripheral circulation.

The causes of shock are shown in Table 15.3. Often shock can result from a combination of these factors (e.g. in sepsis, distributive shock is frequently complicated by hypovolaemia and myocardial depression).

Table 15.3 Causes of shock

Hypovolaemic
Exogenous losses (e.g. haemorrhage, burns)
Endogenous losses

Cardiogenic
(e.g. ischaemic myocardial injury)

Obstructive
Obstruction to outflow (e.g. pulmonary embolus)
Restricted cardiac filling (e.g. cardiac tamponade, tension pneumothorax)

Distributive
(e.g. sepsis. anaphylaxis)
Vascular dilatation
Sequestration
Arteriovenous shunting
Maldistribution of flow
Myocardial depression

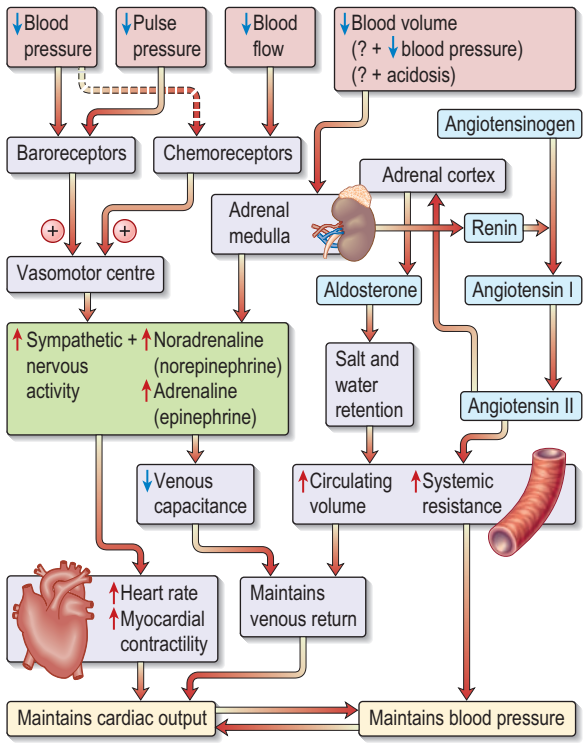


Fig. 15.8 **The sympatho-adrenal response to shock** showing the effect of increased catecholamines on the left of the diagram and the release of angiotensin and aldosterone on the right. Both mechanisms result in maintaining the cardiac output in shock.

PATHOPHYSIOLOGY

The sympatho-adrenal response to shock (Fig. 15.8)

Hypotension stimulates the baroreceptors, and to a lesser extent the chemoreceptors, causing increased sympathetic nervous activity with 'spill-over' of norepinephrine (noradrenaline) into the circulation. Later this is augmented by the release of catecholamines (predominantly epinephrine (adrenaline)) from the adrenal medulla. The resulting vasoconstriction, together with increased myocardial contractility and heart rate, help to restore blood pressure and cardiac output.

Reduction in perfusion of the renal cortex stimulates the juxtaglomerular apparatus to release renin. This converts angiotensinogen to angiotensin I, which in turn is converted in the lungs and by the vascular endothelium to the potent vasoconstrictor angiotensin II. Angiotensin II also stimulates secretion of aldosterone by the adrenal cortex, causing sodium and water retention (p. 1023). This helps to restore the circulating volume (see p. 652).

Neuroendocrine response

- There is *release of pituitary hormones* such as adrenocorticotrophic hormone (ACTH), vasopressin (antidiuretic hormone, ADH) and endogenous opioid peptides. (In septic shock there may be a relative deficiency of vasopressin.)
- There is *release of cortisol*, which causes fluid retention and antagonizes insulin.
- There is *release of glucagon*, which raises the blood sugar level.

Although absolute adrenocortical insufficiency (due, for example, to bilateral adrenal haemorrhage or necrosis) is rare, there is evidence that patients with septic shock have a blunted response to exogenous ACTH (so-called 'relative' or 'occult' adrenocortical insufficiency) and that this may be associated with an impaired pressor response to norepinephrine (noradrenaline) and a worse prognosis. The diagnosis, causes and clinical significance of this phenomenon remain unclear.

Release of pro- and anti-inflammatory mediators (see also Ch. 3)

Severe infection (often with bacteraemia or endotoxaemia), the presence of large areas of damaged tissue (e.g. following trauma or extensive surgery) or prolonged/repeated episodes of hypoperfusion can trigger an exaggerated inflammatory response with systemic activation of leucocytes and release of a variety of potentially damaging 'mediators'. Although beneficial when targeted against local areas of infection or necrotic tissue, dissemination of this 'innate immune' response can produce shock and widespread tissue damage. Characteristically the initial episode of overwhelming inflammation is followed by a period of immune suppression, which in some cases may be profound and during which the patient is at increased risk of developing secondary infections. It also seems that pro- and anti-inflammatory elements of the host response may coexist.

Microorganisms and their toxic products (Fig. 15.9)

In sepsis/septic shock the innate immune response and inflammatory cascade are triggered by the recognition of pathogen-associated molecular patterns (PAMPs) including cell wall components (e.g. endotoxin) and/or exotoxins (antigenic proteins produced by bacteria such as staphylococci, streptococci and *Pseudomonas*). Endotoxin is a lipopolysaccharide (LPS) derived from the cell wall of Gram-negative bacteria and is a potent trigger of the inflammatory response. The lipid A portion of LPS can be bound by a protein normally present in human serum known as lipopolysaccharide binding protein (LBP). The LBP/LPS complex attaches to the cell surface marker CD14 and, combined with a secreted protein (MD2), this complex then binds to a member of the toll-like receptor family (TLR4), which transduces the activation signal into the cell. These receptors act through a critical adaptor molecule, myeloid differentiation factor 88 (MyD 88) to regulate the activity of NFκB pathways. Another mechanism in this complex area involves TREM-I (triggering receptor expressed in myeloid cells, see p. 59) which triggers secretion of pro-inflammatory cytokines.

Specific kinases then phosphorylate inhibitory kappa B (IκB), releasing the nuclear transcription factor NFκB, which passes into the nucleus where it binds to DNA and promotes the synthesis of a wide variety of inflammatory mediators. Cell wall components from Gram-positive bacteria, some of which are similar in structure to LPS (e.g. lipoteichoic acid), can also trigger a systemic inflammatory response, probably through similar (TLR2), but not identical pathways (see Fig. 15.9).

Activation of complement cascade (see p. 55)

One of the many functions of the complement system is to attract and activate leucocytes, which then marginate onto endothelium and release inflammatory mediators such as proteases and toxic free radicals of oxygen and other reactive oxygen and nitrogen species (see below).

Cytokines (see also p. 54)

Pro-inflammatory cytokines such as the interleukins (ILs) and tumour necrosis factor (TNF) are also mediators of the systemic inflammatory response. TNF release initiates many of the responses to endotoxin, for example, and acts synergistically with IL-1, in part through induction of cyclo-oxygenase, platelet-activating factor (PAF) and nitric oxide synthase (see below). The cytokine network is extremely complex, with many endogenous self-regulating mechanisms. For example, naturally occurring soluble TNF receptors are shed from cell surfaces during the inflammatory response, binding to TNF and thereby reducing its biological activity. An endogenous inhibitory protein that binds competitively to the IL-1 receptor has also been identified.

In addition to pro-inflammatory mediators such as TNF, anti-inflammatory cytokines, e.g. IL-10, are released. The ratio of IL-10 to TNF, and of TNF to TNF receptors, has been shown to be related to mortality in severe sepsis/septic shock. When excessive, this compensatory anti-inflammatory response syndrome (CARS) may be associated with an inappropriate immune hyporesponsiveness.

Platelet-activating factor (PAF)

This vasoactive lipid is released from various cell populations, such as leucocytes and macrophages, in shock. Its effects, which are caused both directly and through the secondary release of other mediators, include hypotension, increased vascular permeability and platelet aggregation.

Products of arachidonic acid metabolism (see Fig. 14.32)

Arachidonic acid, derived from the breakdown of membrane phospholipid, is metabolized to form prostaglandins and leukotrienes, which are key inflammatory mediators (see p. 849).

Lysosomal enzymes

These can cause myocardial depression and coronary vasoconstriction. Furthermore, lysosomal enzymes can convert

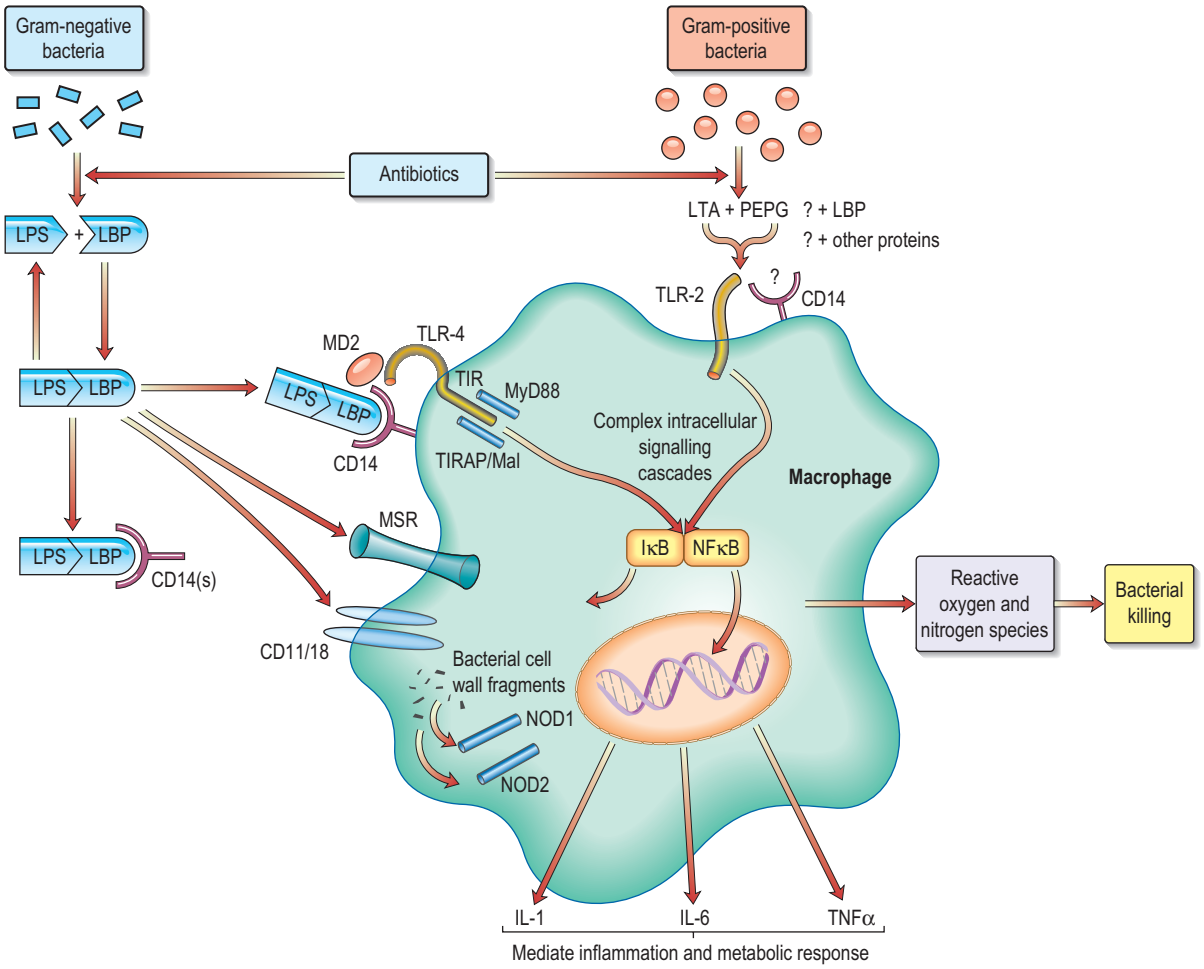


Fig. 15.9 **Induction of the innate immune response** by the lipopolysaccharide–lipopolysaccharide-binding protein (LPS–LBP) complex. This figure illustrates the intracellular events initiated by Gram-negative and Gram-positive bacteria which eventually lead to bacterial killing.

LPS, lipopolysaccharide; LBP, lipopolysaccharide binding protein; LTA, lipoteichoic acid; NFκB, nuclear factor kappa B; IκB inhibitory factor kappa B; PEPG, peptidoglycan-N; TLR, toll-like receptors; MSR, macrophage scavenger receptor; MYD88, myeloid differentiation factor 88; TIR, toll-interleukin receptor; TIRAP, toll-interleukin 1 receptor adaptor protein; MD2 is a secreted protein involved in binding liposaccharide with TLR4; TIRAP/Mal, an adaptor protein for TLR2 and TLR4.

inactive kininogens to vasoactive kinins such as bradykinins. These substances cause vasodilatation and increased capillary permeability, as well as myocardial depression. They can also activate clotting mechanisms.

Adhesion molecules (see also p. 25)

Adhesion of activated leucocytes to the vessel wall and their subsequent extravascular migration is a key component of the sequence of events leading to endothelial injury, tissue damage and organ dysfunction. This process is mediated by inducible intercellular adhesion molecules (ICAMs) found on the surface of leucocytes and endothelial cells. Expression of these molecules can be induced by endotoxin and pro-inflammatory cytokines such as IL-1 and TNF. Several families of molecules are involved in promoting leucocyte–endothelial interaction. The selectins are 'capture' molecules and initiate the process of leucocyte rolling on vascular endothelium, whilst members of the immunoglobulin superfamily (ICAM-1 and vascular cell adhesion molecule-1) are involved in the formation of a more secure bond which leads to leucocyte migration into the tissues (see Fig. 3.13).

Endothelium-derived vasoactive mediators

Endothelial cells synthesize a number of mediators which contribute to the regulation of blood vessel tone and the fluidity of the blood; these include nitric oxide, prostacyclin and endothelin-1 (a potent vasoconstrictor). Nitric oxide (NO) is synthesized from the terminal guanidino-nitrogen atoms of the amino acid L-arginine under the influence of nitric oxide synthase (NOS). NO inhibits platelet aggregation and adhesion and produces vasodilatation by activating guanylate cyclase in the underlying vascular smooth muscle to form cyclic guanosine monophosphate (cGMP) from guanosine triphosphate (GTP) (Fig. 15.10). There are several distinct NOS enzymes.

- *Constitutive or endothelial NOS* (cNOS or eNOS) present in endothelial cells is responsible for the basal release of NO and is involved in the physiological regulation of vascular tone, blood pressure and tissue perfusion.
- *Neuronal NOS* (nNOS). The role of nerves containing nNOS is uncertain but they probably provide neurogenic vasodilator tone. In the central nervous system nNOS

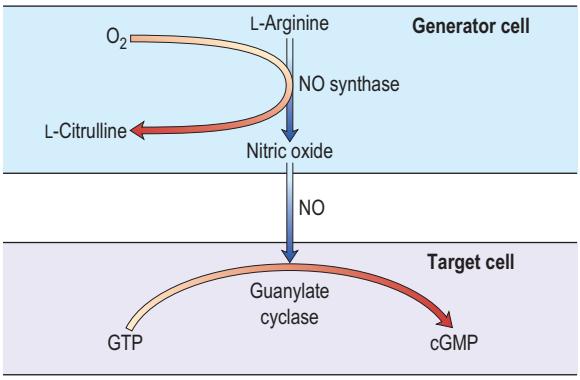


Fig. 15.10 **Synthesis and biochemical action of nitric oxide.** cGMP, cyclic guanosine monophosphate; GTP, guanosine triphosphate.

may be a regulator of local cerebral blood flow as well as fulfilling a number of other physiological functions, such as the acute modulation of neuronal firing behaviour.

- *Inducible NOS* (iNOS) is induced in vascular endothelial smooth muscle cells and monocytes within 4–18 hours of stimulation with certain cytokines, such as TNF, and endotoxin. The resulting prolonged increase in NO formation is believed to be a cause of the sustained vasodilatation, hypotension and reduced reactivity to adrenergic agonists ('vasoplegia') that characterizes septic shock. This mechanism may also be involved in severe prolonged haemorrhage/traumatic shock. The NO generated by macrophages contributes to their role as highly effective killers of intracellular and extracellular pathogens, in part as a consequence of its ability to bind to cytochrome oxidase and inhibit electron transport, but also via the production of the highly reactive radical peroxynitrite.

Redox imbalance

In health the balance between reducing and oxidizing conditions (redox) is controlled by antioxidants which may either prevent radical formation (e.g. transferrin and lactoferrin which bind iron, a catalyst for radical formation) or remove/inactivate reactive oxygen and nitrogen species (e.g. enzymes such as superoxide dismutases, vitamins C and E, and sulphydryl group donors such as glutathione). There are also mechanisms to remove and repair oxidatively damaged molecules and in particular to preserve DNA integrity. In severe systemic inflammation the uncontrolled production of oxygen-derived free radicals and reactive nitrogen species, e.g. superoxide ($O_2^{\bullet -}$), hydroxyl radicals ($OH^{\bullet}$), hydrogen peroxide (H_2O_2) and peroxynitrite ($ONOO^-$), particularly by activated polymorphonuclear leucocytes, can overwhelm these defensive mechanisms and cause:

- lipid and protein peroxidation
- damage to cell membranes
- increased capillary permeability
- impaired mitochondrial respiration
- DNA strand breakage
- apoptosis, which may contribute to the organ damage and immune hypo-responsiveness associated with sepsis.

Influence of genetic variation

Individuals vary considerably in their susceptibility to infection, as well as their ability to recover from apparently similar infections, illnesses or traumatic insults. Evidence now indicates that interindividual variations in susceptibility to, and outcome from, sepsis can be partly explained by polymorphisms of the genes encoding proteins involved in mediating and controlling innate immunity and the inflammatory response.

Haemodynamic and microcirculatory changes

The dominant haemodynamic feature of septic shock is peripheral vascular failure with:

- vasodilatation
- maldistribution of regional blood flow
- abnormalities in the microcirculation:
  - 'stop-flow' capillaries (flow is intermittent)
  - 'no-flow' capillaries (capillaries are obstructed)
  - failure of capillary recruitment
  - increased capillary permeability with interstitial oedema.

Although these *vascular and microvascular abnormalities* may partly account for the reduced oxygen extraction often seen in septic shock, there is also a *primary defect of cellular oxygen utilization* owing to mitochondrial dysfunction (see above). Initially, before hypovolaemia supervenes, or when therapeutic replacement of circulating volume has been inadequate, *cardiac output is usually high and peripheral resistance is low*. These changes may be associated with impaired oxygen consumption, a reduced arteriovenous oxygen content difference, an increased $S_{\bar{v}}o_2$ and a lactic acidosis (so-called 'tissue dysoxia'). Vasodilatation and increased permeability also occur in anaphylactic shock.

In the initial stages of other forms of shock, and sometimes when hypovolaemia and myocardial depression supervene in sepsis and anaphylaxis, cardiac output is low and increased sympathetic activity causes constriction of both precapillary arterioles and, to a lesser extent, the postcapillary venules. This helps to maintain the systemic blood pressure. In addition, the hydrostatic pressure within the capillaries falls and fluid is mobilized from the extravascular space into the intravascular compartment.

Activation of the coagulation system

The inflammatory response to shock, tissue injury and infection is frequently associated with systemic activation of the clotting cascade, leading to platelet aggregation, widespread microvascular thrombosis and inadequate tissue perfusion.

Initially the production of PGI_2 by the capillary endothelium is impaired. Cell damage (for example to the vascular endothelium) leads to exposure to tissue factor (p. 431), which triggers coagulation. In severe cases these changes are compounded by elevated levels of plasminogen activation inhibitor type 1 which impairs fibrinolysis, as well as by deficiencies in physiological inhibitors of coagulation (including antithrombin, proteins C and S and tissue factor-pathway inhibitor). Antithrombin and protein C have a number of anti-inflammatory properties, whereas thrombin is pro-inflammatory.

Plasminogen is converted to plasmin, which breaks down thrombus, liberating fibrin/fibrinogen degradation products (FDPs). Circulating levels of FDPs are therefore increased, the thrombin time, PTT and PT are prolonged and platelet and fibrinogen levels fall. Activation of the coagulation cascade can be confirmed by demonstrating increased plasma levels of D-dimers. The development of *disseminated intravascular coagulation (DIC)* often heralds the onset of multiple organ failure. Because clotting factors and platelets are consumed in DIC, they are unavailable for haemostasis elsewhere and a coagulation defect results – hence the alternative name for DIC is '*consumption coagulopathy*'. In some cases a microangiopathic haemolytic anaemia develops. DIC is particularly associated with septic shock, especially when due to meningococcal infection (see p. 124). Management of the *underlying cause* is most urgent. Supportive treatment may include infusions of fresh frozen plasma, platelets and occasionally factor VIII concentrates.

Reperfusion injury

Restoration of flow to previously hypoxic tissues can exacerbate cell damage through the generation of large quantities of reactive oxygen species and activation of polymorphonuclear leucocytes (see above) (Fig. 15.11). The gut mucosa seems to be especially vulnerable to this 'ischaemia-reperfusion injury'.

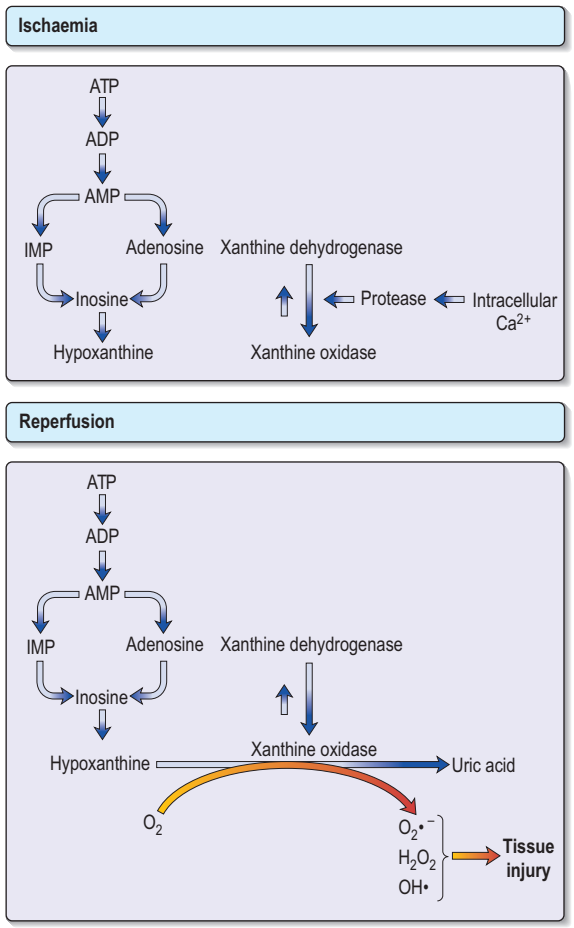


Fig. 15.11 **Generation of reactive oxygen species following ischaemia and reperfusion.** ATP, ADP, AMP, adenosyl tri, di- and mono-phosphate, respectively. IMP, inositol monophosphate.

Metabolic response to trauma, major surgery and severe infection (see also p. 213)

This is initiated and controlled by the neuroendocrine system and various cytokines (e.g. IL-6) acting in concert, and is characterized initially by an increase in energy expenditure ('hypermetabolism'). Gluconeogenesis is stimulated by increased glucagon and catecholamine levels, whilst hepatic mobilization of glucose from glycogen is increased. Catecholamines inhibit insulin release and reduce peripheral glucose uptake. Combined with elevated circulating levels of other insulin antagonists such as cortisol, and downregulation of insulin receptors, these changes ensure that the majority of patients are hyperglycaemic ('insulin resistance'). Later hypoglycaemia may be precipitated by depletion of hepatic glycogen stores and inhibition of gluconeogenesis. Free fatty acid synthesis is also increased, leading to hypertriglyceridaemia.

Protein breakdown is initiated to provide energy from amino acids, and hepatic protein synthesis is preferentially augmented to produce the 'acute phase reactants'. The amino acid glutamine (which is indispensable in this situation) is mobilized from muscle for use as a metabolic fuel in rapidly dividing cells such as leucocytes and enterocytes. Glutamine is also required for hepatic production of the free radical scavenger glutathione. When severe and prolonged, this catabolic response can lead to considerable weight loss. Protein breakdown is associated with wasting and weakness of skeletal and respiratory muscle, prolonging the need for mechanical ventilation and delaying mobilization. Tissue repair, wound healing and immune function may also be compromised.

CLINICAL FEATURES OF SHOCK

Although many clinical features are common to all types of shock, there are certain aspects in which they differ (Box 15.2).

Hypovolaemic shock

- Inadequate tissue perfusion:
  - (a) Skin – cold, pale, slate-grey, slow capillary refill, 'clammy'
  - (b) Kidneys – oliguria, anuria
  - (c) Brain – drowsiness, confusion and irritability.
- Increased sympathetic tone:
  - (a) Tachycardia, narrowed pulse pressure, 'weak' or 'thready' pulse
  - (b) Sweating
  - (c) Blood pressure – may be maintained initially (despite up to a 25% reduction in circulating volume if the patient is young and fit), but later hypotension supervenes.
- Metabolic acidosis – compensatory tachypnoea.

Extreme hypovolaemia may be associated with bradycardia.

Additional clinical features may occur in the following types of shock.

Cardiogenic shock (see p. 742)

Signs of myocardial failure, e.g. raised jugular venous pressure (JVP), pulsus alternans, 'gallop' rhythm, basal crackles, pulmonary oedema.

Box 15.2 Haemodynamic changes in shock

Hypovolaemic shock

Low central venous pressure (CVP) and pulmonary artery occlusion pressure (PAOP)
Low cardiac output
Increased systemic vascular resistance

Cardiogenic shock

Signs of myocardial failure
Increased systemic vascular resistance
CVP and PAOP high (except when hypovolaemic)

Cardiac tamponade

Parallel increases in CVP and PAOP
Low cardiac output
Increased systemic vascular resistance

Pulmonary embolism

Low cardiac output
High CVP, high pulmonary artery pressure but low PAOP
Increased systemic vascular resistance

Anaphylaxis

Low systemic vascular resistance
Low CVP and PAOP
High cardiac output

Septic shock

Low systemic vascular resistance
Low CVP and PAOP
Cardiac output usually high
Myocardial depression – low ejection fraction
Stroke volume maintained by ventricular dilatation
Cardiac output maintained or increased by tachycardia

Obstructive shock

- Elevated JVP.
- Pulsus paradoxus and muffled heart sounds in cardiac tamponade.
- Signs of pulmonary embolism (see p. 784).

Anaphylactic shock (see p. 73)

- Signs of profound vasodilatation:
  (a) Warm peripheries
  (b) Low blood pressure
  (c) Tachycardia.
- Erythema, urticaria, angio-oedema, pallor, cyanosis.
- Bronchospasm, rhinitis.
- Oedema of the face, pharynx and larynx.
- Pulmonary oedema.
- Hypovolaemia due to capillary leak.
- Nausea, vomiting, abdominal cramps, diarrhoea.

Sepsis, severe sepsis and septic shock

- Pyrexia and rigors, or hypothermia (unusual).
- Nausea, vomiting.
- Vasodilatation, warm peripheries.
- Bounding pulse.
- Rapid capillary refill.
- Hypotension (septic shock).
- Occasionally signs of cutaneous vasoconstriction.
- Other signs:
  (a) Jaundice
  (b) Coma, stupor
  (c) Bleeding due to coagulopathy (e.g. from vascular puncture sites, GI tract and surgical wounds)
  (d) Rash and meningism
  (c) Hyper-, and in more severe cases hypoglycaemia.

The diagnosis of sepsis is easily missed, particularly in the elderly when the classical signs may not be present. Mild confusion, tachycardia and tachypnoea may be the only clues, sometimes associated with unexplained hypotension, a reduction in urine output, a rising plasma creatinine and glucose intolerance.

The clinical signs of sepsis are not always associated with bacteraemia and can occur with non-infectious processes such as pancreatitis or severe trauma. The term 'systemic inflammatory response syndrome' (SIRS) describes the disseminated inflammation that can complicate this diverse range of disorders (Box 15.3). Patterns of systemic inflammatory response are shown in Figure 15.12, which illustrates the proinflammatory response (SIRS) and the counter-regulatory anti-inflammatory response (CARS).

MONITORING CRITICALLY ILL PATIENTS

As well as allowing immediate recognition of changes in the patient's condition, monitoring can also be used to establish or confirm a diagnosis, to gauge the severity of the condition,

Box 15.3 Terminology used in systemic inflammation and sepsis

Infection

Invasion of normally sterile host tissue by microorganisms

Bacteraemia

Viable bacteria in blood

Systemic inflammatory response syndrome (SIRS)

The systemic inflammatory response to a variety of severe clinical insults. The response is manifested by two or more of the following:

- Temperature >38°C or <36°C
- Heart rate >90 beats/min
- Respiratory rate >20 breaths/min or P_aco_2 <4.3 kPa
- White cell count >12 × 10^9/L, <4 × 10^9/L or >10% immature forms

Sepsis

SIRS resulting from documented infection

Severe sepsis

Sepsis associated with organ dysfunction, hypoperfusion or hypotension. Hypoperfusion and perfusion abnormalities may include, but are not limited to, lactic acidosis, oliguria or an acute alteration in mental state

Septic shock

Severe sepsis with hypotension (systolic BP <90 mmHg or a reduction of >40 mmHg from baseline) in the absence of other causes for hypotension and despite adequate fluid resuscitation

(Patients receiving inotropic or vasopressor agents may not be hypotensive when perfusion abnormalities are documented)

Refractory shock

Shock unresponsive to conventional therapy (intravenous fluids and inotropic/vasoactive agents) within 1 hour

Compensatory anti-inflammatory response syndrome (CARS)

Release of anti-inflammatory mediators which downregulate the inflammatory response. If excessive, may lead to inappropriate immune hyporesponsiveness

Mixed antagonistic response syndrome (MARS)

Alternating pro- and anti-inflammatory responses

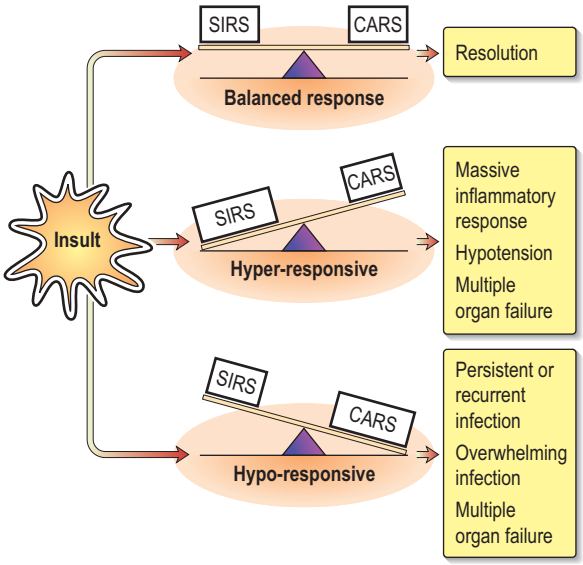


Fig. 15.12 **Pattern of systemic inflammatory response.** CARS, compensatory anti-inflammatory response syndrome; SIRS, systemic inflammatory response syndrome. From Hinds CJ, Watson JD. *Intensive Care: A Concise Textbook,* 3rd edn. Edinburgh: Saunders, 2008, with permission from Elsevier.

to follow the evolution of the illness, to guide interventions and to assess the response to treatment. Invasive monitoring is generally indicated in the more seriously ill patients and in those who fail to respond to initial treatment. These techniques are, however, associated with a significant risk of complications, as well as additional costs and patient discomfort and should therefore only be used when the potential benefits outweigh the dangers. Likewise, invasive devices should be removed as soon as possible.

Assessment of tissue perfusion

- *Pale, cold skin*, delayed capillary refill and the absence of visible veins in the hands and feet indicate poor perfusion. Although peripheral skin temperature measurements can help clinical evaluation, the earliest compensatory response to hypovolaemia or a low cardiac output, and the last to resolve after resuscitation, is vasoconstriction in the splanchnic region.
- *Metabolic acidosis with raised lactate concentration* may suggest that tissue perfusion is sufficiently compromised to cause cellular hypoxia and anaerobic glycolysis. Persistent, severe lactic acidosis is associated with a very poor prognosis. In many critically ill patients, especially those with sepsis, however, lactic acidosis can also be caused by metabolic disorders unrelated to tissue hypoxia and may be exacerbated by reduced clearance owing to hepatic or renal dysfunction.
- *Urinary flow* is a sensitive indicator of renal perfusion and haemodynamic performance.

Blood pressure

Alterations in blood pressure are often interpreted as reflecting changes in cardiac output. However, if there is vasoconstriction with a high peripheral resistance, the blood pressure may be normal, even when the cardiac output is reduced. Conversely, the vasodilated patient may be hypotensive despite a very high cardiac output.

Hypotension may jeopardize perfusion of vital organs. The adequacy of blood pressure in an individual patient must always be assessed in relation to the premorbid value. Blood pressure is traditionally measured using a sphygmomanometer but if rapid alterations are anticipated, continuous monitoring using an intra-arterial cannula is indicated (Practical box 15.1, Fig. 15.13).

Practical Box 15.1 Radial artery cannulation

Technique

1 The procedure is explained to the patient and, if possible, consent obtained
2 The arm is supported, with the wrist extended, by an assistant. (Gloves should be worn)
3 The skin should be cleaned with chlorhexidine
4 The radial artery is palpated where it arches over the head of the radius
5 In conscious patients, local anaesthetic is injected to raise a weal over the artery, taking care not to puncture the vessel or obscure its pulsation
6 A small skin incision is made over the proposed puncture site
7 A small parallel-sided cannula (20 gauge for adults, 22 gauge for children) is used in order to allow blood flow to continue past the cannula
8 The cannula is inserted over the point of maximal pulsation and advanced in line with the direction of the vessel at an angle of approximately 30°
9 'Flashback' of blood into the cannula indicates that the radial artery has been punctured
10 To ensure that the shoulder of the cannula enters the vessel, the needle and cannula are lowered and advanced a few millimetres into the vessel
11 The cannula is threaded off the needle into the vessel and the needle withdrawn
12 The cannula is connected to a non-compliant manometer line filled with saline. This is then connected via a transducer and continuous flush device to a monitor, which records the arterial pressure

Complications

- Thrombosis
- Loss of arterial pulsation
- Distal ischaemia, e.g. digital necrosis (rare)
- Infection
- Accidental injection of drugs – can produce vascular occlusion
- Disconnection – rapid blood loss

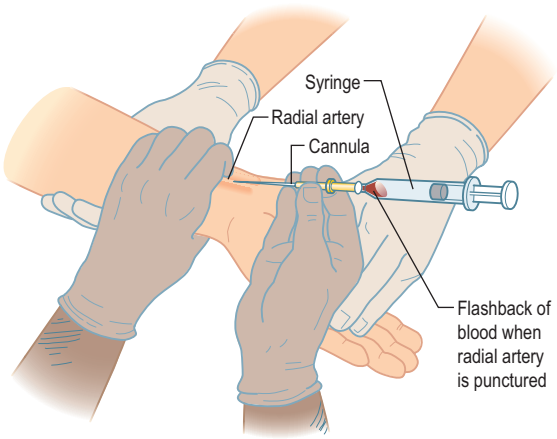


Fig. 15.13 **Percutaneous cannulation of the radial artery.**

Central venous pressure (CVP)

This provides a fairly simple, but approximate method of gauging the adequacy of a patient's circulating volume and the contractile state of the myocardium. The absolute value of the CVP is not as useful as its response to a fluid challenge (the infusion of 100–200 mL of fluid over a few minutes) (Fig. 15.14). The hypovolaemic patient will initially respond to transfusion with little or no change in CVP, together with some improvement in cardiovascular function (falling heart rate, rising blood pressure, increased peripheral temperature and urine output). As the normovolaemic state is approached, the CVP usually rises slightly and stabilizes, while other cardiovascular values begin to normalize. At this stage, volume replacement should be slowed, or even stopped, in order to avoid overtransfusion (indicated by an abrupt and sustained rise in CVP, often accompanied by some deterioration in the patient's condition). In cardiac failure the venous pressure is usually high; the patient will not improve in response to volume replacement, which will cause a further, sometimes dramatic, rise in CVP.

The central venous catheter is usually inserted via a percutaneous puncture of a subclavian or internal jugular vein (Practical box 15.2, Fig. 15.15). Techniques using a guidewire are generally safer and more reliable than the catheter over needle devices (Fig. 15.16). They can also be used in conjunction with a vein dilator for inserting multilumen catheters, double lumen cannulae for haemofiltration or pulmonary artery catheter introducers.

The CVP may be read intermittently using a manometer system or continuously using a transducer and bedside monitor. It is essential that the pressure recorded always be related to the level of the right atrium. Various landmarks are advocated (e.g. sternal notch with the patient supine, sternal angle or mid-axilla when the patient is at 45 degrees), but which is chosen is largely immaterial provided it is used consistently in an individual patient. Pressure measurements should be obtained at end-expiration.

The following are common pitfalls in interpreting central venous pressure readings:

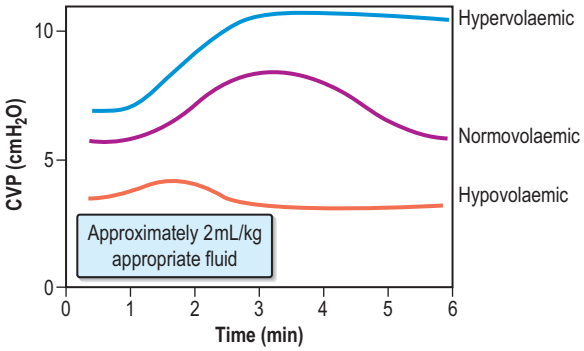


Fig. 15.14 **The effects on the central venous pressure of a rapid administration of a 'fluid challenge' to patients with a CVP within the normal range.** From Sykes MK. Venous pressure as a clinical indication of adequacy of transfusion. *Annals of Royal College of Surgeons of England* 1963; 33: 185–197.

Practical Box 15.2 Internal jugular vein cannulation

Technique

1 The procedure is explained to the patient and, if possible, consent obtained
2 The patient is placed head-down to distend the central veins (this facilitates cannulation and minimizes the risk of air embolism but may exacerbate respiratory distress and is dangerous in those with raised intracranial pressure)
3 The skin is cleaned with an antiseptic solution such as chlorhexidine. Sterile precautions are taken throughout the procedure
4 Local anaesthetic (1% plain lidocaine) is injected intradermally to raise a weal at the apex of a triangle formed by the two heads of sternomastoid with the clavicle at its base
5 A small incision is made through the weal
6 The cannula or needle is inserted through the incision and directed laterally downwards and backwards in the direction of the nipple until the vein is punctured just beneath the skin and deep to the lateral head of sternomastoid. *Ultrasound-guided puncture has been recommended, at least for difficult cases, to reduce the incidence of complications*
7 Check that venous blood is easily aspirated
8 The cannula is threaded off the needle into the vein or the guidewire is passed through the needle (Fig. 15.16)
9 The CVP manometer line is connected to a manometer/transducer
10 A chest X-ray should be taken to verify that the tip of the catheter is in the superior vena cava and to exclude pneumothorax

Possible complications

- Haemorrhage
- Accidental arterial puncture (carotid or subclavian)
- Pneumothorax
- Damage to thoracic duct on left
- Air embolism
- Thrombosis
- Catheter-related sepsis

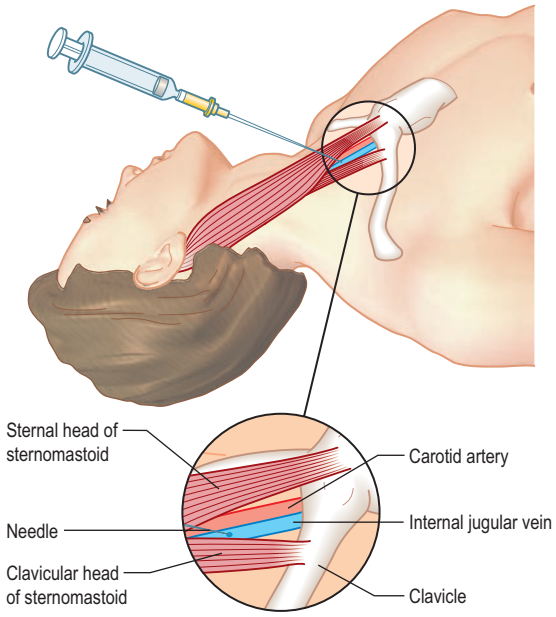


Fig. 15.15 **Cannulation of the right internal jugular vein.**

Ask the authors
www.studentconsult.com

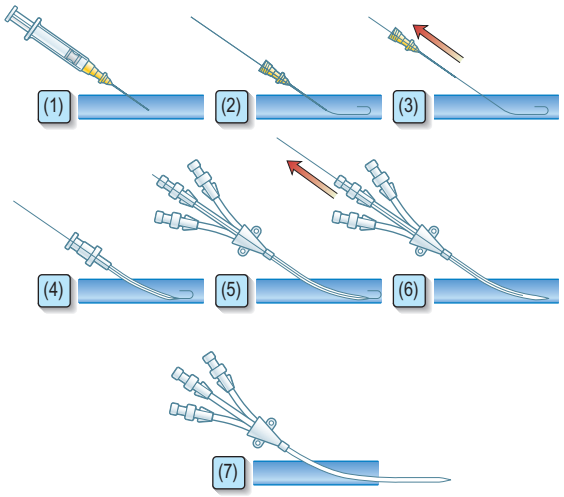


Fig. 15.16 **Seldinger technique – insertion of a catheter over guidewire.** (1) Puncture vessel; (2) advance guidewire; (3) remove needle; (4) dilate vessel; (5) advance catheter over guidewire; (6) remove guidewire; (7) catheter in situ.

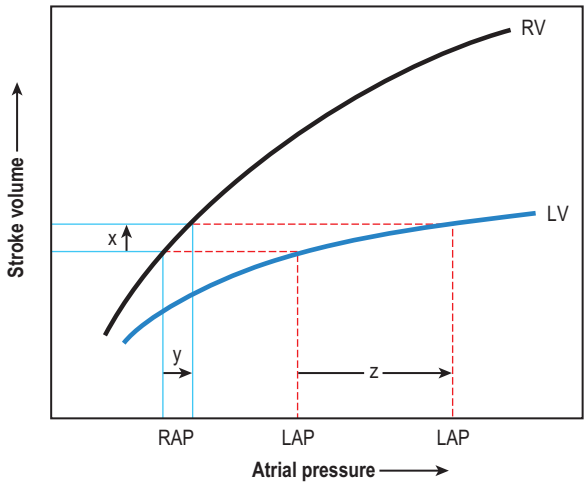


Fig. 15.17 **Left ventricular (LV) and right ventricular (RV) function curves in a patient with left ventricular dysfunction.** Since the stroke volume of the two ventricles must be the same (except perhaps for a few beats during a period of circulatory adjustment), left atrial pressure (LAP) must be higher than right atrial pressure (RAP). Moreover, an increase in stroke volume (x) produced by expanding the circulatory volume may be associated with a small rise in RAP (v) but a marked increase in LAP (z).

Blocked catheter. This results in a sustained high reading, with a damped or absent waveform which often does not correlate with clinical assessment.

Transducer wrongly positioned. Failure to level the system is a common cause of erroneous readings.

Catheter tip in right ventricle. If the catheter is advanced too far, an unexpectedly high pressure with pronounced oscillations is recorded. This is easily recognized when the waveform is displayed.

Arterial pressure variation as a guide to hypovolaemia

Systolic arterial pressure decreases during the inspiratory phase of intermittent positive pressure ventilation (p. 912). The magnitude of this cyclical variability has been shown to correlate more closely with hypovolaemia than other monitored variables, including CVP. Systolic pressure variation during mechanical ventilation can therefore be used as a simple and reliable guide to the adequacy of the circulatory volume.

Left atrial pressure

In uncomplicated cases, careful interpretation of the CVP provides a reasonable guide to the filling pressures of both sides of the heart. In many critically ill patients, however, this is not the case and there is a disparity in function between the two ventricles. Most commonly, left ventricular performance is worst, so that the left ventricular function curve is displaced downward and to the right (Fig. 15.17). High right ventricular filling pressures, with normal or low left atrial pressures, are less common but may occur with right ventricular dysfunction and in situations where the pulmonary vascular resistance (i.e. right ventricular afterload) is raised, such as in acute respiratory failure and pulmonary embolism.

Pulmonary artery pressure

A 'balloon flotation catheter' enables reliable catheterization of the pulmonary artery. These 'Swan–Ganz' catheters can be inserted centrally (see Fig. 15.15) or through the femoral vein, or via a vein in the antecubital fossa. Passage of the catheter from the major veins, through the chambers of the heart, into the pulmonary artery and into the wedged position is monitored and guided by the pressure waveforms recorded from the distal lumen (Fig. 15.18 and Practical box 15.3). A

Practical Box 15.3 Passage of a pulmonary artery balloon flotation catheter through the chambers of the heart into the 'wedged' position

Consent should be obtained from the patient after explanation of the procedure

1. A balloon flotation catheter is inserted through a large vein (see text)
2. Once in the thorax, internal jugular, subclavian or respiratory oscillations are seen. The catheter should be advanced further towards the lower superior vena cava/right atrium **(a)**, where pressure oscillations become more pronounced. The balloon should then be inflated and the catheter advanced
3. When the catheter is in the right ventricle **(b)**, there is no dicrotic notch and the diastolic pressure is close to zero. The patient should be returned to the horizontal, or slightly head-up, position before advancing the catheter further
4. When the catheter reaches the pulmonary artery **(c)** a dicrotic notch appears and there is elevation of the diastolic pressure. The catheter should be advanced further with the balloon inflated
5. Reappearance of a venous waveform indicates that the catheter is 'wedged'. The balloon is deflated to obtain the pulmonary artery pressure. The balloon is inflated intermittently to obtain the pulmonary artery occlusion (also known as pulmonary artery, or capillary, 'wedge') pressure **(d)**

a, b, c, d refer to Fig. 15.18.

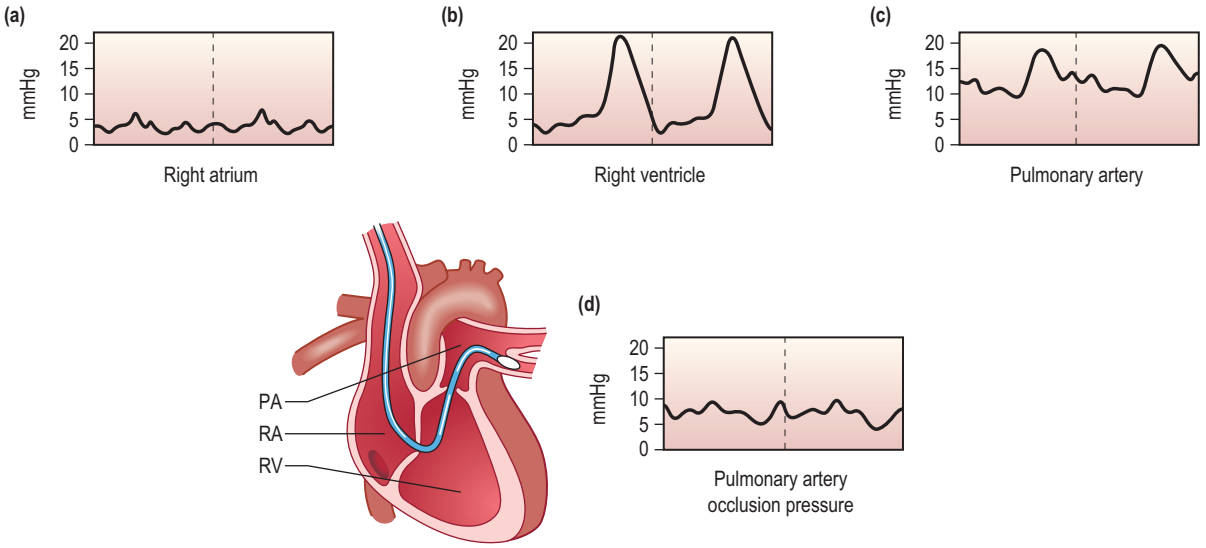


Fig. 15.18 **Passage of pulmonary artery balloon flotation catheter through the chambers of the heart into the 'wedged' position to measure the pulmonary artery occlusion pressure.** See Practical box 15.3.

chest X-ray should always be obtained to check the final position of the catheter. In difficult cases screening with an image intensifier may be required.

Once in place, the balloon is deflated and the pulmonary artery mean, systolic and end-diastolic pressures (PAEDP) can be recorded. The pulmonary artery occlusion pressure (PAOP, previously referred to as the pulmonary artery or capillary 'wedge' pressure) is measured by reinflating the balloon, thereby propelling the catheter distally until it impacts in a medium-sized pulmonary artery. In this position there is a continuous column of fluid between the distal lumen of the catheter and the left atrium, so that PAOP is usually a reflection of left atrial pressure.

The technique is generally safe – the majority of complications are related to user inexperience. Pulmonary artery catheters should preferably be removed within 72 hours, since the incidence of complications, especially infection, then increases progressively (Table 15.4).

Table 15.4 **Pulmonary artery balloon flotation catheters: some complications in addition to those associated with central venous cannulation**

Complication	Comments
Arrhythmias	Occur during passage of catheter through right ventricle Usually benign Can often be prevented with anti-arrhythmic agents
Sepsis	Occurs at insertion site Bacteraemia or endocarditis may develop
Knotting	Occurs when catheter coils in right ventricle and is then withdrawn
Valve trauma	Occurs if catheter is withdrawn with balloon inflated, also tricuspid and pulmonary valves repeatedly open and close on the catheter
Thrombosis/embolism	
Pulmonary infarction	Occurs if catheter remains in 'wedged' position
Pulmonary artery rupture	Usually fatal May occur if balloon is inflated when catheter already 'wedged'
Balloon rupture/leak/embolism	Rare

Cardiac output

Until recently cardiac output was most commonly determined by the thermodilution method, using a modified pulmonary artery catheter with a lumen opening in the right atrium and a thermistor located a few centimetres from its tip. In this method, a known volume (usually 10 mL) of cold 5% dextrose is injected as a bolus into the right atrium. This mixes with, and cools, the blood passing through the heart, and the transient fall in temperature is continuously recorded by the thermistor in the pulmonary artery. The cardiac output is computed from the total amount of indicator (i.e. cold) injected, divided by the average concentration, i.e. the amount of cooling, and the time taken to pass the thermistor. It is also possible to measure cardiac output continuously using a modified pulmonary artery catheter which transmits low heat energy into the surrounding blood and constructs a 'thermodilution curve'. These catheters also optically measure and continuously display $S_{\bar{v}}O_2$.

In general, pulmonary artery catheters enable the clinician to optimize cardiac output and oxygen delivery, whilst minimizing the risk of pulmonary oedema. They can also be used to guide the rational use of inotropes and vasoactive agents. There is increasing evidence, however, to suggest that the use of this monitoring device may not lead to improved outcomes and less invasive techniques are increasingly preferred.

Less invasive techniques for assessing cardiac function

Oesophageal Doppler

Cardiac output and myocardial function can be determined non-invasively using Doppler ultrasonography. A probe is passed into the oesophagus to continuously monitor velocity waveforms from the descending aorta (Fig. 15.19). Although reasonable estimates of stroke volume and cardiac output can be obtained, the technique is best used for trend analysis rather than for making absolute measurements. It is particularly valuable for perioperative optimization of the circulating volume and cardiac performance.

Lithium dilution

Lithium dilution methods do not require pulmonary artery catheterization and are suitable for use in conscious patients. A bolus of lithium chloride is administered via a central venous catheter and the change in arterial plasma lithium concentration is detected by a lithium-sensitive electrode. This sensor can be connected to an existing arterial cannula via a three-way tap. A small battery-powered peristaltic pump is used to create a constant blood flow through the sensor and over the electrode tip. The cardiac output determined in this way can be used to calibrate an arterial pressure waveform ('pulse contour') analysis programme that will continuously monitor changes in cardiac output. Repeated calibration of this system is required.

Echocardiography

Echocardiography is being used increasingly often to provide immediate diagnostic information about cardiac structure and function in the critically ill patient. Transoesophageal echocardiography is preferred because of its superior image clarity.

If there is disagreement between clinical signs and a monitored variable, it should be assumed that the monitor is incorrect until all sources of potential error have been checked and eliminated. Changes and trends in monitored variables are more informative than a single reading.

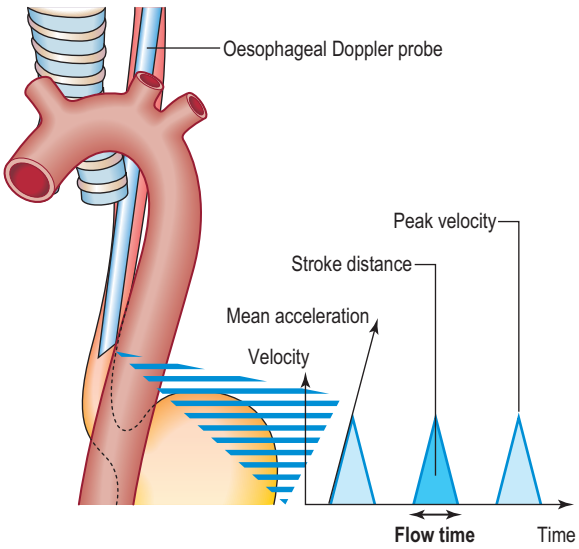


Fig. 15.19 **Doppler ultrasonography.** Stylized velocity waveform traces obtained using oesophageal Doppler ultrasonography. Reproduced from Singer et al. (1991), with permission © 1991 The Williams and Wilkins Co., Baltimore.

MANAGEMENT OF SHOCK (see Fig. 15.20)

Delays in making the diagnosis and in initiating treatment, as well as inadequate resuscitation, contribute to the development of multiple organ failure (MOF) and should be avoided (see p. 909).

A patent airway must be maintained and oxygen must be given. If necessary, an oropharyngeal airway or an endotracheal tube is inserted. The latter has the advantage of preventing aspiration of gastric contents. Very rarely, emergency tracheostomy is indicated (see below). Some patients may require mechanical ventilation.

The underlying cause of shock should be corrected – for example, haemorrhage should be controlled or infection eradicated. In patients with septic shock, every effort must be made to identify the source of infection and isolate the causative organism. As well as a thorough history and clinical examination, X-rays, ultrasonography or CT scanning may be required to locate the origin of the infection. Appropriate samples (urine, sputum, cerebrospinal fluid, pus drained from abscesses) should be sent to the laboratory for microscopy, culture and sensitivities. Several blood cultures should be performed and empirical, broad-spectrum antibiotic therapy (p. 89) should be commenced within the first hour of recognition of sepsis. If an organism is isolated later, therapy can be adjusted appropriately. The choice of antibiotic depends on the likely source of infection, previous antibiotic therapy and known local resistance patterns, as well as on whether infection was acquired in hospital or in the community. Abscesses must be drained and infected indwelling catheters removed.

Whatever the aetiology of the haemodynamic abnormality, tissue blood flow must be restored by achieving and maintaining an adequate cardiac output, as well as ensuring that arterial blood pressure is sufficient to maintain perfusion of vital organs. Recently published guidelines for adult patients suffering severe sepsis or septic shock advocate targeting an MAP > 65 mmHg, CVP 8–12 mmHg (>12 mmHg if mechanically ventilated), urine output >0.5 mL/kg/h and an $S_{cv}O_2 \geq 70\%$ (or $S_{\bar{v}}O_2 \geq 65\%$) as the initial goals of resuscitation.

Preload and volume replacement

Optimizing preload is the most efficient way of increasing cardiac output. Volume replacement is obviously essential in hypovolaemic shock but is also required in anaphylactic and septic shock because of vasodilatation, sequestration of blood and loss of circulating volume because of capillary leak.

In obstructive shock, high filling pressures may be required to maintain an adequate stroke volume. Even in cardiogenic shock, careful volume expansion may, on occasions, lead to a useful increase in cardiac output. On the other hand, patients with severe cardiac failure, in whom ventricular filling pressures may be markedly elevated, often benefit from measures to reduce preload (and afterload) – such as the administration of diuretics and vasodilators (see below). Adequate perioperative volume replacement also reduces morbidity and mortality in high-risk surgical patients.

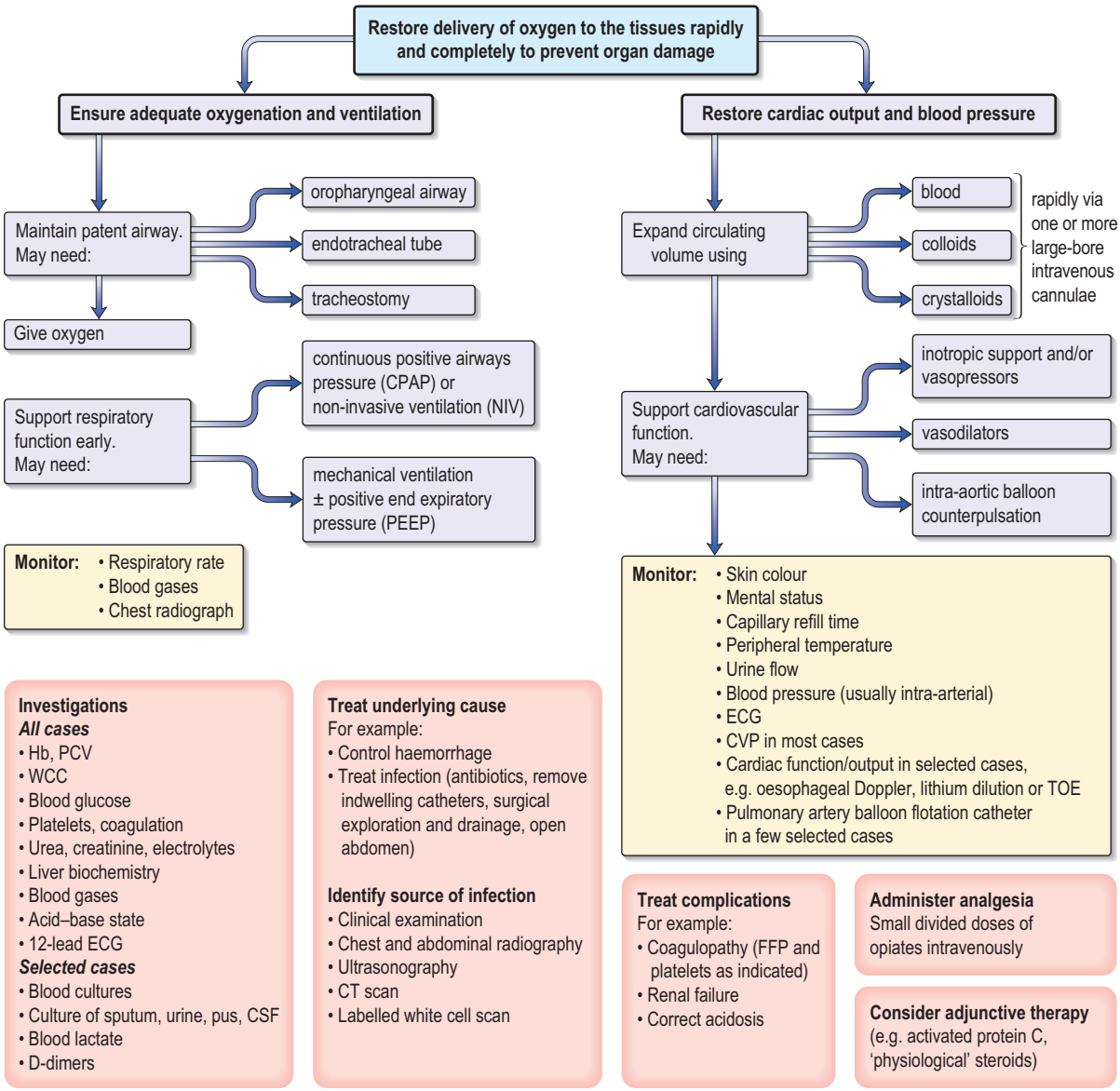


Fig. 15.20 **Management of shock.** CSF, cerebrospinal fluid; CT, computed tomography; CVP, central venous pressure; FFP, fresh frozen plasma.

The circulating volume must be replaced quickly (in minutes not hours) to reduce tissue damage and prevent acute kidney injury. Fluid is administered via wide-bore intravenous cannulae to allow large volumes to be given quickly, and the effect is continuously monitored.

Care must be taken to prevent volume overload, which leads to cardiac dilatation, a reduction in stroke volume, and a rise in left atrial pressure with a risk of pulmonary oedema. Pulmonary oedema is more likely in very ill patients because of a low colloid osmotic pressure (usually due to a low serum albumin) and disruption of the alveolar–capillary membrane (e.g. in acute lung injury).

Choice of fluid for volume replacement

Blood

This is conventionally given for haemorrhagic shock as soon as it is available. In extreme emergencies, group-specific crossmatch can be performed in minutes (see p. 425).

Although red cell transfusion will augment oxygen-carrying capacity, and hence global oxygen delivery, tissue oxygenation is also dependent on microcirculatory flow. This is influenced by the viscosity of the blood and hence the packed cell volume (PCV). Conventionally a PCV of 30–35% has been considered to provide the optimal balance between oxygen-carrying capacity and tissue flow, although it is well recognized that previously fit patients with haemorrhagic shock can tolerate extremely low Hb concentrations provided their circulating volume and cardiac output are maintained. Transfusion of old stored red cells, which are poorly deformable, may be associated with microvascular occlusion and worsening tissue hypoxia.

Complications of blood transfusion are discussed on page 425. Special problems arise when large volumes of stored blood are transfused rapidly. These include:

- *Temperature changes*. Bank blood is stored at 4°C and transfusion may result in hypothermia, peripheral

venoconstriction (which slows the rate of the infusion) and arrhythmias. If possible, blood should therefore be warmed during massive transfusion and in those at risk of hypothermia (e.g. during prolonged major surgery with open body cavity).

- *Coagulopathy*. Stored blood has virtually no effective platelets and is deficient in clotting factors. Large transfusions can therefore produce a coagulation defect. This may need to be treated by replacing clotting factors with fresh frozen plasma and administering platelet concentrates. Occasionally cryoprecipitate may be required.
- *Metabolic acidosis/alkalosis*. Stored blood is preserved in citrate/phosphate/dextrose (CPD) solution, and metabolic acidosis attributable solely to blood transfusion is rare and in any case seldom requires correction. A metabolic alkalosis often develops 24–48 hours after a large blood transfusion, probably mainly owing to metabolism of the citrate. This will be exacerbated if any preceding acidosis has been corrected with intravenous sodium bicarbonate.
- *Hypocalcaemia*. Citrate in stored blood binds calcium ions. During rapid transfusion, citrate may reduce total body ionized calcium levels, causing myocardial depression. This is uncommon in practice, but can be corrected by administering 10 mL of 10% calcium chloride intravenously. Routine treatment with calcium is not recommended.
- *Increased oxygen affinity*. In stored blood, the red cell 2,3-disphosphoglycerate (2,3-DPG) content is reduced, so that the oxyhaemoglobin dissociation curve is shifted to the left. The oxygen affinity of haemoglobin is therefore increased and oxygen unloading is impaired. This effect is less marked with CPD blood. Red cell levels of 2,3-DPG are substantially restored within 12 hours of transfusion.
- *Hyperkalaemia*. Plasma potassium levels rise progressively as blood is stored. However, hyperkalaemia is rarely a problem as rewarming of the blood increases red cell metabolism – the sodium pump becomes active and potassium levels fall.
- *Microembolism*. Microaggregates in stored blood may be filtered out by the pulmonary capillaries. This process is thought by some to contribute to acute lung injury (ALI) (see p. 917).

Red cell concentrates

Standard red cell concentrates are being replaced by supplemental red cell concentrates, in which most of the plasma is removed and the red cells are suspended in 100 mL of nutrient additive solution. The most commonly used additive is SAG (saline, adenine, glucose) to which mannitol (M) is added to reduce spontaneous lysis (SAGM blood).

Concern about the supply, cost and safety of blood, including the risk of disease transmission and immune suppression, has encouraged a more conservative approach to transfusion. There is some evidence to suggest that in normovolaemic critically ill patients a restrictive strategy of red cell transfusion (Hb maintained at >7.0 g/dL) is at least as effective, and may be safer than a liberal transfusion strategy (Hb maintained at 10–12 g/dL). However, in some groups of patients (e.g. the elderly and those with significant cardiac or respiratory disease and patients who are actively bleeding) it may be preferable to maintain Hb closer to the higher level.

Blood substitutes

Attempts to develop an effective and safe oxygen-carrying blood substitute have so far been unsuccessful.

Crystalloids and colloids

The choice of intravenous fluid for resuscitation and the relative merits of crystalloids or colloids has long been controversial. Crystalloid solutions such as saline are cheap, convenient to use and free of side-effects, although they are rapidly lost from the circulation into the extravascular spaces. It has been generally accepted that volumes of crystalloid several times that of colloid are required to achieve an equivalent haemodynamic response and that colloidal solutions produce a greater and more sustained increase in circulating volume, with associated improvements in cardiovascular function and oxygen transport.

Nevertheless, a large, prospective, randomized, controlled trial has demonstrated that in a heterogeneous group of critically ill patients the use of either physiological saline or 4% albumin for fluid resuscitation resulted in similar outcomes.

Polygelatin solutions (Haemaccel, Gelofusin) have an average molecular weight of 35 000, which is iso-osmotic with plasma. They are cheap and do not interfere with crossmatching. Large volumes can be administered, as clinically significant coagulation defects are unusual and renal function is not impaired. However, because they readily cross the glomerular basement membrane, their half-life in the circulation is only approximately 4 hours and they can promote an osmotic diuresis. These solutions are particularly useful during the acute phase of resuscitation, especially when volume losses are continuing. Allergic reactions can, however, occur.

Hydroxyethyl starches (HES). Numerous preparations are now available, characterized by their concentrations (3%, 6%, 10%) and low, medium or high molecular weight. The half-life of high and medium molecular weight solutions is between 12 and 24 hours, whilst that of the low molecular weight solutions is 4–6 hours. Elimination of HES occurs primarily via the kidneys following hydrolysis by amylase. HES are stored in the reticuloendothelial system, apparently without causing functional impairment, but skin deposits have been associated with persistent pruritus. HES, especially the higher molecular weight fractions, have anticoagulant properties and many therefore recommend limiting the volume administered. Recent evidence has implicated HES in the development of acute kidney injury.

Dextrans are polymolecular polysaccharides that have a powerful osmotic effect. They interfere with crossmatching and have a small rate of allergic reactions (0.1–1%) which may be life-threatening. Normally a dose of 1.5 g dextran per kilogram of bodyweight (as Dextran 70) should not be exceeded because of the risk of renal damage. In practice, dextrans are rarely used in the UK because of the availability of other agents.

Human albumin solution (HAS) is a natural colloid which has been used for volume replacement in shock and burns, and for the treatment of hypoproteinaemia. HAS is not generally recommended for routine volume replacement, because supplies are limited and other cheaper solutions are equally effective.

Myocardial contractility and inotropic agents

Myocardial contractility can be impaired by many factors such as hypoxaemia and hypocalcaemia, as well as by some drugs (e.g. beta-blockers, antiarrhythmics, angiotensin-converting enzyme inhibitors and sedatives).

Severe lactic acidosis conventionally is said to depress myocardial contractility and limit the response to vasopressor agents. Attempted correction of acidosis with intravenous sodium bicarbonate, however, generates additional carbon dioxide which diffuses across cell membranes, producing or exacerbating intracellular acidosis. Other disadvantages of bicarbonate therapy include sodium overload and a left shift of the oxyhaemoglobin dissociation curve. Ionized calcium levels may be reduced and, combined with the fall in intracellular pH, this may impair myocardial performance. Treatment of lactic acidosis should therefore concentrate on correcting the cause. Bicarbonate should only be administered to correct *extreme persistent metabolic acidosis* (see p. 675).

If the signs of shock persist despite adequate volume replacement, and perfusion of vital organs is jeopardized, pressor agents should be administered to improve cardiac output and blood pressure. Vasopressor therapy may also be required to maintain perfusion in those with life-threatening hypotension, even when volume replacement is incomplete. All inotropes increase myocardial oxygen consumption, particularly if a tachycardia develops, and this can lead to an imbalance between myocardial oxygen supply and demand, with the development or extension of ischaemic areas. Inotropes should therefore be used with caution, particularly in cardiogenic shock following myocardial infarction and in those known to have ischaemic heart disease.

Many of the most seriously ill patients become increasingly resistant to the effects of pressor agents, an observation attributed to 'downregulation' of adrenergic receptors and NO-induced 'vasoplegia' (p. 895).

All inotropic agents should be administered via a large central vein, and their effects monitored (see also p. 908 and Table 15.5).

Adrenaline (epinephrine)

Epinephrine stimulates both α- and β-adrenergic receptors, but at low doses β effects seem to predominate. This increases heart rate and cardiac index whilst reducing peripheral resistance. If this is associated with an increase in perfusion pressure, urine output may improve. Epinephrine at higher doses can cause excessive (α-mediated) vasoconstriction, with reductions in splanchnic flow, and cardiac output may fall. Prolonged high-dose administration can cause peripheral gangrene and lactic acidosis. The minimum effective dose of adrenaline should therefore be used for as short a time as possible.

Noradrenaline (norepinephrine)

This is predominantly an α-adrenergic agonist. It is particularly useful in those with hypotension associated with a low systemic vascular resistance, for example in septic shock. There is a risk of producing excessive vasoconstriction with impaired organ perfusion and increased afterload. Noradrenaline administration should normally therefore be guided by comprehensive haemodynamic monitoring, including invasive or non-invasive determination of cardiac output (see p. 903) and calculation of systemic vascular resistance.

Dopamine

Dopamine is a natural precursor of adrenaline (epinephrine) which acts on β receptors and α receptors, as well as dopaminergic DA_1 and DA_2 receptors.

In *low doses* (e.g. 1–3 μg/kg/min), dopaminergic vasodilatory receptors in the renal, mesenteric, cerebral and coronary circulations are activated. DA_1 receptors are located on postsynaptic membranes and mediate vasodilatation, whilst DA_2 receptors are presynaptic and potentiate these vasodilatory effects by preventing the release of adrenaline (epinephrine). Renal and hepatic flow increase and urine output is improved. The significance of the renal vasodilator effect of dopamine has, however, been questioned and it has been suggested that the increased urine output is largely attributable to the rise in cardiac output and blood pressure, combined with a decrease in aldosterone and inhibition of tubular sodium reabsorption mediated via DA_1 stimulation.

In *moderate doses* (e.g. 3–10 μg/kg/min), dopamine increases heart rate, myocardial contractility and cardiac output. In some patients the dose of dopamine is limited by β-receptor effects such as tachycardia and arrhythmias.

In *higher doses* (e.g. >10 μg/kg/min) the increased noradrenaline (norepinephrine) produced is associated with vasoconstriction. This increases afterload and raises ventricular filling pressures.

Dopexamine

Dopexamine is an analogue of dopamine which activates β_2 receptors as well as DA_1 and DA_2 receptors. Dopexamine is a weak positive inotrope, but is a powerful splanchnic vasodilator, reducing afterload and improving blood flow to vital organs, including the kidneys. In septic shock, dopexamine can increase cardiac index and heart rate, but causes further reductions in peripheral resistance. It is most useful in those with low cardiac output and peripheral vasoconstriction and has been used as an adjunct to the perioperative management of high-risk patients (see below).

Dobutamine

Dobutamine is closely related to dopamine and has predominantly β_1 activity. Dobutamine has no specific effect on the renal vasculature but urine output often increases as cardiac output and blood pressure improve. It reduces systemic vascular resistance, as well as improving cardiac performance, thereby decreasing afterload and ventricular filling pressures. Dobutamine is therefore useful in patients with cardiogenic shock and cardiac failure. In septic shock, dobutamine can be used to increase cardiac output and oxygen delivery.

Phosphodiesterase inhibitors (e.g. milrinone, enoximone)

These agents have both inotropic and vasodilator properties. Because the phosphodiesterase type III inhibitors bypass the β-adrenergic receptor they do not cause tachycardia and are less arrhythmogenic than β agonists. They may be useful in patients with receptor 'downregulation', those receiving beta-blockers, for weaning patients from cardiopulmonary bypass and for patients with cardiac failure. In vasodilated septic patients, however, they may precipitate or worsen hypotension.

Table 15.5 Receptor actions of sympathomimetic and dopaminergic agents

	β_1	β_2	α_1	α_2	DA_1	DA_2	Dose dependence
Adrenaline (epinephrine)							++++
Low dose	++	+	+	±	–	–	
Moderate dose	++	+	++	+	–	–	
High dose	++(+)	+(+)	++++	+++	–	–	
Noradrenaline (norepinephrine)	++	0	+++	+++	–	–	+++
Isoprenaline	+++	+++	0	0	–	–	
Dopamine							+++++
Low dose	±	0	±	+	++	+	
Moderate dose	++	+	++	+	++(+)	+	
High dose	+++	++	+++	+	++(+)	+	
Dopexamine	+	+++	0	0	++	+	++
Dobutamine	++	+	±	?	0	0	++

Receptor	Action
β_1 – postsynaptic	Positive inotropism and chronotropism Renin release
β_2 – presynaptic	Stimulates noradrenaline (norepinephrine) release
β_2 – postsynaptic	Positive inotropism and chronotropism Vascular dilatation Relaxes bronchial smooth muscle
α_1 – postsynaptic	Constriction of peripheral, renal and coronary vascular smooth muscle Positive inotropism Antidiuresis
α_2 – presynaptic	Inhibition of noradrenaline (norepinephrine) release, vasodilatation
α_2 – postsynaptic	Constriction of coronary arteries Promotes salt and water excretion
DA_1 – postsynaptic	Dilates renal, mesenteric and coronary vessels Renal tubular effect (natriuresis, diuresis)
DA_2 – presynaptic	Inhibits adrenaline (epinephrine) release

Vasopressin

Patients with septic shock have inappropriately low circulating levels of vasopressin. Low-dose vasopressin can increase blood pressure and systemic vascular resistance in patients with vasodilatory septic shock unresponsive to other vasopressors. A recent randomized controlled trial suggests that low-dose vasopressin added to conventional vasopressors may also have some value in less severe septic shock.

Guidelines for use of inotropic and vasopressor agents

Some still consider dopamine in low to moderate doses to be the first-line agent for restoring blood pressure. Dobutamine is particularly indicated in patients in whom the vasoconstriction caused by dopamine could be dangerous (i.e. patients with cardiac disease and septic patients with fluid overload or myocardial failure). Others favour dopexamine as a means of increasing cardiac output and organ blood flow. The combination of dobutamine and (norepinephrine) noradrenaline is popular for the management of patients who are *shocked with a low systemic vascular resistance* (e.g. septic shock). Dobutamine is given to achieve an optimal cardiac output, while noradrenaline (norepinephrine) is used to restore an adequate blood pressure by reducing vasodilatation. In *vasodilated septic patients* with a high cardiac output, noradrenaline (norepinephrine) is used alone. Adrenaline (epinephrine), because of its potency, also remains a useful agent in patients with *refractory hypotension*. Recent evidence, however, indicates that in septic shock, the specific inotrope/vasopressor selected has little influence on outcome.

Targeting 'supranormal' values for oxygen delivery (Do_2) and oxygen consumption (Vo_2)

Although resuscitation has conventionally aimed at achieving normal haemodynamics, survival of many critically ill patients is associated with raised values for cardiac output, Do_2 and Vo_2. However, elevation of Do_2 and Vo_2 to these 'supranormal' levels following admission to intensive care produces no benefit and may be harmful. By contrast, *early* goal-directed therapy in the emergency room, aimed at maintaining a central venous oxygen saturation of more than 70%, significantly improves outcome in patients with severe sepsis or septic shock.

High-risk surgical patients (Box 15.4)

These patients benefit from intensive perioperative monitoring and circulatory support, in particular maintenance of an adequate circulating volume, and postoperative admission to ICU/HDU. Volume replacement and administration of inotropes or vasopressors should be guided by monitoring of stroke volume/cardiac output, usually using an oesophageal Doppler or lithium dilution. The value of the routine use of inodilators such as dopexamine remains unclear.

Vasodilator therapy (see p. 739)

In selected cases, afterload reduction is used to increase stroke volume and decrease myocardial oxygen requirements by reducing the systolic ventricular wall tension. Vasodilatation also decreases heart size and the diastolic ventricular wall tension so that coronary blood flow is improved. The relative magnitude of the falls in preload and

Box 15.4 Patients at risk of developing perioperative multiorgan failure

- Patients with jeopardized cardiorespiratory function
- Patients with trauma to two body cavities requiring multiple blood transfusions
- Patients undergoing surgery involving extensive tissue dissection, e.g. oesophagectomy, pancreatectomy, aortic aneurysm surgery
- Patients undergoing emergency surgery for intra-abdominal or intrathoracic catastrophic states, e.g. faecal peritonitis, oesophageal perforation

Modified from Shoemaker WC et al. *Chest* 1988; **94**: 1176–1178.

afterload depends on the pre-existing haemodynamic disturbance, concurrent volume replacement and the agent selected (see below). Specific vasodilators may also improve microcirculatory flow.

Vasodilator therapy can be particularly helpful in patients with cardiac failure in whom the ventricular function curve is flat (see Fig. 15.3) so that falls in preload have only a limited effect on stroke volume. This form of treatment, combined in selected cases with inotropic support, is therefore useful in cardiogenic shock and in the management of patients with cardiogenic pulmonary oedema, mitral regurgitation or an acute ventricular septal defect.

Nitrate vasodilators are usually used. Nitrates, because of their ability to improve the myocardial oxygen supply/demand ratio, also help to control angina and limit ischaemic myocardial injury. The agents used most commonly to achieve vasodilatation in the critically ill are those which act directly on the vessel wall.

Sodium nitroprusside (SNP) dilates arterioles and venous capacitance vessels, as well as the pulmonary vasculature by donating nitric oxide. SNP therefore reduces the afterload and preload of both ventricles and can improve cardiac output and the myocardial oxygen supply/demand ratio. SNP, however, can exacerbate myocardial ischaemia by producing a 'steal' phenomenon in the coronary circulation. The effects of SNP are rapid in onset and spontaneously reversible within a few minutes of discontinuing the infusion. A large overdose of SNP can cause cyanide poisoning, with intracellular hypoxia caused by inhibition of cytochrome oxidase, the terminal enzyme of the respiratory chain. This is manifested as a metabolic acidosis and a fall in the arteriovenous oxygen content difference.

Nitroglycerine (NTG) and isosorbide dinitrate (ISDN). At low doses, these agents are both predominantly venodilators, but as the dose is increased, they also cause arterial dilatation, thereby decreasing both preload and afterload. Nitrates are particularly useful in the treatment of cardiac failure with pulmonary oedema and are usually used in combination with intravenous furosemide. Both NTG and ISDN reduce pulmonary vascular resistance, an effect that can occasionally be exploited in patients with a low cardiac output secondary to pulmonary hypertension.

Hydralazine predominantly affects arterial resistance vessels. It therefore reduces afterload and blood pressure, while cardiac output and heart rate increase. Hydralazine is usually given as an intravenous bolus to control acute increases in blood pressure.

Mechanical support of the myocardium

Intra-aortic balloon counterpulsation (IABCP) is the technique used most widely for mechanical support of the failing myocardium. It is discussed on page 714.

Sepsis and multiple organ failure (MOF) (also known as multiple organ dysfunction syndrome – MODS)

Sepsis is being diagnosed with increasing frequency and is now the commonest cause of death in non-coronary adult intensive care units. The estimated incidence of severe sepsis has varied from 77 to 300 cases per 100 000 of the population. Mortality rates are high (between 20% and 60%) and are closely related to the severity of illness and the number of organs which fail. Those who die are overwhelmed by persistent or recurrent sepsis, with fever, intractable hypotension and failure of several organs.

Sequential failure of vital organs occurs progressively over weeks, although the pattern of organ dysfunction is variable. In most cases the lungs are the first to be affected (acute lung injury – ALI; acute respiratory distress syndrome – ARDS; see p. 917) in association with cardiovascular instability and deteriorating renal function. Damage to the mucosal lining of the gastrointestinal tract, as a result of reduced splanchnic flow followed by reperfusion, allows bacteria within the gut lumen, or their cell wall components, to gain access to the circulation. The liver defences, which are often compromised by poor perfusion, are overwhelmed and the lungs and other organs are exposed to bacterial toxins and inflammatory mediators released by liver macrophages. Some have therefore called the gut the 'motor of multiple organ failure'. Secondary pulmonary infection, complicating ALI/ARDS, also frequently acts as a further stimulus to the inflammatory response. Later, renal failure and liver dysfunction develop (see p. 910). Gastrointestinal failure, with an inability to tolerate enteral feeding and paralytic ileus, is common. Ischaemic colitis, acalculous cholecystitis, pancreatitis and gastrointestinal haemorrhage may also occur. Features of central nervous system dysfunction include impaired consciousness and disorientation, progressing to coma. Characteristically, these patients initially have a hyperdynamic circulation with vasodilatation and a high cardiac output, associated with an increased metabolic rate. Eventually, however, cardiovascular collapse supervenes and is the usual terminal event.

Adjunctive treatment

Initial attempts to combat the high mortality associated with sepsis concentrated on cardiovascular and respiratory support in the hope that survival could be prolonged until surgery, antibiotics and the patient's own defences had eradicated the infection and injured tissues were repaired. Despite some success, mortality rates remained unacceptably high. So far, attempts to improve outcome by modulating the inflammatory response (including high-dose steroids) or neutralizing endotoxin (Table 15.6) or inhibiting nitric oxide synthesis (e.g. with *N*-monomethyl-L-arginine) have proved disappointing and in some cases may even have been harmful.

Administration of recombinant human activated protein C (rhAPC), an endogenous anticoagulant with anti-inflammatory properties, may improve outcome in some adult patients with severe sepsis and multiple organ dysfunction. Also, it

Table 15.6 Some of the therapeutic strategies tested in randomized, controlled phase II or III trials in human sepsis

High-dose steroids
Endotoxin antibodies
Bactericidal permeability-increasing protein
TNF antibodies
Soluble TNF receptors
Interleukin-1 receptor antagonists
Platelet-activating factor antagonists
N-acetyl cysteine
Nitric oxide synthase inhibition
Antithrombin
Activated protein C

seems that the administration of relatively low, 'stress' doses of hydrocortisone to patients with refractory vasopressor dependent septic shock may assist shock reversal. Careful control of the blood sugar level to below 8.3 mmol/L ('tight glycaemic control') is also recommended.

The aim of current sepsis guidelines is to combine these, and other evidence-based interventions, with early effective resuscitation (aimed especially at achieving an adequate circulating volume, combined with the rational use of inotropes and/or vasoactive agents to maintain blood pressure, cardiac output and oxygen transport) in order to create 'bundles of care' delivered within specific time limits (see http://www.survivingsepsis.org).

FURTHER READING

Brunkhorst FM, Engel C, Bloos F et al. Intensive insulin therapy and pentastarch resuscitation in severe sepsis. *New England Journal of Medicine* 2008; **358:** 125–139.

Dellinger RP, Levy MM, Carlet JM et al. for the International Surviving Sepsis Campaign Guidelines Committee. Surviving Sepsis Campaign: International Guidelines for Management of Severe Sepsis and Septic Shock. *Critical Care Medicine* 2008; **36:** 296–327.

Finfer S, Bellorno R, Boyce N et al. A comparison of albumin and saline for fluid resuscitation in the intensive care unit. *New England Journal of Medicine* 2004; **350:** 2247–2256.

Sprung CL, Annane D, Keh D et al for the CORTICUS Study Group. Hydrocortisone therapy for patients with septic shock. *New England Journal of Medicine* 2008; **358:** 111–124.

RENAL FAILURE (ACUTE KIDNEY INJURY)

Acute kidney injury is a common and serious complication of critical illness which adversely affects the prognosis. The importance of preventing renal failure by rapid and effective resuscitation, as well as the avoidance of nephrotoxic drugs (especially NSAIDs), and control of infection cannot be overemphasized. Shock and sepsis are the most common causes of acute renal failure in the critically ill, but diagnosis of the cause of renal dysfunction is necessary to exclude reversible pathology, especially obstruction (see Ch. 11).

Oliguria is usually the first indication of renal impairment and immediate attempts should be made to optimize cardiovascular function, particularly by expanding the circulating volume and restoring blood pressure. Restoration of the urine output is a good indicator of successful resuscitation. Evidence now suggests that dopamine is not an effective means of preventing or reversing renal impairment and this agent should not be used for renal protection in sepsis (p. 907). If these measures fail to reverse oliguria, administration of diuretics such as furosemide by bolus or infusion, or less often mannitol (for example in rhabdomyolysis) may be indicated (see Ch. 11). If oliguria persists, it is necessary to reduce fluid intake and review drug doses.

Intermittent haemodialysis has a number of disadvantages in the critically ill. In particular it is frequently complicated by hypotension and it may be difficult to remove sufficient volumes of fluid. Peritoneal dialysis is also unsatisfactory in these patients and is contraindicated in those who have undergone intra-abdominal surgery. The use of continuous veno-venous haemofiltration, usually with dialysis (CVVHD), is therefore preferred (see Ch. 11) and is indicated for fluid overload, electrolyte disturbances (especially hyperkalaemia), severe acidosis and, to a lesser extent, uraemia.

If the underlying problems resolve, renal function almost invariably recovers a few days to several weeks later.

FURTHER READING

Friedrich JO, Adhikari N, Herridge MS et al. Meta-analysis: low dose dopamine increases urine output but does not prevent renal dysfunction or death. *Annals of Internal Medicine* 2005; **142:** 510–524.

Ho KH, Sheridan OJ. Meta-analysis of furosemide (to prevent or treat acute renal failure). *British Medical Journal* 2006; **333:** 420–426.

RESPIRATORY FAILURE (see Ch. 14)

Types and causes

The respiratory system consists of a gas-exchanging organ (the lungs) and a ventilatory pump (respiratory muscles/thorax), either or both of which can fail and precipitate respiratory failure. Respiratory failure occurs when pulmonary gas exchange is sufficiently impaired to cause hypoxaemia with or without hypercarbia. In practical terms, respiratory failure is present when the P_aO_2 is <8 kPa (60 mmHg) or the P_aCO_2 is >7 kPa (55 mmHg). It can be divided into:

- type I respiratory failure, in which the P_aO_2 is low and the P_aCO_2 is normal or low
- type II respiratory failure, in which the P_aO_2 is low and the P_aCO_2 is high.

Type I or 'acute hypoxaemic' respiratory failure occurs with diseases that damage lung tissue. Hypoxaemia is due to right-to-left shunts or $\dot{V}/\dot{Q}$ mismatch. Common causes include pulmonary oedema, pneumonia, acute lung injury and, in the chronic situation, lung fibrosis.

Type II or 'ventilatory failure' occurs when alveolar ventilation is insufficient to excrete the volume of carbon dioxide being produced by tissue metabolism. Inadequate alveolar ventilation is due to reduced ventilatory effort, inability to overcome an increased resistance to ventilation, failure to compensate for an increase in deadspace and/or carbon dioxide production, or a combination of these factors. The most common cause is chronic obstructive pulmonary

disease (COPD). Other causes include chest-wall deformities, respiratory muscle weakness (e.g. Guillain–Barré syndrome) and depression of the respiratory centre (e.g. overdose).

Deterioration in the mechanical properties of the lungs and/or chest wall increases the work of breathing and the oxygen consumption/carbon dioxide production of the respiratory muscles. The concept that respiratory muscle fatigue (either acute or chronic) is a major factor in the pathogenesis of respiratory failure is controversial.

MONITORING OF RESPIRATORY FAILURE

A clinical assessment of respiratory distress should be made on the following criteria (those marked with an asterisk* may be indicative of respiratory muscle fatigue):

- the use of accessory muscles of respiration
- intercostal recession
- tachypnoea*
- tachycardia
- sweating
- pulsus paradoxus (rarely present)
- inability to speak, unwillingness to lie flat
- agitation, restlessness, diminished conscious level
- asynchronous respiration (a discrepancy in the timing of movement of the abdominal and thoracic compartments)*
- paradoxical respiration (abdominal and thoracic compartments move in opposite directions)*
- respiratory alternans (breath-to-breath alteration in the relative contribution of intercostal/accessory muscles and the diaphragm).*

Blood gas analysis should be performed to guide oxygen therapy and to provide an objective assessment of the severity of the respiratory failure. The *most sensitive clinical* indicator of increasing respiratory difficulty is a rising respiratory rate. *Measurement of tidal volume* is a less sensitive indicator.

Vital capacity is often a better guide to deterioration and is particularly useful in patients with respiratory inadequacy that is due to neuromuscular problems such as the Guillain–Barré syndrome or myasthenia gravis, in which the vital capacity decreases as weakness increases.

Pulse oximetry

Lightweight oximeters can be applied to an ear lobe or finger. They measure the changing amount of light transmitted through the pulsating arterial blood and provide a continuous, non-invasive assessment of arterial oxygen saturation (S_pO_2). These devices are reliable, easy to use and do not require calibration, although remember that pulse oximetry is not a sensitive guide to *changes* in oxygenation. An S_pO_2 within normal limits in a patient receiving supplemental oxygen does not exclude the possibility of hypoventilation. Readings may be inaccurate in those with poor peripheral perfusion.

Blood gas analysis

Normal values of blood gas analysis are shown in Table 15.2, p. 893.

Errors can result from malfunctioning of the analyser or incorrect sampling of arterial blood. Disposable pre-heparinized syringes are available for blood gas analysis.

- The sample should be analysed immediately. Alternatively, the syringe should be immersed in iced water (the end having first been sealed with a cap) to prevent the continuing metabolism of white cells causing a reduction in PO_2 and a rise in PCO_2.
- The sample must be adequately anticoagulated to prevent clot formation within the analyser. However, excessive dilution of the blood with heparin, which is acidic, will significantly reduce its pH. Heparin (1000 i.u./mL) should just fill the deadspace of the syringe, i.e. approximately 0.1 mL. This will adequately anticoagulate a 2 mL sample.
- Air almost inevitably enters the sample. The gas tensions within these air bubbles will equilibrate with those in the blood, thereby lowering the PCO_2 and usually raising the PO_2 of the sample. However, provided the bubbles are ejected immediately by inverting the syringe and expelling the air that rises to the top of the sample, their effect is insignificant.

Interpretation of the results of blood gas analysis can be considered in two separate parts:

- Disturbances of acid–base balance (see pp. 672 and 893).
- Alterations in oxygenation.

Correct interpretation requires a knowledge of the clinical history, the age of the patient, the inspired oxygen concentration and any other relevant treatment (e.g. the ventilator settings for those on mechanical ventilation and the administration of sodium bicarbonate). The oxygen content of the arterial blood is determined by the percentage saturation of haemoglobin with oxygen. The relationship between the latter and the P_aO_2 is determined by the oxyhaemoglobin dissociation curve (Fig. 15.5).

Capnography

Continuous breath-by-breath analysis of expired carbon dioxide concentration can be used to:

- confirm tracheal intubation
- continuously monitor end-tidal PCO_2, which approximates to P_aCO_2 in normal subjects (may be useful when transporting critically ill patients, for example)
- detect apparatus malfunction
- detect acute alterations in cardiorespiratory function.

MANAGEMENT OF RESPIRATORY FAILURE

Standard management of patients with respiratory failure includes:

- administration of supplemental oxygen
- treatment for airways obstruction
- measures to limit pulmonary oedema
- control of secretions
- treatment of pulmonary infection.

The load on the respiratory muscles should be reduced by improving lung mechanics. Correction of abnormalities which may lead to respiratory muscle weakness, such as hypophosphataemia and malnutrition, is also necessary.

Oxygen therapy

Methods of oxygen administration

Oxygen is initially given via a face mask. In the majority of patients (except patients with COPD and chronically elevated P_aCO_2) the concentration of oxygen given is not vital and oxygen can therefore be given by a 'variable performance' device such as a simple face mask or nasal cannulae (Fig. 15.21).

With these devices the inspired oxygen concentration varies from about 35% to 55%, with oxygen flow rates of between 6 and 10 L/min. Nasal cannulae are often preferred because they are less claustrophobic and do not interfere with feeding or speaking, but they can cause ulceration of the nasal or pharyngeal mucosa. Higher concentrations of oxygen can be administered by using a mask with a reservoir bag attached (Fig. 15.21c). Figure 15.21 should be compared with the fixed-performance mask shown in Figure 14.27, p. 841, with which the oxygen concentration can be controlled. This latter type of mask is used in patients with COPD and chronic type II failure; the dangers of reducing hypoxic drive can be overemphasized – remember, hypoxaemia is more dangerous than hypercapnia.

Oxygen toxicity

Experimentally, mammalian lungs have been shown to be damaged by continuous exposure to high concentrations of oxygen, but oxygen toxicity in humans is less well proven. Nevertheless, it is reasonable to assume that high concentrations of oxygen might damage the lungs, and so the lowest inspired oxygen concentration compatible with adequate arterial oxygenation should be used. Dangerous hypoxia should never be tolerated through a fear of pulmonary oxygen toxicity.

Respiratory support

If, despite the above measures, the patient continues to deteriorate or fails to improve, the institution of some form of respiratory support is necessary (see Table 15.9). Non-invasive ventilation (see p. 916) can be used, particularly in respiratory failure due to COPD, but in the intensive care unit, intermittent positive-pressure ventilation (IPPV) is usually used.

IPPV is achieved by intermittently inflating the lungs with a positive pressure delivered by a ventilator via an endotracheal tube or a tracheostomy. More recently there have been a number of refinements and modifications to the manner in which positive pressure is applied to the airway and in the interplay between the patient's respiratory efforts and mechanical assistance (p. 914).

Controlled mechanical ventilation (CMV), with the abolition of spontaneous breathing, rapidly leads to atrophy of respiratory muscles so that assisted modes that are triggered by the patient's inspiratory efforts (see below) are preferred.

The rational use of mechanical ventilation depends on a clear understanding of its potential beneficial effects, as well as its dangers.

Beneficial effects of mechanical ventilation

- *Improved carbon dioxide elimination*. By increasing the volume of ventilation, the P_aCO_2 can be controlled.
- *Relief from exhaustion*. Mechanical ventilation reduces the work of breathing, 'rests' the respiratory muscles and relieves the extreme exhaustion that may be present in patients with respiratory failure. In some cases, if ventilation is not instituted, this exhaustion may culminate in respiratory arrest.
- *Effects on oxygenation*. Application of positive pressure can prevent or reverse atelectasis. In those with severe pulmonary parenchymal disease, the lungs may be very stiff and the work of breathing is therefore greatly increased. Under these circumstances the institution of respiratory support may significantly reduce total body oxygen consumption; consequently $P_{\bar{v}}O_2$ and thus P_aO_2 may improve. Because ventilated patients are connected to a leak-free circuit, it is possible to administer high concentrations of oxygen (up to 100%) accurately and to apply a positive end-expiratory pressure (PEEP). In selected cases the latter may reduce shunting and increase P_aO_2 (see below).

Indications for mechanical ventilation

- *Acute respiratory failure*, with signs of severe respiratory distress (e.g. respiratory rate >40/min, inability to speak, patient exhausted) persisting despite maximal therapy. Confusion, restlessness, agitation, a decreased conscious level, a rising P_aCO_2 (>8 kPa) and extreme hypoxaemia (<8 kPa), despite oxygen therapy, are further indications.
- *Acute ventilatory failure* due, for example, to myasthenia gravis or Guillain–Barré syndrome. Mechanical ventilation should usually be instituted when the vital

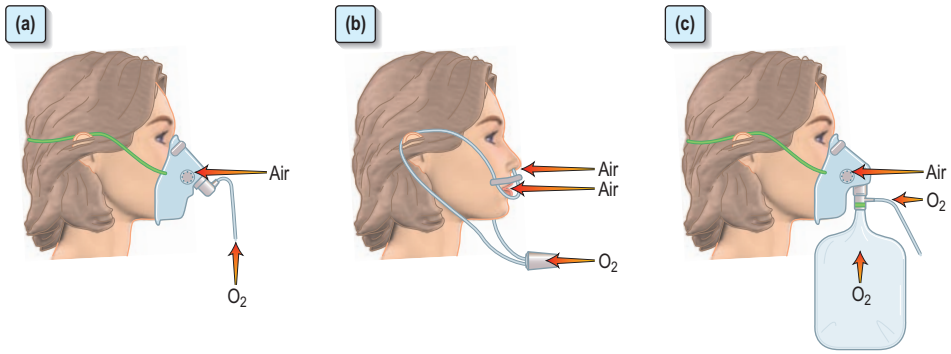


Fig. 15.21 **Methods of administering supplemental oxygen to the unintubated patient. (a)** Simple face mask. **(b)** Nasal cannulae. **(c)** Mask with reservoir bag.

capacity has fallen to 10 mL/kg or less. This will avoid complications such as atelectasis and infection as well as preventing respiratory arrest. The tidal volume and respiratory rate are relatively insensitive indicators of respiratory failure in the above conditions and change late in the course of the disease. A high $P_{a}CO_2$ (particularly if rising) is an indication for urgent mechanical ventilation.

By no means all patients with respiratory failure and/or a reduced vital capacity require ventilation; however, clinical assessment of each individual case is essential. The patient's general condition, degree of exhaustion and level of consciousness are as useful as blood gas values.

Other indications include:

- prophylactic postoperative ventilation in high-risk patients
- head injury – to avoid hypoxia and hypercarbia which increase cerebral blood flow and intracranial pressure
- trauma – chest injury and lung contusion
- severe left ventricular failure with pulmonary oedema
- coma with breathing difficulties, e.g. following drug overdose.

Institution of invasive respiratory support

This requires tracheal intubation. If the patient is conscious, the procedure must be fully explained before anaesthesia is induced and consent obtained. The complications of tracheal intubation are given in Table 15.7.

Intubating patients in severe respiratory failure is an extremely hazardous undertaking and should only be performed by experienced staff. In extreme emergencies it may be preferable to ventilate the patient by hand using an oropharyngeal airway, a face mask and a self-inflating bag with added oxygen until experienced help arrives. An alternative is insertion of a laryngeal mask airway.

The patient is usually hypoxic and hypercarbic, with increased sympathetic activity; the stimulus of laryngoscopy and intubation can precipitate dangerous arrhythmias and even cardiac arrest. Except in an extreme emergency, therefore, the ECG and oxygen saturation should be monitored, and the patient preoxygenated with 100% oxygen before intubation. Resuscitation drugs should be immediately available. If time allows, the circulating volume should be optimized and, if necessary, inotropes commenced before attempting intubation. In some cases it may be appropriate to establish intra-arterial and central venous pressure monitoring before instituting mechanical ventilation, although many patients will not tolerate the supine or head-down position. In some deeply comatose patients, no sedation may be required, but in the majority of patients a short-acting intravenous anaesthetic agent followed by muscle relaxation will be necessary. When available, capnography can confirm tracheal intubation.

Tracheostomy

Tracheostomy may be required for the long-term control of excessive bronchial secretions, particularly in those with a reduced conscious level, and/or to maintain an airway and protect the lungs in those with impaired pharyngeal and laryngeal reflexes. Tracheostomy is also performed when intubation is likely to be prolonged, for patient comfort and to facilitate weaning from mechanical ventilation.

Tracheostomy can be performed at the bedside in an ICU or in theatre. A percutaneous dilatational technique, which is quick and economical, is a suitable technique in critically ill patients (and can be used in an emergency). Alternatively, the trachea can be opened through the second, third and fourth tracheal rings via a transverse skin incision in theatre.

A life-threatening obstruction of the upper respiratory tract that cannot be bypassed with an endotracheal tube can

Table 15.7 Complications of endotracheal intubation

Complication	Comments
Immediate	
Tube in oesophagus	Gives rise to hypoxia and abdominal distension Detected by capnography Requires immediate removal and insertion of the tracheal tube
Tube in one or other (usually the right) bronchus	Avoid by checking both lungs are being inflated; i.e. both sides of the chest move and air entry is heard on auscultation Obtain chest X-ray to check position of tube and to exclude lung collapse
Early	
Migration of the tube out of the trachea Leaks around the tube Obstruction of tube because of kinking or secretions	Dangerous complications The patient becomes distressed, cyanosed and has poor chest expansion The following should be performed immediately: ■ Manual inflation with 100% oxygen ■ Endotracheal suction ■ Check position of tube ■ Deflate cuff ■ Check tube for 'kinks' If no improvement, remove tube, ventilate with face mask and then insert new endotracheal tube
Late	
Sinusitis Mucosal oedema and ulceration Laryngeal injury Tracheal narrowing and fibrosis Tracheomalacia	

be relieved by a *cricothyroidotomy*, which is safer, quicker and easier to perform than a formal tracheostomy.

Tracheostomy has a small but significant mortality rate. Complications of tracheostomy are shown in Table 15.8. With any tracheostomy, care should be taken to ensure that the tube is not blocked by secretions.

Complications associated with mechanical ventilation

Airway complications. There may be complications with tracheal intubation with additional local complications of a tracheostomy (see above) (Tables 15.7 and 15.8).

Disconnection, failure of gas or power supply, mechanical faults. These are unusual but dangerous. A method of manual ventilation, such as a self-inflating bag, and oxygen must always be available by the bedside.

Cardiovascular complications. The application of positive pressure to the lungs and thoracic wall impedes venous return and distends alveoli, thereby 'stretching' the pulmonary capillaries and causing a rise in pulmonary vascular resistance. Both these mechanisms can produce a fall in cardiac output.

Respiratory complications. Mechanical ventilation can be complicated by a deterioration in gas exchange because of $\dot{V}/\dot{Q}$ mismatch and collapse of peripheral alveoli. Traditionally the latter was prevented by using high tidal volumes (10–12 mL/kg) but high inflation pressures, with overdistension of compliant alveoli, perhaps exacerbated by the repeated opening and closure of distal airways, can disrupt the alveolar–capillary membrane. There is an increase in microvascular permeability and release of inflammatory mediators leading to '*ventilator-associated lung injury*'. Extreme overdistension of the lungs during mechanical ventilation with high tidal volumes and PEEP can rupture alveoli and cause air to dissect centrally along the perivascular sheaths. This '*barotrauma*' may be complicated by pneumomediastinum, subcutaneous emphysema, pneumoperitoneum, pneumothorax, and intra-abdominal air. The risk of pneumothorax is increased in those with destructive lung disease (e.g. necrotising pneumonia, emphysema), asthma or fractured ribs.

A *tension pneumothorax* can be rapidly fatal in ventilated patients. Suggestive signs include the development or worsening of hypoxia, hypercarbia, respiratory distress, an unexplained increase in airway pressure, as well as hypotension and tachycardia, sometimes accompanied by a rising CVP. Examination may reveal unequal chest expansion, mediastinal shift away from the side of the pneumothorax (deviated trachea, displaced apex beat) and a hyperresonant hemithorax. Although, traditionally, breath sounds are diminished over the pneumothorax, this sign can be extremely misleading in ventilated patients. If there is time, the diagnosis can be confirmed by chest X-ray prior to definitive treatment with chest tube drainage.

Ventilator-associated pneumonia. Nosocomial pneumonia occurs in as many as one-third of patients receiving mechanical ventilation and may be associated with a significant increase in mortality. It can be difficult to diagnose. Multiple organisms have been isolated such as aerobic Gram-negative bacilli, e.g. *Pseudomonas aeruginosa, Klebsiella pneumoniae, E. coli*, *Acinetobacter* spp. and *Staphylococcus aureus*, including MRSA. Leakage of infected oropharyngeal secretions past the tracheal cuff is thought to be largely responsible. Bacterial colonization of the oropharynx may be promoted by regurgitation of colonized gastric fluid and the risk of nosocomial pneumonia can be reduced by nursing patients in the semi-recumbent, rather than the supine, position and by oropharyngeal decontamination. Treatment is complicated by multiple drug resistance to many antibiotics and mortality is high.

Gastrointestinal complications. Initially, many ventilated patients will develop abdominal distension associated with an ileus. The cause is unknown, although the use of opiates may in part be responsible.

Salt and water retention. Mechanical ventilation, particularly with PEEP, causes increased ADH secretion and possibly a reduction in circulating levels of atrial natriuretic peptide. Combined with a fall in cardiac output and a reduction in renal blood flow, these can cause salt and water retention.

Table 15.8 Complications of tracheostomy

As for tracheal intubation (Table 15.7), plus:

Early
- Death
- Pneumothorax
- Haemorrhage
- Hypoxia
- Hypotension
- Cardiac arrhythmias
- Tube misplaced in pretracheal subcutaneous tissues
- Subcutaneous emphysema

Intermediate
- Mucosal ulceration
- Erosion of tracheal cartilages (may cause tracheo-oesophageal fistula)
- Erosion of innominate artery (may lead to fatal haemorrhage)
- Stomal infection
- Pneumonia

Late
- Failure of stoma to heal
- Tracheal granuloma
- Tracheal stenosis at level of stoma, cuff or tube tip
- Collapse of tracheal rings at level of stoma
- Cosmetic

Techniques for respiratory support (Table 15.9)

Controlled mechanical ventilation (CMV)

This technique is used in patients in whom respiratory efforts are absent, with preset values on the ventilator for tidal volume and respiratory rate.

Table 15.9 Techniques for respiratory support

Intermittent positive-pressure ventilation (IPPV)	**Requires endotracheal tube or tracheostomy**
Controlled mechanical ventilation (CMV)	Can be volume-controlled, or pressure controlled (see text) or pressure limited—volume controlled
■ PEEP (CPPV)	
■ Inverse ratio ventilation (IRV)	
■ Low tidal volume (± 'permissive hypercarbia')	To prevent ventilator induced lung injury
Synchronized intermittent mandatory ventilation (SIMV)	Synchronized ventilation with patient's own respiratory effort Spontaneous respiration with additional ventilator breaths
Continuous positive airway pressure (CPAP)	Positive airway pressure delivered continuously in a spontaneously breathing patient
Pressure support ventilation (PSV)	Spontaneous breaths augmented by positive pressure
Extracorporeal techniques	ECMO – see text $ECCO_2$-R
Non-invasive respiratory support	**Can be given by tight-fitting nasal or facial mask or hood**
Continuous positive airway pressure (CPAP)	Nearly always with CPAP
Non-invasive pressure support ventilation	Inspiratory and expiratory pressures and times set independently
Bilevel positive airway pressure (BiPAP)	

CPPV, Continuous positive-pressure ventilation; ECMO, extracorporeal membrane oxygenation; $ECCO_2$-R, extracorporeal CO_2 removal; PEEP, positive end-expiratory pressure.

Ventilation may be volume limited or pressure limited:

- *Volume controlled ventilation*. The tidal volume, respiratory rate and the positive end-expiratory pressure (PEEP) are preset on the ventilator. The airway pressure varies according to both the ventilatory settings and the patient's lung mechanics (airways resistance and compliance).
- *Pressure controlled ventilation*. Both the inspiratory pressure and the respiratory rate are preset but the tidal volume varies according to the patient's lung compliance and airways resistance.

Positive end-expiratory pressure (PEEP)

A positive airway pressure can be maintained at a chosen level throughout expiration by attaching a threshold resistor valve to the expiratory limb of the circuit. PEEP re-expands underventilated lung units, and redistributes lung water from the alveoli to the perivascular interstitial space, thereby reducing shunt and increasing the P_aO_2. Unfortunately, however, the inevitable rise in mean intrathoracic pressure that follows the application of PEEP may further impede venous return, increase pulmonary vascular resistance and thus reduce cardiac output. This effect is probably least when the lungs are stiff. The fall in cardiac output can be ameliorated by expanding the circulating volume, although in some cases inotropic support or a vasopressor may be required. Thus, although arterial oxygenation is often improved by the application of PEEP, a simultaneous fall in cardiac output can lead to a reduction in total oxygen delivery.

PEEP should be considered if it proves difficult to achieve adequate oxygenation of arterial blood (more than 90% saturation) despite raising the inspired oxygen concentration above 50%. Many use low levels of PEEP (5–7 cmH_2O) in the majority of mechanically ventilated patients in order to maintain lung volume.

Continuous positive airway pressure (CPAP)

The application of CPAP achieves for the spontaneously breathing patient what PEEP does for the ventilated patient. Oxygen and air are delivered under pressure via an endotracheal tube, a tracheostomy, a tightly fitting face mask or a hood. Not only can this improve oxygenation, but the lungs become less stiff, and the work of breathing is reduced.

Pressure support ventilation (PSV)

Spontaneous breaths are augmented by a pre-set level of positive pressure (usually between 5 and 20 cmH_2O) triggered by the patient's spontaneous respiratory effort and applied for a given fraction of inspiratory time or until inspiratory flow falls below a certain level. Tidal volume is determined by the set pressure, the patient's effort and pulmonary mechanics. The level of pressure support can be reduced progressively as the patient improves.

Intermittent mandatory ventilation (IMV)

This technique allows the patient to breathe spontaneously between the 'mandatory' tidal volumes delivered by the ventilator. These mandatory breaths are timed to coincide with the patient's own inspiratory effort (synchronized IMV, or SIMV). SIMV can be used with or without CPAP, and spontaneous breaths may be assisted with pressure support.

'Lung-protective' ventilatory strategies

These are designed to avoid exacerbating or perpetuating lung injury by avoiding overdistension of alveoli, minimizing airway pressures and preventing the repeated opening and closure of distal airways. Alveolar volume is maintained with PEEP, and sometimes by prolonging the inspiratory phase, while tidal volumes are limited to 6 mL/kg predicted bodyweight. Peak airway pressures should not exceed 35–40 cmH_2O. An alternative is to deliver a constant pre-set

Ask the authors
www.studentconsult.com

inspiratory pressure for a prescribed time in order to generate a low tidal volume at reduced airway pressures ('pressure-limited' mechanical ventilation). Respiratory rate can be increased to improve CO_2 removal and avoid severe acidosis (pH <7.2), but hypercarbia is frequent and should be accepted ('permissive hypercarbia'). Both techniques can be used with SIMV. Ventilation with low tidal volumes has been shown to improve outcome in patients with acute lung injury (ALI) or the acute respiratory distress syndrome (ARDS) (see p. 917).

Extracorporeal gas exchange (ECGE)

In patients with severe refractory respiratory failure veno-venous bypass through a membrane lung (extracorporeal membrane oxygenation – ECMO, or extracorporeal carbon dioxide removal – $ECCO_2$-R) has been used to reduce ventilation requirements, thereby minimizing further ventilation-induced lung damage and encouraging resolution of the lung injury. Although randomized controlled trials have indicated that this technique does not improve outcome in adults, some authorities remain convinced that, when used by experienced teams in specialist centres, extracorporeal gas exchange can significantly reduce the high mortality associated with severe ARDS.

Non-invasive ventilation (NIV)

NIV can be used as a therapeutic trial with a view to tracheal intubation if it fails. It is also used as the ceiling of treatment in patients who are not candidates for intubation. NIV is suitable for patients who are conscious, cooperative and able to protect their airway; they must also be able to expectorate effectively. Positive pressure is applied to the airways using tight-fitting full-face or nasal masks or a hood so that tracheal intubation is avoided. Techniques include CPAP, and/or bilevel positive airway pressure (BiPAP). With the latter technique, inspiratory and expiratory pressure levels and times are set independently and unrestricted spontaneous respiration is possible throughout the respiratory cycle. BiPAP can also be patient triggered. There is a reduced risk of ventilator-associated pneumonia and improved patient comfort, with preservation of airway defence mechanisms, speech and swallowing which allows better nutrition. Spontaneous coughing and expectoration are not hampered, allowing effective physiotherapy, and sedation is unnecessary. Institution of non-invasive respiratory support can rest the respiratory muscles, reduce respiratory acidosis and breathlessness, improve clearance of secretions and re-expand collapsed lung segments. The intubation rate, length of ICU and hospital stay and, in some categories of patient, mortality, may all be reduced. NIV is particularly useful in acute hypercapnic respiratory failure associated with COPD, provided the patient is not profoundly hypoxic. NIV may also be useful as a means of avoiding tracheal intubation in immunocompromised patients with acute respiratory failure. Box 15.5 shows some indications for the use of NIV when standard medical treatment has failed. Remember, NIV should not be used as a substitute for invasive ventilation when the latter is clearly more appropriate.

Box 15.5 Some indications for the use of non-invasive ventilation (NIV)

- Acute exacerbation of COPD (pH <7.35)
- Cardiogenic pulmonary oedema
- Chest wall deformity/neuromuscular disease (hypercapnic respiratory failure)
- Obstructive sleep apnoea
- Severe pneumonia (see Box 14.5)
- Asthma (occasionally)
- Weaning patients from invasive ventilation (p. 917)

Contraindications include facial or upper airway surgery, upper gastrointestinal surgery, inability to protect the airway.

Modified from BTS guidelines after 1997. http://www.britishthoracicsociety.co.uk.

Weaning

Weakness and wasting of respiratory muscles is an inevitable consequence of the catabolic response to critical illness and may be exacerbated by the reduction in respiratory work during mechanical ventilation. Often abnormalities of gas exchange and lung mechanics persist. Not surprisingly, therefore, many patients experience difficulty in resuming spontaneous ventilation. In a significant proportion of patients who have undergone a prolonged period of respiratory support the situation is further complicated by the development of a neuropathy, a myopathy or both.

Critical illness polyneuropathy and myopathy

Polyneuropathies have most often been described in association with persistent sepsis and multiple organ failure (see below). Polyneuropathy is characterized by a primary axonal neuropathy involving both motor and, to a lesser extent, sensory nerves. Clinically the initial manifestation is often difficulty in weaning the patient from respiratory support. There is muscle wasting, the limbs are weak and flaccid, and deep tendon reflexes are reduced or absent. Cranial nerves are relatively spared. Nerve conduction studies confirm axonal damage. The cerebrospinal fluid (CSF) protein concentration is normal or minimally elevated. These findings differentiate critical illness neuropathy from Guillain–Barré syndrome, in which nerve conduction studies show evidence of demyelination and CSF protein is usually high.

The cause of critical illness polyneuropathy is not known and there is no specific treatment. Weaning from respiratory support and rehabilitation are likely to be prolonged. With resolution of the underlying critical illness, recovery can be expected after 1–6 months but muscle weakness and fatigue frequently persist.

Myopathies can also occur. A severe quadriplegic myopathy has been particularly associated with the administration of steroids and muscle relaxants to mechanically ventilated patients with acute, severe asthma.

Often the most severely ill patients will have a combined neuropathy and myopathy.

Criteria for weaning patients from mechanical ventilation

Clinical assessment is the best way of deciding whether a patient can be weaned from the ventilator. The patient's

conscious level, psychological state, metabolic function, the effects of drugs and cardiovascular performance must all be taken into account. A subjective evaluation by an experienced clinician of the patient's response to a short period of spontaneous breathing (spontaneous breathing trial) is the most reliable predictor of weaning success or failure. Objective criteria are based on an assessment of pulmonary gas exchange (blood gas analysis), lung mechanics and muscular strength.

Techniques for weaning

Patients who have received mechanical ventilation for less than 24 hours – for example, after elective major surgery – can usually resume spontaneous respiration immediately and no weaning process is required. This procedure can also be adopted for those who have been ventilated for longer periods but who tolerate a spontaneous breathing trial and clearly fulfil objective criteria for weaning. Techniques of weaning include:

- SIMV as a gradual, controlled method of weaning from mechanical ventilation.
- Gradual reduction of the level of pressure support, with no mandatory breaths, is currently considered by many to be the preferred technique.
- CPAP can prevent the alveolar collapse, hypoxaemia and fall in compliance that might otherwise occur when patients start to breathe spontaneously. It is therefore often used during weaning with SIMV/pressure support and in spontaneously breathing patients prior to extubation.
- Non-invasive ventilation via tightly fitting facial or nasal masks (BiPAP, CPAP).
- Tracheostomy is used frequently in critically ill patients to facilitate weaning from mechanical ventilation.
- The traditional method is to allow the patient to breathe entirely spontaneously for a short time, following which respiratory support is reinstituted. The periods of spontaneous breathing are gradually increased and the periods of respiratory support are reduced. Initially it is usually advisable to ventilate the patient throughout the night. This method can be stressful and tiring for both patients and staff, although this approach is sometimes successful when other methods have failed.

Extubation and tracheostomy decannulation

This should not be performed until patients can cough, swallow, protect their own airway and are sufficiently alert to be cooperative. Patients who fulfil these criteria can be extubated provided their respiratory function has improved sufficiently to sustain spontaneous ventilation indefinitely. Similar considerations guide the elective removal of tracheostomy tubes.

ACUTE LUNG INJURY/ACUTE RESPIRATORY DISTRESS SYNDROME

Definition and causes (see Table 15.10)

Acute lung injury (ALI) and the more severe acute respiratory distress syndrome (ARDS) are diagnosed in an appropriate clinical setting with one or more recognized risk factors. ALI/ARDS can be defined as follows:

- Respiratory distress.
- Stiff lungs (reduced pulmonary compliance resulting in high inflation pressures).
- Chest radiograph: new bilateral, diffuse, patchy or homogeneous pulmonary infiltrates.
- Cardiac: no apparent cardiogenic cause of pulmonary oedema (pulmonary artery occlusion pressure <18 mmHg if measured or no clinical evidence of left atrial hypertension).
- Gas exchange abnormalities: ALI – arterial oxygen tension/fractional inspired oxygen (P_aO_2/ F_IO_2) ratio <40 kPa (<300 mmHg); ARDS – P_aO_2/ F_IO_2 <26.6 kPa (<200 mmHg) (in both cases despite normal arterial carbon dioxide tension and regardless of positive end-expiratory pressure). The criterion for arterial oxygen tension/fractional inspired oxygen is arbitrary and the value of differentiating ALI from ARDS has been questioned.

ALI/ARDS can occur as a non-specific reaction of the lungs to a wide variety of direct pulmonary and indirect non-pulmonary insults. By far the commonest predisposing factor is sepsis, and 20–40% of patients with severe sepsis will develop ALI/ARDS (Table 15.10).

PATHOGENESIS AND PATHOPHYSIOLOGY OF ALI/ARDS

Acute lung injury can be considered as an early manifestation of a generalized inflammatory response with endothelial dysfunction and is therefore frequently associated with the development of multiple organ dysfunction syndrome (MODS).

Non-cardiogenic pulmonary oedema

This is the cardinal feature of ALI and is the first and clinically most evident sign of a generalized increase in vascular permeability caused by the microcirculatory changes and release of inflammatory mediators described previously (see p. 895),

Table 15.10 Disorders associated with acute respiratory distress syndrome

Direct lung injury	Indirect lung injury
Common causes	
Pneumonia	Sepsis
Aspiration of gastric contents	Severe trauma with shock and multiple transfusions
Less common causes	
Pulmonary contusion	Cardiopulmonary bypass
Blast injury	Drug overdose (heroin, barbiturates)
Fat embolism	Acute pancreatitis
Near-drowning	Transfusion-associated lung injury
Inhalational injury (smoke or corrosive gases)	Eclampsia
Reperfusion lung injury after lung transplantation or pulmonary embolectomy	High altitude
Amniotic fluid embolism	

FURTHER READING

Canadian Critical Care Trials Group. A randomised trial of diagnostic techniques for ventilator-associated pneumonia. *New England Journal of Medicine* 2006; **355:** 2619–2630.

Truwit JD, Bernard GR. Noninvasive ventilation. *New England Journal of Medicine* 2004; **350:** 2512–2515.

with activated neutrophils playing a particularly key role. The pulmonary epithelium is also damaged in the early stages, reducing surfactant production and lowering the threshold for alveolar flooding.

Pulmonary hypertension

This is a common feature. Initially, mechanical obstruction of the pulmonary circulation may occur as a result of vascular compression by interstitial oedema, whilst local activation of the coagulation cascade leads to thrombosis and obstruction in the pulmonary microvasculature. Later, pulmonary vasoconstriction may develop in response to increased autonomic nervous activity and circulating substances such as catecholamines, serotonin, thromboxane and complement. Those vessels supplying alveoli with low oxygen tensions constrict (the 'hypoxic vasoconstrictor response'), diverting pulmonary blood flow to better oxygenated areas of lung, thus limiting the degree of shunt.

Haemorrhagic intra-alveolar exudate

This exudate is rich in platelets, fibrin, fibrinogen and clotting factors and may inactivate surfactant and stimulate inflammation, as well as promoting hyaline membrane formation and the migration of fibroblasts into the air spaces.

Resolution, fibrosis and repair

Within days of the onset of lung injury, formation of a new epithelial lining is underway and activated fibroblasts accumulate in the interstitial spaces. Subsequently, interstitial fibrosis progresses, with loss of elastic tissue and obliteration of the lung vasculature, together with lung destruction and emphysema. In those who recover, the lungs are substantially remodelled.

Physiological changes

Shunt and deadspace increase, compliance falls, and there is evidence of airflow limitation. Although the lungs in ALI and ARDS are diffusely injured, the pulmonary lesions, when identified as densities on a CT scan, are predominantly located in dependent regions (Fig. 15.22 a and b). This is partly explained by the effects of gravity on the distribution of extravascular lung water and areas of lung collapse. Pleural effusions are common.

CLINICAL PRESENTATION OF ALI/ARDS

The first sign of the development of ALI/ARDS is often an unexplained tachypnoea, followed by increasing hypoxaemia, with central cyanosis, and breathlessness. Fine crackles are heard throughout both lung fields. Later, the chest X-ray shows bilateral diffuse shadowing, interstitial at first, but subsequently with an alveolar pattern and air bronchograms that may then progress to the picture of complete 'white-out' (Fig. 15.23). The differential diagnosis includes cardiac failure and lung fibrosis.

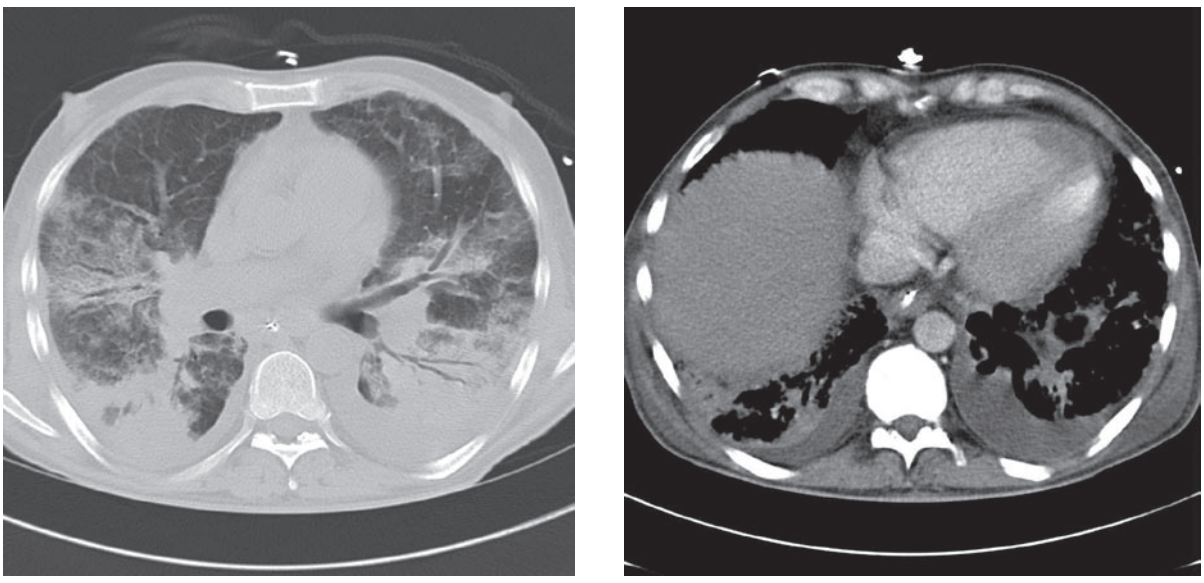
MANAGEMENT OF ALI/ARDS

This is based on treatment of the underlying condition (e.g. eradication of sepsis), avoidance of complications such as ventilator-associated pneumonia, and supportive measures.

Mechanical ventilation

Strategies designed to minimize ventilator-induced lung injury and encourage lung healing should be used (see p. 915).

Pulmonary oedema limitation. Pulmonary oedema formation should be limited by minimizing left ventricular filling pressure with fluid restriction, the use of diuretics and, if these measures fail, preventing fluid overload by haemofiltration. The aim should be to achieve a consistently negative fluid balance. Some recommend that plasma oncotic pressure should be maintained by using colloidal solutions to expand



(a) (b)

Fig. 15.22 **Acute respiratory distress syndrome. (a)** Lung computed tomography scan showing ground-glass opacification in non-dependent regions with atelectasis and consolidation in dependent regions. There are small pleural effusions. **(b)** Same patient as shown in (a) using soft-tissue window settings to demonstrate small bilateral effusions layering in the dependent region of both hemithoraces. From Hinds CJ, Watson JD. Intensive Care: A Concise Textbook, 3rd edn. Edinburgh: Saunders, 2008, with permission from Elsevier. Courtesy of Dr SPG Padley.

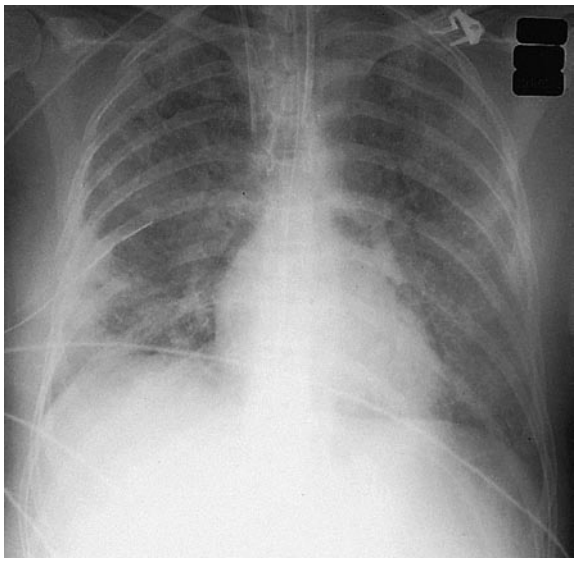


Fig. 15.23 **Chest radiograph appearances in acute respiratory distress syndrome.** Bilateral diffuse alveolar shadowing with air bronchograms and no cardiac enlargement.

the intravascular volume. In patients with ALI/ARDS, however, colloids are unlikely to be retained within the vascular compartment; once they enter the interstitial space, the transvascular oncotic gradient is lost and the main determinants of interstitial oedema formation become the microvascular hydrostatic pressure and lymphatic drainage. There is therefore some controversy concerning the relative merits of colloids or crystalloids for volume replacement in patients likely to develop ALI/ARDS, or in whom the condition is established. Cardiovascular support and the reduction of oxygen requirements are also necessary.

Prone position. When the patient is changed from the supine to the prone position, lung densities in the dependent region are redistributed and shunt fraction is reduced. A reduced pleural pressure gradient, more uniform alveolar ventilation, caudal movement of the diaphragm, redistribution of perfusion and recruitment of collapsed alveoli may all contribute to the improvement in gas exchange. Body position changes can be achieved with minimal complications despite the presence of multiple indwelling vascular lines. Repeated position changes between prone and supine may allow reductions in airway pressures and the inspired oxygen fraction. The response to prone positioning is, however, variable and it seems that this strategy does not improve overall outcome (and perhaps therefore should be reserved for those with severe refractory hypoxaemia).

Inhaled nitric oxide. This vasodilator, when inhaled, may improve $\dot{V}/\dot{Q}$ matching by increasing perfusion of ventilated lung units, as well as reducing pulmonary hypertension. It has been shown to improve oxygenation in so-called 'responders' with ALI/ARDS but has not been shown to increase survival. Its administration requires specialized monitoring equipment, as products of its combination with oxygen include toxic nitrogen dioxide.

Aerosolized prostacyclin. This appears to have similar effects to inhaled NO and is easier to monitor and deliver. As with inhaled NO, the response to aerosolized prostacyclin is, however, variable and although it has been shown to improve oxygenation its effect on outcome has yet to be established.

Aerosolized surfactant. Surfactant replacement therapy reduces morbidity and mortality in neonatal respiratory distress syndrome and is beneficial in animal models of ALI/ARDS. In adults with ARDS, however, the value of surfactant administration remains uncertain.

Steroids. Administration of steroids to patients with established ALI/ARDS does not appear to improve outcome.

Prognosis

Mortality from ALI/ARDS has fallen over the last decade, from around 60% to between 30% and 40%, perhaps as a consequence of improved general care, the increasing use of management protocols, and attention to infection control and nutrition, as well as the introduction of novel treatments and lung-protective strategies for respiratory support. Prognosis is, however, still very dependent on aetiology. When ARDS occurs in association with intra-abdominal sepsis, mortality rates remain very high, whereas much lower mortality rates are to be expected in those with 'primary' ARDS (pneumonia, aspiration, lung contusion). Mortality rises with increasing age and failure of other organs. Most of those dying with ARDS do so as a result of MODS and haemodynamic instability rather than impaired gas exchange.

PATIENT SELECTION – WITHHOLDING AND WITHDRAWING TREATMENT

For many critically ill patients, intensive care is undoubtedly life-saving and resumption of a normal lifestyle is to be expected. In the most seriously ill patients, however, immediate mortality rates are high, a significant number die soon after discharge from the intensive care unit, and the quality of life for some of those who do survive may be poor. Moreover, intensive care is expensive, particularly for those with the worst prognosis, and resources are limited.

Inappropriate use of intensive care facilities has other implications. The patient may experience unnecessary suffering and loss of dignity, while relatives may also have to endure considerable emotional pressures. In some cases treatment may simply prolong the process of dying, or sustain life of dubious quality, and in others the risk of interventions may outweigh the potential benefits.

Both for a humane approach to the management of critically ill patients and to ensure that limited resources are used appropriately, it is necessary to avoid admitting patients who cannot benefit from intensive care and to limit further aggressive therapy when the prognosis is clearly hopeless. Such decisions can be extremely difficult and every case must be assessed individually, taking into account the patient's previous health and quality of life, the primary diagnosis, the medium- and long-term prognosis of the underlying condition (both in terms of survival and quality of life) and the

FURTHER READING

Rubenfeld GD, Caldwell E, Peabody E et al. Incidence and outcomes of acute lung injury. *New England Journal of Medicine* 2005; **353:** 1685–1693.

Wheeler AP, Bernard GR. Acute lung injury and the acute respiratory distress syndrome: a clinical review. *Lancet* 2007; **369:** 1556–1565.

survivability of the acute illness. Age alone should not be a consideration. When in doubt, active measures should continue but should be reviewed regularly in the light of response to treatment and any further information which may become available. Decisions to limit therapy, not to resuscitate in the event of cardiorespiratory arrest, or to withdraw treatment should be made jointly by the medical staff of the unit, the primary physician or surgeon, the nurses and if possible the patient, normally in consultation with the patient's family. Withdrawal or limitation of active treatment should not be viewed negatively as the cessation of all medical or nursing care. Rather, a caring approach should be adopted to ensure that the patient dies with dignity, free of pain and distress, and that family and friends receive support and comfort.

Scoring systems

A variety of scoring systems have been developed that can be used to evaluate the severity of a patient's illness. Some have included an assessment of the patient's previous state of health and the severity of the acute disturbance of physiological function (*acute physiology, age, chronic health evaluation – APACHE*, and *simplified acute physiology score – SAPS*). Other systems have been designed for particular categories of patient (e.g. the injury severity score for trauma victims).

The APACHE and SAPS scores are widely applicable and have been extensively validated. They can quantify accurately the severity of illness and predict the overall mortality for large groups of critically ill patients, and are therefore useful for defining the 'casemix' of patients when auditing a unit's clinical activity, for comparing results nationally or internationally, and as a means of characterizing groups of patients in clinical studies. Although the APACHE and SAPS methodologies can also be used to estimate risks of mortality, no scoring system has yet been devised that can predict with certainty the outcome in an individual patient. They must not, therefore, be used in isolation as a basis for limiting or discontinuing treatment.

FURTHER READING

Carlet J, Thys LG, Antonelli M et al. Challenges in end-of-life care in the ICU. Statement of the 5th International Consensus Conference in Critical Care. *Intensive Care Medicine* 2004; **30:** 770–784.

Knaus WA, Wagner DP, Draper EA et al. The APACHE III prognostic system: risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991; **100:** 1619–1636.

Lautrette A, Darmen M, Megarlane B et al. A communication strategy and brochure for relatives for patients dying in the ICU. *New England Journal of Medicine* 2007; **356:** 469–478.

BRAIN DEATH

Brain death means 'the irreversible loss of the capacity for consciousness combined with the irreversible loss of the capacity to breathe'. Both of these are essentially functions of the brainstem. Death, if thought of in this way, can arise either from causes outside the brain (i.e. respiratory and cardiac arrest) or from causes within the cranial cavity. With the advent of mechanical ventilation it became possible to support such a dead patient temporarily, although in all cases cardiovascular failure eventually supervenes and progresses to asystole.

Before considering a diagnosis of brainstem death it is essential that certain preconditions and exclusions be fulfilled.

Preconditions

- The patient must be in apnoeic coma (i.e. unresponsive and on a ventilator, with no spontaneous respiratory efforts).
- Irremediable structural brain damage due to a disorder that can cause brainstem death must have been diagnosed with certainty (e.g. head injury, intracranial haemorrhage).

Exclusions

- The possibility that unresponsive apnoea is the result of poisoning, sedative drugs or neuromuscular blocking agents must be excluded.
- Hypothermia must be excluded as a cause of coma. The central body temperature should be more than 35°C.
- There must be no significant metabolic or endocrine disturbance that could produce or contribute to coma or cause it to persist.
- There should be no profound abnormality of the plasma electrolytes, acid–base balance, or blood glucose levels.

Diagnostic tests for the confirmation of brainstem death

All brainstem reflexes are absent in brainstem death.

Tests

The following tests should not be performed in the presence of seizures or abnormal postures.

- Oculocephalic reflexes should be absent: In a comatose patient whose brainstem is intact, the eyes will rotate relative to the orbit (i.e. doll's eye movements will be present). In a brainstem dead patient, when the head is rotated from side to side, the eyes move with the head and therefore remain stationary relative to the orbit.
- The pupils are fixed and unresponsive to bright light. Both direct and consensual light reflexes are absent. The size of the pupils is irrelevant, although most often they will be dilated.
- Corneal reflexes are absent.
- There are no vestibulo-ocular reflexes on caloric testing (see p. 1109).
- There is no motor response within the cranial nerve territory to painful stimuli applied centrally or peripherally. Spinal reflex movements may be present.
- There is no gag or cough reflex in response to pharyngeal, laryngeal or tracheal stimulation.
- Spontaneous respiration is absent. The patient should be ventilated with 100% O_2 (or 5% CO_2 in 95% O_2) for 10 minutes and then temporarily disconnected from the ventilator for up to 10 minutes. Oxygenation is maintained by insufflation with 100% oxygen via a catheter placed in the endotracheal tube. The patient is observed for any signs of spontaneous respiratory

efforts. A blood gas sample should be obtained during this period to ensure that the P_aCO_2 is sufficiently high to stimulate spontaneous respiration (>6.7 kPa (50 mmHg)).

The examination should be performed and repeated by two senior doctors.

In the UK it is not considered necessary to perform confirmatory tests such as EEG and carotid angiography.

The primary purpose of establishing a diagnosis of brainstem death is to demonstrate beyond doubt that it is futile to continue mechanical ventilation and other life-supporting measures.

In suitable cases, and provided the assent of relatives has been obtained (easier if the patient was carrying an organ donor card or is on the organ donor register), the organs of those in whom brainstem death has been established may be used for transplantation. In the UK each region has a transplant coordinator who can help with the process, as well as providing information, training and advice about organ donation. They should be informed of all potential donors. In all cases in the UK the coroner's consent must be obtained.

FURTHER READING

Wijdicks EF. The diagnosis of brain death. *New England Journal of Medicine* 2001; **344:** 1215–1221.

CHAPTER BIBLIOGRAPHY

Hinds CJ, Watson D. *Intensive Care: A Concise Textbook*. Edinburgh: Elsevier, 2008.

Society of Critical Care Medicine: http://www.sccm.org/professional_resources/guidelinestable-of-contents/index.asp.

SIGNIFICANT WEBSITES

http://www.ics.ac.uk

UK Intensive Care Society

http://www.esicm.org

European Society of Intensive Care Medicine

http://www.survivingsepsis.org

Surviving Sepsis Campaign

This page intentionally left blank

16 Drug therapy and poisoning

Drug therapy 923
The patient 923
The drug 924
The dose 925
Affordability 927
Adverse drug reactions 927
Evidence-based medicine 929
Statistical analyses 932
Information 933

Poisoning 933
The approach to the patient 934
Principles of management 935
Care of the unconscious patient 935
Specific management 937
Specific poisons: drugs and other chemicals 939
Specific poisons: marine animals 950
Specific poisons: venomous animals 951
Specific poisons: plants 952
Specific poisons: mushrooms 952

DRUG THERAPY

Introduction

This chapter provides an introduction to the principles of rational therapeutics whose cardinal features are encompassed within the 1984 'Nairobi Declaration'. This emphasizes the importance of prescribing:

- to the right patient
- the right drug
- at the right dose
- and at an affordable cost.

THE PATIENT

The prerequisite of any form of therapeutic intervention is a reliable diagnosis or, at least, an assessment of clinical need. An accurate diagnosis ensures that a patient is not exposed, unnecessarily, to the hazards or costs of a particular intervention. Nevertheless there are some circumstances when treatment is used in the absence of a clear diagnosis. Examples include:

- the symptomatic treatment of severe pain
- the initiation of 'blind' antimicrobial therapy where delay would expose a patient to hazard, or discomfort (e.g. antimicrobial therapy to a patient in the community with suspected meningococcal meningitis or a suspected lower urinary tract infection).

In some instances a particular medicine is only effective in subgroups of patients with a particular disorder. Trastuzumab, for example, is only effective in women with breast cancer whose malignant cells express the HER2 epidermal growth factor receptor. It is likely, in the future, that there will be fragmentation of many diseases we currently consider as single entities, with a more precise definition of sites against which drugs can be targeted.

Medicines are also given to otherwise healthy individuals. In such circumstances there must be a very clear imperative to ensure that the benefits to the individual outweigh the harm. Examples include:

- immunization against serious microbial infections (e.g. influenza vaccination)
- the reduction of individual risk factors for overt disease (e.g. the use of antihypertensive, or lipid-lowering, agents)
- the use of oral contraceptives in sexually active women who wish to avoid pregnancy.

Co-morbidity may also significantly alter the way in which conditions are managed, particularly in the elderly. Some examples are shown in Table 16.1.

Table 16.1 Examples of drugs to be avoided in patients with co-morbidity

Co-morbidity	Avoid	Effect
Parkinson's disease	Neuroleptics	Exacerbates Parkinsonian symptoms (including tremor)
Hypertension	Non-steroidal anti-inflammatory drugs	Sodium retention
Asthma	Beta-blockers, adenosine	Bronchospasm
Respiratory failure	Morphine, diamorphine	Respiratory depression
Atrioventricular block	Digoxin, beta-blockers	Heart block
Renovascular disease	ACE inhibitors/ antagonists	Reduction in glomerular filtration
Epilepsy	Tricyclic antidepressants	Lower convulsive threshold

Prescribing in neonates, infants and children

The use of drugs in the newborn, infants and children poses special problems. Extrapolating from adult dosage regimens, merely adjusting for weight, leads to excessive (and potentially toxic) doses because:

- The rates of hepatic metabolism and renal excretion of drugs are reduced in neonates and infants.
- Premature babies have approximately 1% of their body weight as fat (compared to 20% in adults), leading to a marked increase in plasma drug levels of fat-soluble drugs.

Other difficulties in paediatric prescribing:

- Many treatments have never been subject to formal trials in children and their benefits and risks have not, therefore, been adequately assessed in this age group. Efforts are being made, internationally, to redress this.
- For many drugs, there are no paediatric preparations or formulations. Instead, adult products are used.
- Precise oral dosing is often impossible in babies who spit out unpleasant-tasting products!

Prescribing for the elderly

The use of drugs in the elderly is often a problem because:

- Rates of hepatic drug metabolism and renal excretion decline with age. Extrapolation of drug dosages, from those appropriate in younger adults, may therefore lead to toxic plasma levels.
- Changes in drug distribution, associated with a reduction in body mass, changes in body composition, and the preferential distribution of the cardiac output to the brain, may also predispose to toxicity.
- Co-morbidity, often associated with polypharmacy, leads to increased opportunities for drug interactions.
- Concordance with treatment regimens diminishes as the number of prescribed drugs increases, and is especially poor in the face of cognitive impairment.
- Exaggerated pharmacodynamic effects of drugs acting on the central nervous, cardiovascular and gastrointestinal systems are common.

Examples of common problems encountered in the use of drugs amongst older people are shown in Table 16.2.

Table 16.2 Common adverse effects of drugs in the elderly

Drug	Effect
Beta-blockers (including eye drops)	Bradycardia
Digoxin	
Nitrates	Postural hypotension
α-Adrenoceptor-blockers	
Diuretics	
Diuretics (thiazides)	Glucose intolerance, gout
Anti-muscarinic drugs	Confusion, cognitive, dysfunction
Tricyclic antidepressants	
Neuroleptics	
Minor tranquillizers	
Anticonvulsants	
Hypnotics	
Opioids	
Bisphosphonates (mainly alendronic acid)	Oesphageal ulceration and stricture formation
NSAIDs	Gastric erosions
	Upper gastrointestinal bleeding
	Perforated peptic ulcer
	Renal impairment

Drug use in pregnancy

Clinicians should be extremely cautious about prescribing drugs to pregnant women, and only essential treatments should be given. When a known teratogen is needed during pregnancy (e.g. an anticonvulsant drug or lithium) the potential adverse effects should be discussed with the parents preferably before conception. If they decide to go ahead with the pregnancy, they should be offered an appropriate ultrasound scan to assess whether there is any fetal damage. Some known human teratogens are shown in Table 16.3.

Table 16.3 Some human teratogens

Drug	Effect
ACE inhibitors/antagonists	Oligohydramnios
Retinoids, e.g. acitretin	Multiple abnormalities
Amiodarone	Neonatal goitre
Carbimazole	Neonatal hypothyroidism
Warfarin	Dysmorphia Abnormalities of bone growth
Antiepileptics (see p. 1143)	
Carbamazepine	Cleft palate
Phenytoin	
Valproate	Neural tube defects
NSAIDs	Delayed closure of the ductus arteriosus
Cytotoxic drugs	Most are presumed teratogens
Lithium	Ebstein's anomaly

NB: All drugs should be avoided in pregnancy unless benefit clearly outweighs the risk.

Breast-feeding

Although most drugs can be detected in breast milk, the quantity is generally small. This is because, for most drugs, the concentration in milk is in equilibrium with plasma water (i.e. the non-protein-bound fraction). A few drugs (e.g. aspirin, carbimazole) may, however, cause harm to the infant if ingested in breast milk. Relevant drug literature should be consulted when prescribing to a nursing mother.

THE DRUG

Selecting the right drug involves three elements:

- the drug's clinical efficacy for the proposed use
- the balance between the drug's efficacy and safety
- patient preference.

The 'gold standard' (see p. 929) to demonstrating the clinical efficacy of a drug is the randomized controlled trial (RCT), although other approaches (see p. 931) can be informative. The demonstration of absolute efficacy (against placebo) may, itself, be insufficient. Where there is more than one

treatment for the same indication these should be compared with one another, taking account of the magnitude of their benefits, their individual adverse reaction profiles and their costs. Comparative RCTs are particularly useful in this respect.

Patient's preferences should be discussed to enable them to be equal partners in decision-making about whether and how they wish to undergo treatment. A full understanding of the reasons for considering treatment, the likely benefits and the possible adverse reactions, has repeatedly been shown to improve 'concordance' with treatment regimens.

THE DOSE

Appropriate drug dosages will have usually been determined from the results of so-called 'dose-ranging' studies during the original development programme. Such studies are generally conducted as RCTs covering a range of potential doses. Drug doses and dosage regimens may be fixed or adjusted.

Fixed dosage regimens

Drugs suitable (in adults) for prescribing at fixed doses for all patients share common features. Efficacy is optimal in virtually all patients; and the risks of dose-related (type A) adverse reactions (see p. 927) are normally low. These drugs have a high 'therapeutic ratio' (i.e. the ratio between toxic and therapeutic doses). Examples of drugs prescribed at a fixed dose are shown in Table 16.4.

Titrated dosage regimens

For many drugs there are wide interindividual variations in response. As a consequence, whilst a particular dose may in one person lack any therapeutic effect, the same dose in another may cause serious toxicity. The reasons for such variability are partly due to pharmacokinetic factors (differences in the rates of drug absorption, distribution or metabolism) and partly due to pharmacodynamic factors (differences in the sensitivity of target organs).

Pharmacokinetics

The intensity of a drug's action, immediately after parenteral administration, is largely a function of its volume of distribution. This, in turn, is predominantly governed by body composition and regional blood flow. Dosage adjustments, for bodyweight or surface area, are therefore common, for example in cancer chemotherapy in order to optimize treatment.

The main determinants of a drug's plasma concentrations are its bioavailability (the proportion of the unchanged drug that reaches the systemic circulation) and its rate of systemic clearance (by hepatic metabolism or renal excretion). After oral administration, a drug's bioavailability depends on the extent to which it is:

- Destroyed in the gastrointestinal tract.
- Crosses the gastrointestinal epithelium.
- Metabolized by the liver before reaching the systemic circulation (so called presystemic or first pass metabolism). First pass metabolism can be avoided by the intravascular (i.v.), intramuscular (i.m.) or sublingual route.

Liver drug metabolism occurs in two stages:

- *Phase I* is the modification of a drug, by oxidation, reduction or hydrolysis. Of these, oxidation is the most frequent route and is largely undertaken by a family of isoenzymes known as the cytochrome P450 system (see p. 926). Inhibition or induction of cytochrome P450 isoenzymes are major causes of drug interaction (Table 16.5).
- *Phase II* involves conjugation with glucuronic acid, sulphate, acetate or other substances to render it more soluble and therefore able to be excreted in the urine.

Genetic causes of altered pharmacokinetics

Both presystemic hepatic metabolism, and the rate of systemic hepatic clearance, may vary markedly between healthy individuals.

Variability in the genes encoding drug metabolizing (Table 16.6) enzymes are major determinants of the inter-individual differences in the therapeutic and adverse responses to drug treatment. The most common involve polymorphisms of the cytochrome P450 family of enzymes. The first to be discovered was the polymorphism in the hydroxylation of the anti-hypertensive agent debrisoquin (CYP2D6). Defective

Table 16.4 Examples of fixed dose prescribing

Drug	Indication
Aspirin	Secondary prevention of myocardial infarction
Bendroflumethiazide	Hypertension
Amoxicillin	Lower urinary tract infection
	Upper and lower respiratory tract infection
Ferrous sulphate	Iron-deficiency anaemia
Oral contraceptives	Oral contraception

Table 16.5 Some inducers and inhibitors of cytochrome P450

Inducers	Carbamazepine
	Hyperforin*
	Nifedipine
	Non-nucleoside reverse transcriptase inhibitors (NNRTIs)
	Omeprazole
	Paclitaxol
	Phenytoin
	Rifampicin
	Ritonavir (see p. 193)
Inhibitors	Allopurinol
	Cimetidine
	Erythromycin, clarithromycin
	Grapefruit juice (contains flavinoids)
	Imidazoles
	Quinolones
	Sulphonamides

*Hyperforin is one of the ingredients of the herbal product known as St John's Wort used by herbalists to treat depression. Although it is marketed as a licensed medicine, it is a reminder that drug interactions can occur with alternative, as well as conventional, medicines.

catabolism was shown to be a monogenetically inherited trait, involving 5–10% of Caucasian populations and leading to an exaggerated hypotensive response. Figure 16.1 shows the distribution of debrisoquin metabolic ratios in a UK population. Those with a ratio of more than 10 are homozygous slow metabolizers; those with a ratio of less than 0.5 are homozygous rapid metabolizers. Those with intermediate metabolic ratios are predominantly heterozygotes.

Table 16.6 Some genetic polymorphisms involving drug metabolism

Enzyme	Drug
P450	
Cytochrome CYP1A2	Amitriptyline
	Clozapine
Cytochrome CYP3A4	Quinidine
	Ciclosporin
	Lidocaine
	Verapamil
	Statins
	Protease inhibitors
Cytochrome CYP2C9	Warfarin
	Tolbutamide
	Phenytoin
	Glipizide
	Losartan
Cytochrome CYP2D6	Amitriptyline
	Venlafaxine
	SSRIs*
	Codeine
	Beta-blockers
	Flecainide
Cytochrome CYP2C19	Diazepam
	Omeprazole/lansoprazole
Plasma pseudocholinesterase	Succinylcholine
	Mivacurium
Thiopurine methyltransferase	Mercaptopurine
	Azathioprine
UDP-glucuronosyl transferase	Irinotecan
N-acetyl transferase	Procainamide
	Isoniazid
	Hydralazine

*Selective serotonin reuptake inhibitors.

A very large number of drugs – estimated at 15–25% of all medicines in use – are substrates for CYP2D6. The frequency of the variant alleles show racial variation and a small proportion of individuals have two or more copies of the active gene. The phenotypic consequences of the defective CYP2D6 include the increased risk of toxicity with those antidepressants or antipsychotics undergoing metabolism by this pathway. Conversely, in those with multiple copies of the active gene, there is extremely rapid rates of metabolism and therapeutic failure at conventional doses.

Warfarin is predominantly metabolized by CYP2C9. In most populations, between 2% and 10% are homozygous for an allele that results in low enzyme activity. Such individuals will therefore metabolize warfarin more slowly leading to higher plasma levels, a greater risk of bleeding, and a requirement for lower doses if the international normalized ratio (INR) is to be maintained within the therapeutic range.

Individual differences in the activity of thiopurine methyltransferase (TPMT) determine the dose of mercaptopurine and azathioprine that are used. TMPT activity is therefore undertaken routinely in children undergoing treatment for acute lymphatic leukaemia and patients with Crohn's disease (see p. 290).

Many drugs undergo metabolism by more than one member of the cytochrome P450 family. Individuals deficient in one enzyme may have normal, or even overexpressed, activities of others. Current knowledge (and cost) does not therefore permit predictions of an individual's dosage requirements, to a wide range of drugs, based on scans of all the genes concerned with drug metabolism. This may, though, become possible in the future, leading – in part – to the prospect of so-called 'personalized prescribing'.

Other causes of altered pharmacokinetics

Rates of hepatic drug clearance can also be influenced by environmental factors including diet, alcohol consumption and concomitant therapy with drugs capable of inducing or inhibiting (Table 16.5) drug metabolism. Hepatic drug clearance also decreases with age. By contrast, renal drug clearance does not show substantial variation between healthy individuals although it declines with age and in patients with intrinsic renal disease.

Pharmacodynamics

Pharmacodynamic sources of variability in the intensity of drug action are at least partly due to drug receptor and

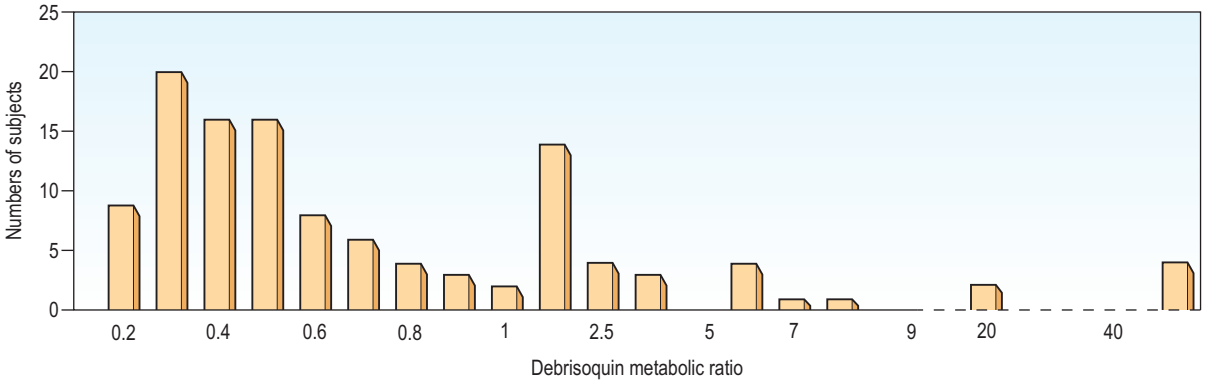


Fig. 16.1 **Distribution of debrisoquin metabolic ratios** (ratio of urinary debrisoquin:hydroxydebrisoquin after single 10 mg oral doses) in a normal UK population. By kind permission of Professor Ann Daly.

receptor polymorphisms (Table 16.7). At present, the only pharmacodynamic tests used routinely in clinical practice are the expressions of the oestrogen and HER2 receptors in women with breast cancer.

The prospect for 'personalized prescribing' will be enhanced further, when pharmacodynamic polymorphisms can be elicited by gene scanning. The interplay between pharmacokinetics and pharmacodynamics will influence drug selection and dosing.

Monitoring the effects of treatment

The combination of pharmacokinetic and pharmacodynamic causes of variability makes monitoring of the effects of treatment essential. Three approaches are used.

Pre-treatment dose selection

In patients with known, or suspected, impaired renal function it is usually possible to predict their dose requirements from the serum creatinine concentration. If treatment needs to be started before the serum creatinine concentration is available, or in patients with very advanced renal impairment, or if renal function is fluctuating, then start with conventional doses but be prepared to make adjustments within 24 hours. Frequent assessments of renal function, coupled with measurements of plasma drug concentrations, may be necessary.

Measuring plasma drug concentrations

For a few drugs, dosages can be effectively monitored by reference to their plasma concentrations (Table 16.8). This technique is only useful, however, if both the following criteria are fulfilled:

- There is a reliable and available drug assay.
- Plasma concentrations correlate well with therapeutic efficacy and toxicity.

Measuring drug effects

For many drugs, dosage adjustments are made in line with patients' responses. Monitoring can involve dose titration against a therapeutic end-point or a toxic effect. Objective measures (such as monitoring antihypertensive therapy by measuring blood pressure, or cytotoxic therapy with serial white blood cell counts) are most helpful, but subjective ones are necessary in many instances (as with antipsychotic therapy in patients with schizophrenia).

Table 16.7 Some pharmacodynamic genetic polymorphisms

	Drug	Drug effect
ACE	ACE inhibitors, e.g. ramipril	Blood pressure reduction
Bradykinin B_2 receptor	ACE inhibitors	Cough
β_2-Adrenoceptor	Salbutamol	Bronchodilatation
Dopamine receptors (D_2, D_3, D_4)	Haloperidol Clozapine Risperidone	Antipsychotic response Tardive dyskinesia Akathisia
Serotonin transporter	Fluoxetine Paroxetine	Antidepressant response
Oestrogen receptor-α	Oestrogens	Bone mineral density

AFFORDABILITY

The money available for healthcare varies widely across the world and there are marked differences (Fig. 16.2). All healthcare systems try to provide their populations with the highest standards of care within the resources they have at their disposal. The expenditure of large sums on a few people may deprive many of cost-effective remedies – a phenomenon known as the 'opportunity cost'.

In many countries cost-containment measures are encouraged (or mandated). For example, to reduce costs all drugs should be prescribed by their generic (approved) names rather than their 'brand' ones because, once their patents have expired, they are cheaper. Despite occasional claims to the contrary, generic products are required to go through the same stringent regulatory processes as their branded counterparts.

Some countries, including Australia, Canada and Britain, assess the cost effectiveness of new drugs (value for money) before they are available under their publicly funded healthcare systems.

ADVERSE DRUG REACTIONS

Adverse drug reactions (ADRs), defined as 'the unwanted effects of drugs occurring under normal conditions of use', are a significant cause of morbidity and mortality. Around 5% of acute medical emergencies are admitted with ADRs, and around 10–20% of hospital inpatients suffer an ADR during their stay. Unwanted effects of drugs are five to six times more likely in the elderly, compared to young adults; and the risk of an ADR rises sharply with the number of drugs administered.

Table 16.8 Drugs for which therapeutic drug monitoring is used

Drug	Therapeutic plasma concentration range	Toxic level	Optimum post-dose sampling time (hours)
Carbamazepine	20–50 μmol/L	>50 μmol/L	>8
Ciclosporin	50–200 μg/L	>200 μg/L	Pre-dose
Digoxin	1.3–2.6 nmol/L	>2.6 nmol/L	>8
Gentamicin	Trough < 2 mg/L	>14 mg/L	Pre-dose
	Peak 5–10 mg/L	>12 mg/L	>1
Lithium	0.6–1.0 mmol/L	>1.5 mmol/L	>10
Phenytoin	40–80 μmol/L	>80 μmol/L	>10
Theophylline	55–110 μmol/L	>110 μmol/L	>4
Vancomycin	15–20 mg/L	—	Pre-dose

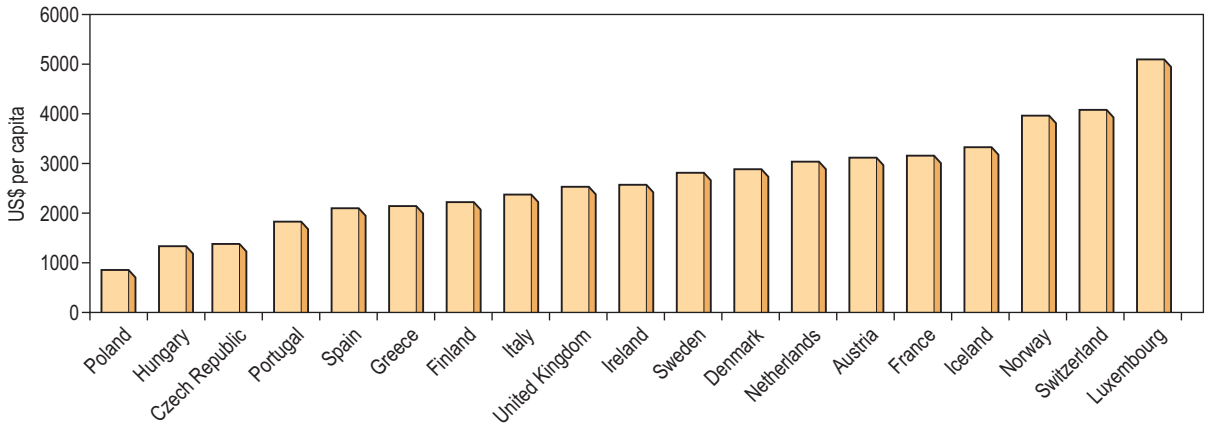


Fig. 16.2 **Annual expenditure on healthcare, as US$ per head of the population, in some developed countries.**

Table 16.9 Examples of adverse drug reactions

Drug	Adverse reaction
Type A (augmented)	
Anticoagulants	Bleeding
Insulin	Hypoglycaemia
Angiotensin-converting enzyme inhibitors/ antagonists	Hypotension
Antipsychotics	Acute dystonia and dyskinesia Parkinson's disease Tardive dyskinesia
Tricyclic antidepressants	Dry mouth
Amiodarone	Hyperthyroidism Hypothyroidism Pulmonary fibrosis
Cytotoxic agents	Bone marrow dyscrasias Cancer
Glucocorticoids	Osteoporosis
Type B (idiosyncratic)	
Benzylpenicillin Radiological contrast media	Anaphylaxis
Amoxicillin	Maculopapular rash
Sulphonamides Lamotrigine	Toxic epidermal necrolysis
Volatile anaesthetics Suxamethonium	Malignant hyperthermia
Diclofenac Isoflurane, sevoflurane Isoniazid Rifampicin Phenytoin	Hepatotoxicity

Classification

Two types of ADR are recognized.

Type A (augmented) reactions (Table 16.9) are:

- qualitatively normal, but quantitatively abnormal, manifestations of a compound's pharmacological or toxicological properties
- predictable from a compound's known pharmacological or toxicological actions
- generally dose-dependent
- usually common
- only occasionally serious.

Whilst some such reactions as hypotension with ACE inhibitors may occur after a single dose, others may develop only after months (pulmonary fibrosis with amiodarone) or years (second cancers with cytotoxic drugs).

Type B (idiosyncratic) reactions (Table 16.9) have no resemblance to the recognized pharmacological or toxicological effects of the drug. They are:

- qualitatively abnormal responses to the drug
- unpredictable from a compound's known pharmacological or toxicological actions
- generally dose-independent
- usually rare
- often serious.

Diagnosis

All ADRs mimic some naturally occurring disease, and the distinction between an iatrogenic aetiology and an event unrelated to the drug is often difficult. Although some effects are obviously iatrogenic (e.g. acute anaphylaxis occurring a few minutes after intravenous penicillin), many are less so. There are six characteristics that can help distinguish an adverse reaction from an event due to some other cause:

- *Appropriate time interval*. The time interval between the administration of a drug and the suspected adverse reaction should be appropriate. Acute anaphylaxis usually occurs within a few minutes of administration, whilst aplastic anaemia will only become apparent after a few weeks (because of the life-span of erythrocytes). Drug-induced malignancy, however, will take years to develop.
- *Nature of the reaction*. Some conditions (maculopapular rashes, angio-oedema, fixed drug eruptions, toxic epidermal necrolysis) are so typically iatrogenic that an adverse drug reaction is very likely.
- *Plausibility*. Where an event is a manifestation of the known pharmacological property of the drug, its recognition as a type A adverse drug reaction can be made (e.g. hypotension with an antihypertensive agent, or hypoglycaemia with an antidiabetic drug). Unless there have been previous reports in the literature, the recognition of type B reactions may be very difficult. The first cases of depression with isotretinoin, for example, were difficult to recognize as an ADR even though a causal association is now acknowledged.

- *Exclusion of other causes*. In some instances, particularly suspected hepatotoxicity, an iatrogenic diagnosis can only be made after the exclusion of other causes of disease.
- *Results of laboratory tests*. In a few instances, the diagnosis of an adverse reaction can be inferred from the plasma concentration (Table 16.8). Occasionally, an ADR produces diagnostic histopathological features. Examples include putative reactions involving the skin and liver.
- *Results of dechallenge and rechallenge*. Failure of remission when the drug is withdrawn (i.e. 'dechallenge') is unlikely to be an ADR. The diagnostic reliability of dechallenge, however, is not absolute: if the ADR has caused irreversible organ damage (e.g. malignancy) then dechallenge will result in a false-negative response. Rechallenge, involving re-institution of the suspected drug to see if the event recurs, is often regarded as an absolute diagnostic test. This is, in many instances, correct but there are two caveats. First, it is rarely ethically justified to subject a patient to further hazard. Second, some adverse drug reactions develop because of particular circumstances which may not necessarily be replicated on rechallenge (e.g. hypoglycaemia with an antidiabetic agent).

Management

As a general rule, type A reactions can usually be managed by a reduction in dosage whilst type B reactions almost invariably require the drug to be withdrawn (and never re-instituted).

Specific therapy is sometimes required for ADRs such as bleeding with warfarin (vitamin K), acute dystonias (benzatropine) or acute anaphylaxis (see Emergency Box 3.1, p. 73).

FURTHER READING

Relling M, Giacomini KM. Pharmacogenetics. In: Brunton LL, Lazo JS, Parker KL (eds) *Gilman & Goodman's The Pharmacological Basis of Therapeutics,* 11th edn. McGraw-Hill: New York, 2006, 93–115.

Wilkinson GR. Drug metabolism and variability among patients in drug response. *New England Journal of Medicine* 2005; **352**: 2211–2221.

Wynne HA. Sensible prescribing for older people. *Clinical Medicine* 2003; **3**: 409–412.

EVIDENCE-BASED MEDICINE

There is general acceptance that clinical practice should, as far as possible, be based on scientific evidence of benefit rather than theoretical speculation, anecdote or pronouncement.

One of the main applications of 'evidence-based medicine' is in therapeutics. Treatments should be introduced into, and used in, routine clinical care only if they have been demonstrated to be effective in formal clinical trials. Three approaches have been used:

- randomized controlled trials
- controlled observational trials
- uncontrolled observational studies.

Randomized controlled trials

Parallel group design. In this type of study, patients with a particular condition are given, prospectively, one of two (or more) treatments. Treatments are usually allocated randomly (a 'randomized' controlled trial). In order to reduce patient bias, the patients themselves are generally unaware of their treatment allocations (a 'single-blind' trial); and in order to reduce doctor bias the treatment allocation is also withheld from the investigators (a 'double-blind' trial). To recruit sufficient numbers of patients, and to examine the effects of treatment in different settings, it is often necessary to conduct the trial at several locations (a 'multicentre' trial). The 'gold standard' for demonstrating the efficacy of a treatment is, thus, the *prospective, randomized, double-blind, multicentre, controlled trial*.

Prospective randomized controlled trials are designed either to show that one treatment is better than another (a 'superiority' trial); or that one treatment is similar to another (an 'equivalence' trial). In a *superiority trial* the study treatment is usually compared to placebo, no treatment or to current standard practice. With drugs, the comparators may include different doses of the study ('active') drug in order to define the optimum treatment regimen. In an *equivalence trial* the treatment under study is usually compared to another treatment for the same condition.

Although RCTs were originally introduced to investigate the efficacy of drugs, the methodology is used for surgical procedures and medical devices.

Cross-over design. In some circumstances patients can receive both the active and the comparator, in random sequence, thus acting as their own control. Cross-over trials have the advantage that, since each patient is his/her own control, fewer numbers are required to demonstrate efficacy. Such designs, however, are usually only appropriate in circumstances where the drug is intended to produce improvement in a chronic condition. Examples include hypnotics, bronchodilators and antihypertensive agents.

Assessing randomized controlled trials

In assessing the relevance and reliability of an RCT a number of features need to be taken into account.

- *Was ethical approval obtained*? All clinical trials should have received approval, before the start of the study, from a properly constituted research ethics committee. In particular, patients taking part should have given their full and informed consent to participate.
- *Randomization*. In any RCT the method of randomization should be robust. In particular, the investigator should be unaware of which treatment a patient entering a trial will receive. This avoids selection bias.
- *Maintaining blindness*. Although, ideally, in RCTs neither the investigator nor the patient is aware of the treatment allocation until the end of the study, this is not always possible. Adverse drug reactions, for example, may make it obvious which treatment a patient has been given. Nevertheless maintaining 'blindness' is necessary where the outcome is subjective (e.g. relief of pain, alleviation of depression) if bias is to be avoided.
- *Were the treated and control groups comparable*? Were the treated and control groups similar in their 'baseline'

characteristics? Were they, for example, of similar age, severity and duration of illness? If not, are the differences likely to bias the results? Or has the statistical analysis (using analysis of covariance, or Cox's proportional hazards model) (see below) tried to adjust for them? Table 16.10 shows some of the baseline characteristics of a trial comparing prednisolone with placebo in the treatment of Bell's palsy (idiopathic facial paralysis).

- *Outcomes*. There are two ways to look at the outcomes of an RCT. One is to include only those who completed the study ('per protocol analysis') and the other is to include all patients from the time of randomization ('intention-to-treat analysis'). Ideally there should be no difference but in reality the results of a per protocol analysis are usually more advantageous to a treatment than an intention-to-treat analysis. The reason is that the intention-to-treat analysis will take account of patients who have withdrawn from the trial because of intolerance of the treatment or adverse drug reactions. It is therefore a much more robust approach. The results of the intention-to-treat analysis, in the trial of prednisolone in Bell's palsy, are shown in Table 16.10. The trial results indicate a high probability that treatment of Bell's palsy with prednisolone will increase the chances of a full recovery of facial nerve function.
- *Are the results generalizable*? Were the patients enrolled into the study a reasonable reflection of those likely to be treated in routine clinical practice (a so-called pragmatic trial)? Or were they a selected population that excluded significant patient groups (such as the elderly)? If the latter, view the results with caution.
- *Analysis of a superiority trial*. The aim of a superiority trial is to determine whether one treatment (or dose) is better (or more effective) than another. It is usual to estimate the probability that there is 'no difference' between the treatments; if this is less than 1 in 20 ($p < 0.05$) it is regarded as 'statistically significant'. There are two caveats. First, any difference may still be due to chance and it is better to await the results of two independent studies before adopting a new treatment. Second, a trial may show no 'statistically significant' difference, when one in fact exists, because too few patients have been included. The 'power' of a study (the number of patients needed in each treatment group to detect a predefined difference) should have been defined at the outset. If not, the results of the study should be interpreted with extreme care.
- *Effect size*. Considerable care needs to be taken in interpreting the magnitude of the reported differences in a superiority trial. The results of the well-designed trial in Table 16.10 show, very convincingly, that the treatment of Bell's palsy with prednisolone increases the likelihood of a complete recovery of facial nerve function, at 12 months, from 81.6% to 94.4%. One expression of the benefit of prednisolone, in this indication, can be derived from the *number needed to treat* (NNT). In the study shown in Table 16.10 the NNT to enable one patient to regain normal facial nerve function is 8.
- *Analysis of an equivalence trial*. The aim in this trial is to determine whether two (or possibly more) treatments produce similar benefits. During the design of such trials, it is necessary to decide what difference is unimportant and then to calculate the number of patients needed in order to have an 80% or 90%

Table 16.10 **Summary of a multicentre randomized placebo-controlled trial of prednisolone (25 mg twice daily), for 10 days, in the treatment of Bell's palsy**

	Baseline characteristics	
	Placebo (N = 245)	**Prednisolone (N = 251)**
Sex – no. (%)		
Male	118 (48.2)	135 (53.8)
Female	127 (51.8)	116 (46.2)
Mean age (years ± SD)	44.9 (±16.6)	43.2 (±16.2)
Score on House-Brackmann scale*	3.8 (±1.3)	3.5 (±1.2)
Time between onset of symptoms and start of treatment – no. (%)		
Within 24 hours	147 (60.0)	120 (47.8)
>24 to ≤48 hours	64 (26.1)	95 (37.8)
>48 to ≤72 hours	18 (7.3)	25 (10.0)
Unknown (but <72 hours)	16 (6.5)	11 (4.4)

	Results at 3 and 12 months		
	Placebo	**Prednisolone**	**P-value**
Grade 1 on House-Brackmann scale*– no. (%)			
At 3 months	152/239 (63.1)	205/247 (83.0)	<0.001
At 12 months	200/245 (81.6)	237/251 (94.4)	<0.001

*The House-Brackmann scale scores facial nerve function from 1 = normal to 6 = complete paralysis.
From Sullivan FM, Swan IRC, Donnan PT et al. Early treatment with prednisolone of aciclovir in Bell's palsy. *New England Journal of Medicine* 2007; **357**: 1598–1607.

Ask the authors
www.studentconsult.com

chance of showing this. In equivalence trials such power calculations show that the number of patients required is invariably greater than those needed for superiority trials. The results of equivalence trials are usually reported as a *'hazard ratio'* (the ratio of the response rate to the study treatment to that of the comparator) with its 95% confidence intervals. A hazard ratio of around unity (and with a confidence interval of – say – 0.9–1.1) would indicate that the two treatments were indeed likely to be equivalent. A hazard ratio of 0.5 (confidence interval 0.3–0.7) would suggest inequivalence, and superiority of the study treatment. By contrast, a hazard ratio of 2.0 (confidence interval 1.7–2.3) would suggest that the new treatment was inferior. In equivalence trials, it should be obvious that the comparator itself must have been shown to be effective.

- *Meta-analysis*. An analysis of all controlled trials that have been performed in a particular area can minimize random errors in the assessment of treatment effects because more patients and treatments are included than in any individual trial. Meta-analysis should be performed (and interpreted) carefully because of the heterogeneity of studies.

Controlled observational trials

Three types of observational study have been used to test the clinical effectiveness of therapeutic interventions – historical controlled trials, case-control studies and before-and-after studies.

Historical controlled trials. Despite the pre-eminence of the prospective randomized controlled trial there are many treatments that have never been subjected to this technique, yet their efficacy is unquestioned. Examples include insulin in the treatment of diabetic ketoacidosis, thyroxine for hypothyroidism, vitamin B_{12} in pernicious anaemia and defibrillation for ventricular fibrillation.

In a historical controlled trial the outcome in patients treated with the study drug is compared to that of previously untreated patients with the same disease. The circumstances when it is legitimate for a treatment to be accepted into routine use on the basis of favourable comparisons with historical controls are generally limited to the following:

- There should be a biologically plausible reason why the treatment might be effective. From a knowledge of the underlying nature of the condition, and the properties of the treatment, it should be reasonable to infer likely benefit.
- There should be no other effective form of treatment. If there is, then the study treatment should be compared with the alternative in a prospective, randomized controlled trial.
- The disease, if untreated, should result in death or permanent disability.
- The condition should have a known, and predictable, natural history.

New treatments that fulfil these stringent criteria are uncommon. Historical controlled trials are most frequently (though inappropriately) used in the assessment of, for example, new anticancer treatments; in most instances prospective, comparative, randomized, and controlled trials would be more informative.

Case-control studies. This type of study design compares patients with a particular condition (the 'cases') with those without (the 'controls'). The approach has predominantly been used to identify 'risk factors' for specific conditions such as lung cancer (smoking), sudden infant death syndrome (lying prone) or deep venous thrombosis (oral contraceptives). Such a study allows an estimation of the odds ratio (OR), e.g. the ratio of the probability of an event occurring to the probability of the event not occurring:

	Cases	Controls
Risk factor present	*a*	*c*
Risk factor absent	*b*	*d*

The odds ratio (OR) = $a \div b / c \div d$

An OR that is significantly greater than unity indicates a statistical association that may be causal. The OR for deep venous thrombosis and current use of oral contraceptives equals 2–4 (depending on the preparation): this indicates that the risk of developing a deep venous thrombosis on oral contraceptives is between 2 and 4 times greater than the background rate.

In some studies, the OR for a particular observation has been found to be significantly less than unity, suggesting 'protection' from the condition under study. Some studies of women with myocardial infarction indicated protection in those using hormone-replacement therapies but it has been subsequently shown that the result was due to bias. On the other hand, case-control studies have consistently shown that aspirin and other non-steroidal anti-inflammatory drugs are associated with a reduced risk of colon cancer. This seems to be a causal effect.

Case-control studies claiming to demonstrate the efficacy of a drug need to be interpreted with great care: the possibility of bias and confounding is substantial as was seen in the studies of hormone-replacement therapy and myocardial infarction. Confirmation from one or more RCTs is usually essential.

Before-and-after studies. It has sometimes been inferred that observed improvements seen in patients before, and after, the application of a particular treatment is evidence of efficacy. Such an approach is fraught with difficulties: the combination of a placebo effect, as well as regression to the mean, is likely to negate most studies using this type of design. Nevertheless there are some circumstances where genuine efficacy can be confidently observed with such designs: the consequences of hip replacement, and cataract surgery, are good examples. Such instances can be regarded as special examples of the use of implicit historical controls.

Uncontrolled observational studies

Uncontrolled case series cannot be considered as providing primary evidence of efficacy unless they are undertaken in circumstances that are virtually those of historical controlled trials. When used in this way to demonstrate clinical effectiveness their validity relies on the use of implicit historical trials. Case series can, however, sometimes be of value in demonstrating the generalizability of the results of RCTs.

Dangers

Drug trials are carried out in specific groups of selected patients under strict supervision. The results, particularly when dramatic, are often used outside these strict criteria. The dramatic effect of spironolactone in heart failure (30% reduction in all-cause mortality) has not always been replicated in clinical practice because the wrong patients have been treated, often with higher doses, leading to hyperkalaemia and death.

Evaluation of new drugs

New drugs are subjected to a vigorous programme of preclinical and clinical testing before they are licensed for general use (Table 16.11) and are also monitored for safety following licensing. Doctors are recommended to fill in yellow cards when they suspect an adverse reaction has taken place.

Table 16.11 Evaluation of new drugs

Phase I
Healthy human subjects (usually men)
First use in man
Evaluation of safety and toxicity
Pharmacokinetic assessment
Sometimes pharmacodynamic assessment
Approximately 100 subjects

Phase II
First assessment in patients
Safety and toxicity evaluated
Dose range identified
Pharmacokinetic and pharmacodynamic monitoring
Approximately 500 subjects

Phase III
Use in wider patient population
Efficacy main objective
Safety and toxicity also carefully monitored
Often multicentre trials
Approximately 2000 patients involved

Phase IV
Postmarketing surveillance
All patients prescribed drug monitored
Efficacy, safety and toxicity measured
Quantification of unusual drug adverse effects
Yellow Card and Prescription Event Monitoring
Often very large numbers of patients observed

FURTHER READING

Benson K, Hartz AJ. A comparison of observational studies and randomized trials. *New England Journal of Medicine* 2000; **342**: 1878–1886.

Collins R, MacMahon S. Reliable assessment of the effects of treatment on mortality and major morbidity. I Clinical trials, II Observational studies. *Lancet* 2001; **357**: 373–380, 455–462.

Concato J, Shah N, Horwitz RI. Randomized controlled trials, observational studies and the hierarchy of research designs. *New England Journal of Medicine* 2000; **342**: 1887–1892.

Treating individuals (a series of articles starting in January 2005). *Lancet* **365**: 826.

von Elm E, Altman DG, Egger M et al. The Strengthening the Reporting of Observational studies in Epidemiology (STROBE) Statement. *Lancet* 2007; **370**: 1453–1457.

STATISTICAL ANALYSES

The relevance of statistics is not confined to those who undertake research but also to anyone who wants to understand the relevance of research studies to their clinical practice.

The average

Clinical studies may describe, quantitatively, the value of a particular variable (e.g. height, weight, blood pressure, haemoglobin) in a sample of a defined population. The 'average' value (or 'central tendency' in statistical language) can be expressed as the mean, median or mode depending on the circumstances:

- The *mean* is the average of a distribution of values that are grouped symmetrically around the central tendency.
- The *median* is the middle value of a sample. It is used, particularly, where the values in a sample are asymmetrically distributed around the central tendency.
- The *mode* is the interval, in a frequency distribution of values, that contains more values than any other.

In a symmetrically distributed population the mean, median and mode are the same.

The average value of a sample, on its own, is of only modest interest. Of equal (and often greater) relevance is the confidence we can place on the sample average as truly reflecting the average value of the population from which it has been drawn. This is most often expressed as a *confidence interval*, which describes the probability of a sample mean being a certain distance from the population mean. If, for example, the mean systolic blood pressure of 100 undergraduates is 124 mmHg, with a 95% confidence interval of ± 15 mmHg, we can be confident that if we replicated the study 100 times the value of the mean would be within the range 109–139 mmHg on 95 occasions. It is intuitively obvious that the larger the sample the smaller will be the size of the confidence interval.

Correlation

In clinical studies two, or more, independent variables may be measured in the same individuals in a sample population (e.g. weight and blood pressure). The degree of correlation between the two can be investigated by calculating the *correlation coefficient* (often abbreviated to '*r*'). The correlation coefficient measures the degree of association between the two variables and may range from 1 to –1. If $r = 1$, there is complete and direct concordance between the two variables: if $r = -1$, there is complete but inverse concordance; and where $r = 0$ there is no concordance.

Statistical tables are available to inform investigators as to the probability that *r* is due to chance. As in other areas of statistics, if the probability is less than 1 in 20 ($p < 0.05$) then by custom and practice it is regarded as '*statistically significant*'. There are, however, two caveats. First, the 1 in 20 rule is a convention and does not exclude the possibility that a presumed association is due to chance. Second, the

fact that there is an association between two variables does not necessarily mean that it is causal. A correlation between blood pressure and weight, with $r = 0.75$ and $p < 0.05$, does not mean that weight has a direct effect on blood pressure (or vice versa).

Correlation analyses can become complicated. The simplest (least squares regression analysis) presumes a straight-line relationship between the two variables. More complicated techniques can be used to estimate r where a non-linear relationship is presumed (or assumed); where the distributions deviate from normal; where the scales of one or both variables are intervals or ranks; or where a correlation between three or more variables is sought.

Hypothesis testing

Much of statistics is concerned with testing hypotheses. The basic assumption – known as the '*null hypothesis*' – is that there is no difference between the groups. The reason for this confusing terminology is that statistical techniques are designed to assess the extent to which a zero difference might be due to the play of chance (such as a sampling 'error'). In the analysis of the RCT in Table 16.10, the null hypothesis asserts that there is no difference between the results in patients treated with placebo and those on active treatment. Statistical tests are used to determine the probability that the observed difference is due to chance. Where this is less than 1 in 20 ($p < 0.05$) it is described as 'statistically significant'. Again, the 1 in 20 rule is arbitrary but is a convention that is widely adopted.

The choice of statistical test to examine the null hypothesis is a complicated one. It is dependent on the type of data collected (whether it is ordinal, cardinal or categorical); and whether it conforms to a normal (parametric) or other (non-parametric) distribution. Those most commonly used include *Student's t test (parametric)* and *chi-squared* (χ^2) test, analysis of variance and various tests for non-parametric data. In some circumstances it is possible to use the confidence intervals of the means of two (or more) groups to test the null hypothesis.

When the results of a statistical test indicate that it is reasonable to reject the null hypothesis, and that the probability of the results being due to chance is less than 1 in 20 ($p < 0.05$), it means that 95 times out of 100 we will be right, but that 5 times out of 100 we will be wrong. Statisticians call erroneous rejection of the null hypothesis a *type 1 error*. In this situation the null hypothesis is actually true, although we believe it to be false. Erroneous acceptance of the null hypothesis when there is indeed a difference is known as a *type 2 error*. Type 1 errors can be reduced by requiring a higher level of significance (e.g. $p < 0.01$ or $p < 0.005$) but can never be absolutely excluded. Type 2 errors are usually the result of too few participants in the study and can be avoided by estimating the numbers needed to examine specific levels of difference before the trial is started. Such estimates are known as '*power calculations*'.

Other statistical techniques

Statisticians have developed a range of sophisticated methods to handle a wide variety of biomedical problems. Unless an investigator is supremely (and usually unwisely) confident it is wise to seek professional advice in analysing numerical data that look complicated. In doing so, it is invariably wiser to do so at the time the study is being designed rather than after the results have been generated!

FURTHER READING

Armitage P, Berry G, Matthews JNS. *Statistical Methods in Medical Research*. Oxford: Blackwell Science, 2002.

Bland M. *An Introduction to Medical Statistics*, 3rd edn. Oxford: Oxford University Press, 2000.

Matthews JNS. *Introduction to Randomized Controlled Trials*, 2nd edn. London: Chapman & Hall, 2006.

Schulz KF, Grimes DA. Epidemiology I. *Lancet* 2005; **365**: 1348–1353.

INFORMATION

Pharmacotherapy moves at a very rapid pace and it is impossible for anyone to keep up with contemporary advances. Details of current prescribing advice can be found in:

- the Summary of Product Characteristics (SmPCs) produced by manufacturers but vetted by the UK drug regulatory authority (the Medicines and Healthcare Products Regulatory Agency)
- the British National Formulary (BNF) produced, jointly, by the British Medical Association and the Royal Pharmaceutical Society. Many countries have their own formularies
- the Technology Appraisals Guidance series from the National Institute for Health and Clinical Excellence (NICE).

Advice on the management of individual conditions, in the form of clinical guidelines (systematically developed statements to assist practitioner and patient decisions about appropriate healthcare for specific clinical circumstances), can be accessed from:

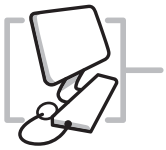
- the National Institute for Clinical Excellence (NICE): http://www.nice-org.uk
- the Scottish Intercollegiate Guidelines Network (SIGN): http://www.sign.ac.uk/
- the National Guidelines Clearing House in the USA: http://www.guideline.gov.

Patient information leaflets are supplied with all prescribed medication.

POISONING

The nature of the problem

Exposure to a substance is often equated with poisoning. However, absorption is necessary for there to be a toxic effect and, even if this occurs, poisoning does not necessarily result, because the amount absorbed may be too small. In developed countries, poisoning causes approximately 10% of acute hospital medical presentations. In such cases poisoning is usually by self-administration of prescribed and over-the-counter medicines, or illicit drugs. Poisoning in children aged less than 6 months is most commonly iatrogenic and involves overtreatment with, for example, paracetamol.

www.studentconsult.com
More online


Children between 8 months and 5 years of age also ingest poisons accidentally, or they may be administered deliberately to cause harm, or for financial or sexual gain. Occupational poisoning as a result of dermal or inhalational exposure to chemicals is a common occurrence in the developing world and still occurs in the developed world. Sometimes inappropriate treatment of a patient by a doctor is responsible for the development of poisoning, for example, in the case of digoxin toxicity.

In adults, self-poisoning is commonly a 'cry for help'. Those involved are most often females under the age of 35 who are in good physical health. They take an overdose in circumstances where they are likely to be found, or in the presence of others. In those older than 55 years of age, men predominate and the overdose is usually taken in the course of a depressive illness or because of poor physical health.

The type of agent taken in overdose is also heavily influenced by availability and culture. In the UK, paracetamol poisoning is responsible for approximately one-third of all admissions, whereas in Sri Lanka, for example, the agents ingested are more often pesticides or plants, such as oleander, and in South India, copper sulphate is a problem. In addition, ingestion of heating fuels (e.g. petroleum distillates), antimalarials, anti-tuberculous drugs and traditional medicine is reported frequently in the developing world.

A third of patients admitted with an overdose in the UK state that they are unaware of the toxic effects of the substance involved; the majority take whatever drug is easily available at home (Box 16.1). Studies of the agents involved reveal that:

- Acute overdoses usually involve more than one agent.
- Alcohol is the most commonly implicated second agent in mixed self-poisonings – 60% of men and 45% of women consume some alcohol at the same time as the drug.
- There is often a poor correlation between the drug history and the toxicological analytical findings. Therefore, a patient's statement about the type and amount of drug ingested cannot always be relied upon.

The majority of cases of self-poisoning do not require intensive medical management, but all patients require a sympathetic and caring approach, a psychiatric and social assessment and, sometimes, psychiatric treatment. However, as the majority of patients ingest relatively non-toxic agents, receive good supportive care and, when appropriate, the administration of specific antidotes, the in-hospital mortality in most developed countries is now less than 1%. Fatalities in the UK are due predominantly to carbon monoxide, antidepressants, paracetamol, analgesic combinations containing paracetamol and an opioid, heroin, methadone or cocaine. Deaths from poisoning in children are usually accidental and due to inappropriate storage of drugs such as digoxin and quinine and from drugs of abuse purchased or prescribed for a parent/carer.

Box 16.1 Prevention of self-poisoning

Patients usually take what is readily available at home.

- Small amounts only of drugs should be bought
- Foil-wrapped drugs are less likely to be taken
- Keep drugs in a safe place
- Keep drugs and liquids in their original containers
- Child-resistant drug containers should be used
- Doctors should be careful in prescribing all drugs
- Prescriptions for any susceptible patient (e.g. the depressed) must be monitored carefully
- Household products should be labelled and kept safely, away from children

Self-poisoning can kill

All people must be aware of the dangers of drugs and chemicals. Education on safe storage and careful handling of household and workplace chemicals is necessary on a continual basis.

THE APPROACH TO THE PATIENT (Table 16.12)

History

More than 80% of adults are conscious on arrival at hospital in the UK and the diagnosis of self-poisoning can usually be made from the history (Table 16.12). In the unconscious patient a history from friends or relatives is helpful, and the diagnosis can often be inferred from the medicine containers or a 'suicide note' brought by the paramedics. It should be emphasized that in any patient with an altered level of consciousness, drug overdose must always be considered in the differential diagnosis.

Examination

On arrival at hospital the patient must be assessed urgently (**A**irways, **B**reathing and **C**irculation). The following should be evaluated:

- *Level of consciousness* – The Glasgow Coma Scale should be used (see p. 1123).
- *Ventilation* – Pulse oximetry can be used to measure oxygen saturation. The displayed reading may be inaccurate when the saturation is below 70%, there is poor peripheral perfusion and in the presence of carboxyhaemoglobin and methaemoglobin. Only measurement of arterial blood gases will indicate the presence both of hypercapnia and hypoxia.
- *Blood pressure and pulse rate.*
- *Pupil size and reaction to light.*
- Evidence of intravenous drug abuse.
- A *head injury* complicating poisoning.

If the patient is unconscious the following should also be checked:

- *Cough and gag reflex* – present or absent.
- *Temperature* – measured with a low-reading rectal thermometer.

Some of the physical signs that may aid identification of the agents responsible for poisoning are shown in Table 16.13. The cluster of features on presentation may be distinctive and diagnostic. For example, sinus tachycardia, fixed

Table 16.12 Diagnostic process in acute poisoning

- Obtain history, if possible, from the patient, relative, friend or paramedics
- Is there circumstantial evidence of an overdose?
- Are the circumstances in which the patient has been found suggestive?
- Was a suicide note left?
- Are the symptoms suggestive of an overdose?
- Do the physical signs suggest an overdose (Tables 16.13 and 16.14)

Table 16.13 Some physical signs of poisoning

Features	Likely poisons
Constricted pupils (miosis)	Opioids, organophosphorus insecticides, nerve agents
Dilated pupils (mydriasis)	Tricyclic antidepressants, amfetamines, cocaine, antimuscarinic drugs
Divergent strabismus	Tricyclic antidepressants
Nystagmus	Carbamazepine, phenytoin
Loss of vision	Methanol, quinine
Papilloedema	Carbon monoxide, methanol
Convulsions	Tricyclic antidepressants, theophylline, opioids, mefenamic acid, isoniazid, amfetamines
Dystonic reactions	Metoclopramide, phenothiazines
Delirium and hallucinations	Amfetamines, antimuscarinic drugs, cannabis, recovery from tricyclic antidepressant poisoning
Hypertonia and hyper-reflexia	Tricyclic antidepressants, antimuscarinic drugs
Tinnitus and deafness	Salicylates, quinine
Hyperventilation	Salicylates, phenoxyacetate herbicides, theophylline
Hyperthermia	Ecstasy (MDMA), salicylates
Blisters	Usually occur in comatose patients

MDMA: 3,4-methylenedioxymethamfetamine.

Table 16.14 Common feature clusters in acute poisoning

Feature clusters	Poisons
Coma, hypertonia, hyperreflexia, extensor plantar responses, myoclonus, strabismus, mydriasis, sinus tachycardia	Tricyclic antidepressants; less commonly antihistamines, orphenadrine, thioridazine
Coma, hypotonia, hyporeflexia, plantar responses (flexor or non-elicitable), hypotension	Barbiturates, benzodiazepine and alcohol combinations, tricyclic antidepressants
Coma, miosis, reduced respiratory rate	Opioid analgesics
Nausea, vomiting, tinnitus, deafness, sweating, hyperventilation, vasodilatation, tachycardia	Salicylates
Hyperthermia, tachycardia, delirium, agitation, mydriasis	Ecstasy (MDMA) or other amfetamine
Miosis, hypersalivation, bronchorrhoea	Organophosphorus and carbamate insecticides, nerve agents

dilated pupils, exaggerated tendon reflexes, extensor plantar responses and coma suggest tricyclic antidepressant poisoning (Tables 16.13 and 16.14).

PRINCIPLES OF MANAGEMENT (Table 16.15)

Table 16.15 Management strategy in acute poisoning

- Provide supportive treatment
- Is the use of an antidote appropriate? (Table 16.16)
- Is it appropriate to attempt to reduce poison absorption?
- Is it appropriate to perform toxicological investigations?
- Will non-toxicological investigations assist? (Table 16.17)
- Should urine alkalinization, multiple-dose activated charcoal, haemodialysis or haemoperfusion be employed to increase poison elimination?

Most patients with self-poisoning require only general care and support of the vital systems. However, for a few drugs additional therapy is required.

CARE OF THE UNCONSCIOUS PATIENT (see also p. 1124)

In all cases the patient should be nursed in the lateral position with the lower leg straight and the upper leg flexed; in this position the risk of aspiration is reduced. A clear passage for air should be ensured by the removal of any obstructing object, vomit or dentures, and by backward pressure on the mandible. Nursing care of the mouth and pressure areas should be instituted. Immediate catheterization of the bladder in unconscious patients is usually unnecessary as it can be emptied by gentle suprapubic pressure. Insertion of a venous cannula is usual, but administration of intravenous fluids is unnecessary unless the patient has been unconscious for more than 12 hours or is hypotensive.

Ventilatory support

If respiratory depression is present, as determined by pulse oximetry or preferably by arterial blood gas analysis, an oropharyngeal airway should be inserted, and supplemented oxygen should be administered. Pulse oximetry alone will not detect hypercapnia. Loss of the cough or gag reflex is the prime indication for intubation. The gag reflex can be assessed by positioning the patient on one side and making him or her gag using a suction tube. In many severely poisoned patients the reflexes are depressed sufficiently to allow intubation without the use of sedatives or relaxants. The complications of endotracheal tubes are discussed on page 914. If ventilation remains inadequate after intubation, as shown by hypoxaemia and hypercapnia, intermittent positive-pressure ventilation (IPPV) should be instituted.

Cardiovascular support

Although hypotension (systolic blood pressure below 80 mmHg) is a recognized feature of acute poisoning, the classic features of shock – tachycardia and pale cold skin – are observed only rarely.

Hypotension and shock may be caused by:

- a direct cardio-depressant action of the poison (e.g. beta-blockers, calcium channel-blockers, tricyclic antidepressants)

- vasodilation and venous pooling in the lower limbs (e.g. ACE inhibitors, phenothiazines)
- decrease in circulating blood volume because of gastrointestinal losses (e.g. profuse vomiting in theophylline poisoning), increased insensible losses (e.g. salicylate poisoning), increased renal losses (e.g. poisoning due to diuretics) and increased capillary permeability.

Hypotension may be exacerbated by coexisting hypoxia, acidosis and dysrhythmias. In patients with marked hypotension, volume expansion with crystalloids should be used, guided by monitoring of central venous pressure (CVP). Urine output (aiming for 35–50 mL/h) is also a useful guide to the adequacy of the circulation. If a patient fails to respond to the above measures, more intensive therapy is required. In such patients, it is helpful to undertake invasive haemodynamic monitoring to confirm that adequate volume replacement has been administered. Volume replacement and the use of inotropes are discussed on page 905.

Systemic hypertension can be caused by a few drugs when taken in overdose. If this is mild and associated with agitation, a benzodiazepine may suffice. In more severe cases, for example those due to a monoamine oxidase inhibitor, there may be a risk of arterial rupture, particularly intracranially. To prevent this, an α-adrenergic blocking agent (e.g. phentolamine, 5 mg i.v. every 10–15 minutes), or intravenous isosorbide dinitrate 2–10 mg/hour up to 20 mg/hour if necessary, or sodium nitroprusside 0.5–1.5 μg/kg/min by intravenous infusion, should be administered until the blood pressure is controlled.

Arrhythmias occur, e.g. tachyarrythmias following ingestion of a tricyclic antidepressant or theophylline; bradyarrythmias with digoxin poisoning. Known arrhythmogenic factors such as hypoxia, acidosis and hypokalaemia should be corrected. All patients with shock should have ECG monitoring.

Other problems

Hypothermia and hyperthermia

A rectal temperature below 35°C is a recognized complication of poisoning, especially in older patients or those who are comatose. The patient should be covered with a 'space blanket' and, if necessary, given intravenous and intragastric fluids at normal body temperature. The administration of heated (37°C), humidified oxygen delivered by face mask is also useful.

Rarely, body temperature may increase to potentially fatal levels after poisoning with central nervous system stimulants such as cocaine, amfetamines including ecstasy (MDMA), monoamine oxidase inhibitors or theophylline. Muscle tone is often increased and convulsions and rhabdomyolysis are common. Cooling measures, sedation with diazepam and, in severe cases, i.v. dantrolene 1 mg/kg body weight should be given.

Skin blisters

Skin blisters may be found in poisoned patients who are, or have been, unconscious. Such lesions are not diagnostic of specific poisons, but are sufficiently common in poisoned patients (and sufficiently uncommon in patients unconscious from other causes) to be of diagnostic value.

Rhabdomyolysis

Rhabdomyolysis can occur from pressure necrosis in drug-induced coma, or it may complicate, for example, ecstasy (MDMA) abuse in the absence of coma. Patients with rhabdomyolysis are at risk of developing, firstly, acute kidney injury from myoglobinaemia, particularly if they are hypovolaemic and have an acidosis and, secondly, wrist or ankle drop from the development of a compartment syndrome (see p. 516).

Convulsions

These may occur, for example, in poisoning due to tricyclic antidepressants, mefenamic acid or opioids. Usually the seizures are short-lived but, if they are prolonged, diazepam 10–20 mg i.v. or lorazepam 4 mg i.v. should be administered. Persistent fits must be controlled rapidly to prevent severe hypoxia, brain damage and laryngeal trauma. If diazepam or lorazepam in repeated dose is ineffective, the patient should also receive a loading dose of phenytoin (18 mg/kg) administered intravenously at not more than 50 mg/min, with ECG monitoring.

Stress ulceration and bleeding

Measures to prevent stress ulceration of the stomach should be started on admission in all patients who are unconscious and require intensive care. A proton pump inhibitor should be administered intravenously.

'Body packers' and 'body stuffers'

'Body packers' (also known as 'mules') are individuals who attempt to smuggle drugs across national borders or into prisons by ingesting multiple small packets that are swallowed for later retrieval from vomit or faeces, or are inserted into the vagina or rectum. Body packers often swallow large numbers of packages, with each package containing a potentially fatal dose of drug. Acute intestinal obstruction may result, and poisoning is a hazard if a packet bursts. The commonest drugs carried are cocaine, heroin, amphetamines and cannabis, though most deaths are opiate-induced.

Drugs are packaged in 2–4 layers of wrapping (condoms, latex gloves, balloons) and seen on plain X-ray as circular white densities. Alternatively, drugs may be covered by 6–7 layers of tied tubular latex (less susceptible to breakage) and seen on X-ray as soft tissue densities with surrounding gas haloes. Hardened drug paste may be wrapped in aluminium foil and overwrapped with 3–5 layers of tubular latex; these packages are least likely to break and may be difficult to detect on X-ray.

'Body stuffers', by contrast, are individuals who ingest drugs to avoid detection or arrest by the police. These drugs are usually either unpackaged or poorly packaged, and bodystuffers are at risk of developing features of toxicity soon after the drug has been ingested.

Management

All suspected cases should have an abdominal X-ray on presentation. CT of the abdomen should be arranged as soon as possible. A urine screen for drugs of abuse should be performed on admission. A screen that is positive for one or more drugs of abuse suggests ingestion of the abused drug in the previous few days, or a leaking packet. A negative screen strongly suggests that no packet is leaking. Packets in the stomach should not be removed by endoscopy or by inducing emesis, as these are potentially dangerous. Packets

in the vagina can usually be removed manually. With packets in the small bowel (if there is no clinical, analytical or radiological evidence to support leakage), sorbitol or lactulose to encourage transit through the gut is appropriate. Alternatively, for faster results, whole-bowel irrigation using polyethylene glycol electrolyte solutions can be used. Liquid paraffin should not be used because it can weaken the rubber, leading to bursting of the packets. Activated charcoal is contraindicated as it induces constipation when used in substantial doses. Packets in the colon or rectum are probably best managed by giving sorbitol or lactulose and allowing them to pass spontaneously, with least risk of rupture. Immediate surgery is indicated if acute intestinal obstruction develops, or when packets can be seen radiologically and there is radiological, clinical or analytical evidence to suggest leakage, particularly if the drug involved is a CNS stimulant (e.g. cocaine).

SPECIFIC MANAGEMENT

Antidotes

Specific antidotes are available for only a small number of poisons (Table 16.16).

Table 16.16 Antidotes of value in poisoning

Poison	Antidotes
Aluminium (aluminum)	Desferrioxamine (deferoxamine)
Arsenic	DMSA*, dimercaprol
Benzodiazepines	Flumazenil
β-adrenoceptor blocking drugs	Atropine, glucagon
Calcium channel blockers	Atropine
Carbamate insecticides	Atropine
Carbon monoxide	Oxygen
Copper	D-penicillamine, DMPS†
Cyanide	Oxygen, dicobalt edetate, hydroxocobalamin, sodium nitrite, sodium thiosulphate
Diethylene glycol	Ethanol, fomepizole
Digoxin and digitoxin	Digoxin-specific antibody fragments
Ethylene glycol	Ethanol, fomepizole
Hydrogen sulfide	Oxygen
Iron salts	Desferrioxamine
Lead (inorganic)	Sodium calcium edentate, succimer (DMSA*)
Methaemoglobinaemia	Methylthioninium chloride (Methylene blue)
Methanol	Ethanol, fomepizole
Mercury (inorganic)	Unithiol (DMPS†)
Nerve agents	Atropine, HI-6, obidoxime, pralidoxime
Oleander	Digoxin-specific antibody fragments
Opioids	Naloxone
Organophosphorus insecticides	Atropine, HI-6, obidoxime, pralidoxime
Paracetamol	N-acetylcysteine
Thallium	Berlin (Prussian) blue
Warfarin and other anticoagulants	Phytomenadione (Vitamin K)

*Dimercaptosuccinic acid.
†Dimercaptopropanesulphonate.

Antidotes may exert a beneficial effect by:

- *Forming an inert complex* with the poison (e.g. desferrioxamine (deferoxamine), D-penicillamine, dicobalt edetate, digoxin-specific antibody fragments, dimercaprol, HI-6, hydroxocobalamin, obidoxime, pralidoxime, protamine, Prussian (Berlin) blue, sodium calcium edetate, succimer [DMSA], unithiol [DMPS]).
- *Accelerating detoxification* of the poison (e.g. *N*-acetylcysteine, sodium thiosulphate).
- *Reducing the rate of conversion* of the poison to a more toxic compound (e.g. ethanol, fomepizole).
- *Competing with the poison* for essential receptor sites (e.g. oxygen, naloxone, phytomenadione).
- *Blocking essential receptors* through which the toxic effects are mediated (e.g. atropine).
- *Bypassing the effect* of the poison (e.g. oxygen, glucagon).

Reducing poison absorption

To reduce poison absorption through the lungs, remove the casualty from the toxic atmosphere, making sure that rescuers themselves are not put at risk. Contaminated clothing should be removed to reduce dermal absorption and contaminated skin washed thoroughly with soap and water.

Gut decontamination. While it appears logical to assume that removal of unabsorbed drug from the gastrointestinal tract will be beneficial (gut decontamination), the efficacy of current methods remains unproven and efforts to remove small amounts of non-toxic drugs are clinically not worthwhile or appropriate.

Gastric lavage. Gastric lavage should not be employed as the amount of marker removed by gastric lavage is highly variable and diminishes with time; it may cause significant morbidity. It should only be performed if a patient has ingested a potentially life-threatening amount of a poison, e.g. iron, and the procedure can be undertaken within 60 minutes of ingestion. Intubation is required if airway protective reflexes are lost. Lavage is also contraindicated if a hydrocarbon with high aspiration potential or a corrosive substance has been ingested.

In some countries gastric lavage is still widely practised.

Syrup of ipecacuanha contains two alkaloids, emetine and cephaeline, which induce vomiting by a central, as well as a local (emetine) action. It should not be used as the amount removed is highly variable, diminishes with time and there is no evidence that it improves the outcome of poisoned patients.

Single-dose activated charcoal. Activated charcoal has a highly developed internal pore structure which is able to adsorb a wide variety of compounds. Exceptions are strong acids and alkalis, ethanol, ethylene glycol, iron, lithium, mercury and methanol.

In studies in volunteers given 50 g activated charcoal, the mean reduction in absorption was 40%, 16% and 21% at 60 minutes, 120 minutes and 180 minutes respectively after ingestion. Based on these studies, activated charcoal should

be given in those who have ingested a potentially toxic amount of a poison (known to be adsorbed by charcoal) up to 1 hour previously. There are insufficient data to support or exclude its use after 1 hour. There is no evidence that administration of activated charcoal improves the clinical outcome.

Cathartics alone have no role in the management of the poisoned patient and are not recommended as a method of gut decontamination. Their use in combination with activated charcoal has produced conflicting data and there are no data on the reduction of the bioavailability of drugs or improvement in outcome of poisoned patients.

Whole bowel irrigation (WBI) requires the insertion of a nasogastric tube into the stomach and the introduction of polyethylene glycol electrolyte solution 1500–2000 mL/h in an adult. WBI is continued until the rectal effluent is clear.

WBI should not be used routinely in the management of the poisoned patient as there are no controlled clinical trials and no conclusive evidence that it improves the outcome. Based on volunteer studies, WBI may be used for potentially toxic ingestions of sustained-release or enteric-coated drugs. There is insufficient data to support or exclude the use of WBI for potentially toxic ingestions of iron, lead, zinc, or packets of illicit drugs. WBI is contraindicated in patients with bowel obstruction, perforation, ileus, and in patients with haemodynamic instability or compromised unprotected airways.

FURTHER READING

Barceloux D, McGuigan M, Hartigan-Go K et al. Position paper: cathartics. *Clinical Toxicology* 2004; **42**: 243–253.

Chyka PA, Seger D, Krenzelok EP et al. Position paper: single-dose activated charcoal. *Clinical Toxicology* 2005; **43**: 61–87.

Krenzelok EP, McGuigan M, Lheureux P et al. Position paper: ipecac syrup. *Clinical Toxicology* 2004; **42**: 133–143.

Kulig K, Vale JA. Position paper: gastric lavage. *Clinical Toxicology* 2004; **42**: 933–943.

Tenenbein M, Lheureux P. Position paper: whole bowel irrigation. *Clinical Toxicology* 2004; **42**: 843–854.

Increasing poison elimination

Multiple-dose activated charcoal (MDAC) involves the repeated administration of oral activated charcoal to increase the elimination of a drug that has already been absorbed into the body. Drugs are secreted in the bile and re-enter the gut by passive diffusion if the concentration in the gut is lower than that in the blood. The rate of passive diffusion depends on the concentration gradient and the intestinal surface area, permeability and blood flow. Occasionally, drugs such as digoxin may be secreted actively by the intestinal mucosa, though this effect is small. Activated charcoal will bind any drug that is in the gut lumen.

Elimination of drugs with a small volume of distribution (<1 litre/kg), low pK_a (which maximizes transport across membranes), low binding affinity and prolonged elimination half-life following overdose is particularly likely to be enhanced by MDAC. MDAC also improves total body clearance of the drug when endogenous processes are compromised by liver and/or renal failure.

Although MDAC has been shown to significantly increase drug elimination, it has not reduced morbidity and mortality in controlled studies. At present, MDAC should only be used in patients who have ingested a life-threatening amount of carbamazepine, dapsone, phenobarbital, quinine and theophylline.

Dosage. In *adults*, charcoal should be administered in an initial dose of 50–100 g and then at a rate of not less than 12.5 g/h, preferably via a nasogastric tube. Smaller initial doses (10–25 g) can be used in *children* because, generally, smaller overdoses have been ingested and the capacity of the gut lumen is smaller. If the patient has ingested a drug that induces protracted vomiting (e.g. theophylline), intravenous ondansetron is effective as an antiemetic and thus enables administration of MDAC.

Urine alkalinization. Increasing the urine pH enhances elimination of salicylate, phenobarbital, chlorpropamide and chlorophenoxy herbicides (e.g. 2,4-dichlorophenoxyacetic acid) by mechanisms which are not clearly understood. However, with the exception of moderate and severe salicylate poisoning, urine alkalinization is not recommended as first-line therapy for poisoning with these agents as MDAC is superior for phenobarbital, and supportive care is invariably adequate for chlorpropamide. A substantial diuresis is required in addition to urine alkalinization to achieve clinically relevant elimination of chlorophenoxy herbicides.

Urine alkalinization is a metabolically invasive procedure requiring frequent biochemical monitoring and medical and nursing expertise. Before commencing urine alkalinization, correct plasma volume depletion, electrolytes (administration of sodium bicarbonate exacerbates pre-existing hypokalaemia) and metabolic abnormalities. Urine alkalinization is a metabolically invasive procedure requiring frequent biochemical monitoring and medical and nursing expertise. Before commencing urine alkalinization, it is necessary to correct plasma volume depletion, electrolytes (administration of sodium bicarbonate exacerbates pre-existing hypokalaemia) and metabolic abnormalities. sufficient bicarbonate is administered to ensure that the pH of the urine, which is measured by narrowrange indicator paper or a pH meter, is more than 7.5 and preferably close to 8.5. In one study sodium bicarbonate 225 mmol was the mean amount required initially. This is most conveniently administered as 225 mL of an 8.4% solution (1 mmol bicarbonate/mL) i.v. over 1 hour.

Haemodialysis, haemodialfiltration and haemoperfusion. Haemodialysis, haemodialfiltration (haemodialysis and haemofiltration done in parallel) and haemoperfusion are of little value in patients poisoned with drugs with large volumes of distribution (e.g. tricyclic antidepressants), because the plasma contains only a small proportion of the total amount of drug in the body. These methods to increase poison elimination are indicated in patients with both severe clinical features and high plasma toxin concentrations.

Haemodialysis significantly increases elimination of ethanol, ethylene glycol, isopropanol, lithium, methanol and salicylate, and is the treatment of choice in all cases of severe poisoning with these agents. *Haemodialfiltration* is more widely and readily available and increases elimination of poisons such as ethylene glycol and methanol, though it is

FURTHER READING

Proudfoot AT, Krenzelok EP, Brent J, et al. Does urine alkalinization increase salicylate elimination? If so, why? *Toxicological Reviews* 2003; **22**: 129–136.

Proudfoot AT, Krenzelok EP, Vale JA. Position paper on urine alkalinization. *Clinical Toxicology* 2004; **42**: 1–26.

less efficient than *haemodialysis*. Charcoal haemoperfusion can significantly reduce the body burden of phenobarbital, carbamazepine and theophylline, but MDAC is as effective and simpler to use.

Toxicological investigations

On admission, or at an appropriate time post overdose, a timed blood sample should be taken if it is suspected that aspirin, digoxin, ethylene glycol, iron, lithium, methanol, paracetamol, paraquat, quinine or theophylline has been ingested. The determination of the concentrations of these drugs will be valuable in management. Drug screens on blood and urine are occasionally indicated in severely poisoned patients in whom the cause of coma is unknown. A poison information service will advise.

Non-toxicological investigations (Table 16.17)

Some routine investigations are of value in the differential diagnosis of coma or the detection of poison-induced hypokalaemia, hyperkalaemia, hypoglycaemia, hyperglycaemia, hepatic or renal failure or acid–base disturbances (Table 16.18). Measurement of carboxyhaemoglobin, methaemoglobin and cholinesterase activities are of assistance in the diagnosis and management of cases of poisoning due to carbon monoxide, methaemoglobin-inducing agents such as nitrites, and organophosphorus insecticides respectively.

Table 16.17 Relevant non-toxicological investigations

Serum sodium (e.g. hyponatraemia in MDMA poisoning) and potassium (e.g. hypokalaemia in theophylline poisoning and hyperkalaemia in digoxin poisoning) concentrations
Plasma creatinine concentration (e.g. renal failure in ethylene glycol poisoning)
Acid-base disturbances, including metabolic acidosis (Table 16.18)
Blood sugar concentration (e.g. hypoglycaemia in insulin poisoning or hyperglycaemia in salicylate poisoning)
Serum calcium concentration (e.g. hypocalcaemia in ethylene glycol poisoning)
Liver function (e.g. in paracetamol poisoning)
Carboxyhaemoglobin concentration (in carbon monoxide poisoning)
Methaemoglobinaemia (e.g. in nitrite poisoning)
Cholinesterase activities (e.g. organophosphorus insecticide and nerve agent poisoning)
ECG (e.g. wide QRS in tricyclic antidepressant poisoning) – see page 949
X-ray (including identification of complications)

Table 16.18 Some poisons inducing metabolic acidosis

Carbon monoxide
Cocaine
Cyanide
Diethyleneglycol
Ethanol
Ethylene glycol
Iron
Methanol
Paracetamol
Tricyclic antidepressants
Calcium channel blockers
Metformin
Topiramate

ECG

Routine ECG is of limited diagnostic value, but continuous ECG monitoring should be undertaken in those ingesting potentially cardiotoxic drugs; for example, sinus tachycardia with prolongation of the PR and QRS intervals in an unconscious patient suggests tricyclic antidepressant overdose. Q–T interval prolongation is an adverse effect of several drugs (e.g. quetiapine and quinine).

Radiology

Routine radiology is of little diagnostic value. It can confirm ingestion of metallic objects (e.g. coins, button batteries) or injection of globules of metallic mercury. Rarely, hydrocarbon solvents (e.g. carbon tetrachloride) may be seen as a slightly opaque layer floating on the top of the gastric contents with the patient upright, or outlining the small bowel. Some enteric-coated or sustained-release drug formulations may be seen on plain abdominal radiographs, but, with the exception of iron salts, ordinary formulations are seldom seen. Ingested packets of illicit substances can sometimes been seen on CT (see p. 936). Radiology can confirm complications of poisoning, for example, aspiration pneumonia, non-cardiogenic pulmonary oedema (salicylates), bronchiolitis obliterans (nitrogen oxides), ARDS, fibrosis (paraquat).

SPECIFIC POISONS: DRUGS AND OTHER CHEMICALS

In this section only specific treatment regimens will be discussed. The general principles of management of self-poisoning will always be required.

Amfetamines including ecstasy (MDMA)

The medicinal product is usually the dextro-isomer, dexamfetamine. The *N*-methylated derivative, metamfetamine, is now abused widely, and the crystalline form of this salt is known as ‘crystal meth’ or ‘ice’. Since the 1980s, the so-called designer amfetamine, 3,4-methylenedioxymethamfetamine (MDMA), commonly known as ecstasy, has been abused world-wide.

Amfetamines are CNS and cardiovascular stimulants. These effects are mediated by increasing synaptic concentrations of epinephrine and dopamine. Poisoning is usually the result of their use for pleasurable purposes.

Clinical features

Amfetamines cause euphoria, extrovert behaviour, a lack of desire to eat or sleep, tremor, dilated pupils, tachycardia and hypertension. More severe intoxication is associated with agitation, paranoid delusions, hallucinations and violent behaviour. Convulsions, rhabdomyolysis, hyperthermia and cardiac arrhythmias may develop in severe poisoning. Rarely, intracerebral and subarachnoid haemorrhage occur and can be fatal.

Mild MDMA poisoning is characterized by agitation, tachycardia, hypertension, widely dilated pupils, trismus and

sweating. In more severe cases, hyperthermia, disseminated intravascular coagulation, rhabdomyolysis, acute kidney injury and hyponatraemia (secondary to inappropriate antidiuretic hormone secretion), predominate.

Treatment

Agitation is controlled by diazepam 10–20 mg i.v. or chlorpromazine 50–100 mg i.m. The peripheral sympathomimetic actions of amfetamines are antagonized by β-adrenoceptor blocking drugs. If hyperthermia is present dantrolene 1 mg/kg body weight i.v. is used.

FURTHER READING

Freedman RR, Johanson CE, Tancer ME. Thermoregulatory effects of 3,4-methylenedioxymethamphetamine (MDMA) in humans. *Psychopharmacology* 2005; **183**: 248–256.

Anticonvulsants

Clinical features

Carbamazepine. In overdose, carbamazepine causes dry mouth, coma, convulsions, nystagmus, ataxia and incoordination. The pupils are often dilated, divergent strabismus may be present and complete external ophthalmoplegia has been reported. Hallucinations occur, particularly in the recovery phase.

Phenytoin. Acute overdose of phenytoin results in nausea, vomiting, headache, tremor, cerebellar ataxia, nystagmus, and, rarely, loss of consciousness.

Sodium valproate. Sodium valproate most frequently causes drowsiness, impairment of consciousness and respiratory depression. In very severe poisoning myoclonic jerks and seizures occur and cerebral oedema has been reported. Liver damage, hyperammonaemia and metabolic acidosis are uncommon complications.

FURTHER READING

Isbister GK, Balit CR, Whyte IM et al. Valproate overdose: a comparative cohort study of self poisonings. *British Journal of Clinical Pharmacology* 2003; **55**: 398–404.

Gabapentin and pregabalin. Lethargy, ataxia, slurred speech and gastrointestinal symptoms may develop.

Lamotrigine. Lethargy, coma, ataxia, nystagmus, seizures and cardiac conduction abnormalities have been reported.

Levetiracetam. Lethargy, coma and respiratory depression have been observed.

Tiagabine. Lethargy, facial grimacing, nystagmus, posturing, agitation, coma, hallucinations and seizures have been reported.

Topiramate. Lethargy, ataxia, nystagmus, myoclonus, coma, seizures and a non-anion gap metabolic acidosis have been observed. Metabolic acidosis can appear within hours of ingestion and persist for days.

Management

A single dose of charcoal 50–100 g administered within 1 hour of ingestion will reduce absorption substantially. Multiple-dose activated charcoal has been shown to significantly increase elimination of carbamazepine. Although haemodialysis is effective in removing sodium valproate the indications for its use remain uncertain.

FURTHER READING

von Mach MA, Gauer M, Meyer S et al. Antidiabetic medications in overdose: a comparison of the inquiries made to a regional poisons unit regarding original sulfonylureas, biguanides and insulin. *International Journal of Clinical Pharmacology and Therapeutics* 2006; **44**: 51–56.

Antidiabetic agents

In all cases of poisoning with insulin and sulfonylureas, prompt diagnosis and treatment are essential if death or cerebral damage from neuroglycopenia is to be prevented. Metformin overdose rarely causes hypoglycaemia, since its mode of action is to increase glucose utilization, but lactic acidosis is a potentially serious complication. Thiazolidinediones cause mild hypoglycaemia but the risk of hepatic injury (recognized in regular treatment) is low after acute single ingestion. Moderate to severe poisoning occurs in overdose with insulin (≈ 15%), biguanides (≈ 12%) and also sulfonylureas (≈ 5%). Overdose with the newer antidiabetic drugs (see p. 1039) has not been a problem so far.

Clinical features

Features include drowsiness, coma, twitching, convulsions, depressed limb reflexes, extensor plantar responses, hyperapnoea, pulmonary oedema, tachycardia and circulatory failure. Hypoglycaemia is to be expected, and hypokalaemia, cerebral oedema and metabolic acidosis may occur. Neurogenic diabetes insipidus and persistent vegetative states are possible long-term complications. Cholestatic jaundice has been described as a late complication of chlorpropamide poisoning.

Treatment

The blood or plasma glucose concentration should be measured urgently and intravenous glucose given, if necessary. Glucagon may be ineffective, although can reduce the amount of glucose required (see p. 1043). Recurring hypoglycaemia is highly likely. A continuous infusion of glucose (with K^+ 10–20 mmol/L) together with carbohydrate-rich meals are required in cases of severe insulin poisoning, though there may be difficulty in maintaining normoglycaemia.

Even if the blood sugar is normal, gastric lavage should only be used if the patient has presented within 1 hour of a very large ingestion of a sulphonyurea. In the case of sulfonylurea poisoning, the administration of glucose only serves to increase already high circulating insulin concentrations. Instead, octreotide 50 μg in adults (1 μg/kg in children) i.v. should be given as it has been shown to be effective in patients developing sulfonylurea-induced hypoglycaemia.

Antimalarials

Clinical features

Chloroquine. Hypotension is often the first clinical manifestation of chloroquine poisoning. It may progress to acute heart failure, pulmonary oedema and cardiac arrest. Agitation and acute psychosis, convulsions and coma may ensue. Hypokalaemia is common and is due to chloroquine-induced potassium channel blockade. Bradyarrhythmias and tachyarrhythmias are common and ECG conduction abnormalities are similar to those seen in quinine poisoning.

Quinine. Cinchonism (tinnitus, deafness, vertigo, nausea, headache and diarrhoea) is common. In more severe poisoning, convulsions, hypotension, pulmonary oedema and cardiorespiratory arrest is seen (due to ventricular arrhythmias which are often preceded by ECG conduction abnormalities, particularly QT prolongation). Quinine cardiotoxicity is due to sodium channel blockade**.** Patients may also develop ocular features, including blindness, which can be permanent.

Primaquine. The main concern regarding primaquine is its propensity to cause methaemoglobinaemia and haemolytic anaemia following overdose.

Management

Gastric lavage or activated charcoal 50–100 g is sometimes used if the patient presents within 1 hour of taking a large dose of drugs. Multiple-dose oral activated charcoal increases quinine and probably chloroquine clearance. Hypokalaemia should be corrected. Sodium bicarbonate 50–100 mmol i.v. is given if the ECG shows intraventricular block but will exacerbate hypokalaemia which should be corrected first. Mechanical ventilation, the administration of an inotrope and high doses of diazepam (1 mg/kg as a loading dose and 0.25–0.4 mg/kg/h maintenance) may reduce the mortality in severe chloroquine poisoning. Overdrive pacing may be required if torsade de pointes (p. 728) occurs in quinine poisoning and does not respond to magnesium sulphate infusion. If clinically significant methaemoglobinaemia (generally above 30%) develops in primaquine poisoning, methylthioninium (methylene blue) 1–2 mg/kg body weight should be administered.

β-Adrenoceptor blocking drugs

Clinical features

In mild poisoning sinus bradycardia is the only feature, but if a substantial amount has been ingested, coma, convulsions and hypotension develop. Less commonly delirium, hallucinations and cardiac arrest supervene.

Treatment

Glucagon 50–150 μg/kg (typically 5–10 mg in an adult) followed by an infusion of 1–5 mg/h is the most effective agent. It acts by bypassing the blocked beta-receptor thus activating adenyl cyclase and promoting formation of cyclic AMP from ATP; cyclic AMP in turn exerts a direct beta-stimulant effect on the heart. Atropine 0.6–1.2 mg i.v. can be used but is usually less effective.

Batteries

Children occasionally ingest disc button batteries, less commonly larger batteries. In the past there was concern that disc batteries might leak and release mercuric oxide but this metal is not used now in Europe. Most disc batteries will pass through the gut in 2 or 3 days. If they lodge in the oesophagus, removal is by endoscopy.

Benzodiazepines

Benzodiazepines are commonly taken in overdose but rarely produce severe poisoning except in the elderly or those with chronic respiratory disease.

Clinical features

Benzodiazepines produce drowsiness, ataxia, dysarthria and nystagmus. Coma and respiratory depression develop in severe intoxication.

Treatment

If respiratory depression is present in patients with severe benzodiazepine poisoning, flumazenil 0.5–1.0 mg i.v. is given in an adult and this dose often needs repeating. Flumazenil is used to avoid ventilation. It is contraindicated in mixed tricyclic antidepressant (TCA)/benzodiazepine poisoning and in those with a history of epilepsy because it may cause convulsions.

Calcium channel blockers

Calcium channel blockers (phenylalkylamines, e.g. verapamil; benzothiazepines, e.g. diltiazem; dihydropyridines, see p. 806) all act by blocking voltage-gated calcium channels. Dihydropyridines (e.g. amlodipine, felodipine, nifedipine) are predominantly peripheral vasodilators while verapamil and, to a lesser extent, diltiazem also have significant cardiac effects. Overdose of these drugs, particularly with verapamil and diltiazem, are serious with heart block and hypotension causing a significant fatality rate.

Clinical features

Hypotension occurs due to peripheral vasodilatation, myocardial depression and conduction block. The electrocardiogram may progress from sinus bradycardia through first, then higher, degrees of block, to asystole. Cardiac and non-cardiac pulmonary oedema may ensue in severely poisoned patients. Other features include nausea, vomiting, seizures and a lactic acidosis. When a sustained-release preparation has been ingested the onset of severe features are delayed, sometimes for more than 12 hours. Overdose with even small amounts can have profound effects.

Treatment

Activated charcoal 50–100 g should be given in an adult to reduce absorption. Treat hypotension initially with intravenous crystalloid. If significant hypotension persists despite volume replacement, administer glucagon (see p. 1043) as it activates myosin kinase independent of calcium. Give i.v. glucagon 10 mg (150 μg/kg) as a slow bolus and repeat every 3–5 minutes. If there is a favourable response in blood pressure, an infusion 5–10 mg/h is commenced; if there is no response after repeated initial boluses, discontinue. Monitor blood glucose.

If hypotension persists despite the above measures, high dose *dopamine* (10–30 μg/kg/min) is a reasonable choice if hypotension is thought to be mainly due to reduced systemic vascular resistance. Alternatively, *adrenaline* (epinephrine) 0.1–1.0 μg/kg/min, which has both alpha and beta adrenergic effects, should be administered. *Insulin–dextrose euglycaemia* has been shown to improve myocardial contractility and systemic perfusion and may be used in addition to sympathomimetic amine therapy. Insulin is given as a bolus dose of 1 U/kg, followed by an infusion of 0.5 U/kg/h with 10% glucose and frequent monitoring of blood glucose and potassium.

Acidosis impairs L-type channel function (see p. 684) and is corrected by the administration of sodium bicarbonate, which has been shown experimentally to improve myocardial contractility and cardiac output.

Intravenous atropine 0.6–1.2 mg in an adult, repeated as required, should be given for bradycardia and heart block. The response to atropine may be improved following parenteral 10% calcium chloride, 5–10 mL (at 1–2 mL/min). The initial dose can be repeated every 3–5 minutes but if there is no response in pulse rate or blood pressure after three such doses it is unlikely that further boluses will be helpful. If there is an initial response to calcium, a continuous infusion may be warranted; this may be given as 10% calcium chloride, 1–10 mL/h. Cardiac pacing may have a role if there is evidence of AV conduction delay but there may be failure to capture.

FURTHER READING

Clemessy JL, Taboulet P, Hoffman JR. Treatment of acute chloroquine poisoning: a 5-year experience. *Critical Care Medicine* 1996; **24,** 1189–1195.

FURTHER READING

DeWitt CR, Waksman JC. Pharmacology, pathophysiology and management of calcium channel blocker and beta-blocker toxicity. *Toxicological Reviews* 2004; **23**: 223–238.

More online
www.studentconsult.com

FURTHER READING

Buckley N, Dawson A, Whyte I. Calcium channel blockers. *Medicine* 2007; **35**: 599–602.

Cannabis (marijuana)

FURTHER READING

Buckley NA, Isbister GK, Stokes B et al. Hyperbaric oxygen for carbon monoxide poisoning: a systematic review and critical analysis of the evidence. *Toxicological Reviews* 2005; **24**: 75–92.

Cannabis is usually smoked but may be ingested as a 'cake', made into a tea or injected intravenously. Apart from alcohol, it is the drug most widely abused in developed countries. The major psychoactive constituent is delta-11-tetrahydrocannabinol (THC). THC possesses activity at the benzodiazepine, opioid and cannabinoid receptors. Street names include pot, grass, ganga, reefs and spliff. It is prepared as marijuana (ganga) from the female flowers; hashish or charas – a concentrated resin of glandular trichomes; kief – chopped female plants; and bhang – a drink prepared from cannabis leaves boiled in milk with spices.

Clinical features

Initially there is euphoria, followed by distorted and heightened images, colours and sounds, altered tactile sensations and sinus tachycardia. Visual and auditory hallucinations and acute psychosis are particularly likely to occur after substantial ingestion in naïve cannabis users. Intravenous injection leads to watery diarrhoea, tachycardia, hypotension and arthralgia.

Heavy users suffer impairment of memory and attention and poor academic performance. There is an increased risk of anxiety and depression. Regular users are at risk of dependence. Cannabis use results in an overall increase in the relative risk for later schizophrenia and psychotic episodes (see p. 1216). Cannabis smoke is probably carcinogenic.

Treatment

Reassurance is usually the only treatment required, though sedation with intravenous diazepam 10–20 mg i.v. in an adult or chlorpromazine 50–100 mg i.m. in an adult is sometimes required. Hypotension requires i.v. fluids.

Carbamate insecticides

Carbamate insecticides inhibit acetylcholinesterase but the duration of this effect is comparatively short-lived since the carbamate–enzyme complex tends to dissociate spontaneously. The clinical features are similar to those of organophosphorus insecticide poisoning (see p. 947). Atropine 0.6–2 mg i.v. is required for bradycardia and recovery invariably occurs within 24 hours.

Carbon monoxide

FURTHER READING

Hollander JE, Henry TD. Evaluation and management of the patient who has cocaine-associated chest pain. *Cardiology Clinics* 2006; **24**: 103–114.

The commonest source of carbon monoxide is an improperly maintained and poor, ventilated heating system. In addition, inhalation of methylene chloride (found in paint strippers) may also lead to carbon monoxide poisoning as methylene chloride is metabolized in vivo to carbon monoxide. The affinity of haemoglobin for carbon monoxide is some 240 times greater than that for oxygen. Carbon monoxide combines with haemoglobin to form carboxyhaemoglobin, thereby reducing the total oxygen carrying capacity of the blood and increasing the affinity of the remaining haem groups for oxygen. This results in tissue hypoxia. In addition, carbon monoxide also inhibits cytochrome oxidase a_3.

Clinical features

Symptoms of mild to moderate exposure to carbon monoxide may be mistaken for a viral illness. A peak carboxyhaemoglobin (COHb) concentration of less than 10% is not normally associated with symptoms and peak COHb concentrations of 10–30% result only in headache and mild exertional dyspnoea. Higher concentrations of COHb are associated with coma, convulsions and cardiorespiratory arrest. Neuropsychiatric features occur after apparent recovery from carbon monoxide intoxication.

Treatment

In addition to removing the patient from carbon monoxide exposure, high flow oxygen should be administered using a tightly fitting face mask. Endotracheal intubation and mechanical ventilation is required in those who are unconscious. Several controlled studies of hyperbaric oxygen have now been published but none have shown long-term clinical benefit.

Cocaine

Cocaine hydrochloride ('street' cocaine, 'coke') is a water-soluble powder or granule that can be taken orally, intravenously or intranasally. 'Freebase' or 'crack' cocaine comprises crystals of relatively pure cocaine without the hydrochloride moiety and is obtained in rocks (150 mg of cocaine). It is more suitable for smoking in a pipe or mixed with tobacco and can also be heated on foil and the vapour inhaled (approximately 35 mg of drug per 'line' or a 'rail'. The 'effects' of cocaine are experienced almost immediately with i.v. or smoking routes, about 10 minutes in the intranasal route and 45–90 minutes when taken orally. The effects start resolving in about 20 minutes and may last up to 90 minutes. In severe poisoning, death may occur in minutes but survival beyond 3 hours is not usually fatal. Cocaine blocks the reuptake of biogenic amines. Inhibition of dopamine reuptake is responsible for the psychomotor agitation which commonly accompanies cocaine use. Blockade of norepinephrine reuptake produces tachycardia, and inhibition of serotonin reuptake may induce hallucinations. Cocaine also enhances CNS arousal by potentiating the effects of excitatory amino acids. Cocaine is also a powerful local anaesthetic and vasoconstrictor.

Clinical features

After initial euphoria, cocaine produces agitation, tachycardia, hypertension, sweating, hallucinations, convulsions, metabolic acidosis, hyperthermia, rhabdomyolysis and ventricular arrhythmias. Dissection of the aorta, myocarditis, myocardial infarction, dilated cardiomyopathy, subarachnoid haemorrhage, and cerebral haemorrhage and infarction also occur. If a young person presents with a stroke or myocardial infarction, cocaine overdosage is a possibility because of its vasoconstrictor effect.

Treatment

Diazepam 10 mg i.v. is used to control agitation and convulsions. Active external cooling should be used for hyperthermia. β-adrenoceptor blockers are contraindicated for the treatment of hypertension as they may cause paradoxical hypertension; phentolamine 2–5 mg i.v. can be used. Early use of a benzodiazepine, aspirin and nitrates is effective in relieving cocaine-associated chest pain. β-adrenoceptor blocking drugs are contraindicated because they may exacerbate cocaine-induced coronary artery vasoconstriction.

Cyanide

Cyanide and its derivatives are used widely in industry. Hydrogen cyanide is also released during the thermal decomposition of polyurethane foams. Cyanide reversibly inhibits cytochrome oxidase a_3 so that cellular respiration ceases.

Clinical features

Inhalation of hydrogen cyanide gas produces symptoms within seconds and death within minutes. By contrast, the ingestion of a cyanide salt may not produce features for 1 hour. After exposure initial symptoms are non-specific and include a feeling of constriction in the chest and dyspnoea. Coma, convulsions and metabolic acidosis may then supervene.

Treatment

Dicobalt edetate (and the free cobalt contained in the preparation) complexes free cyanide. Sodium thiosulphate and hydroxocobalamin enhance endogenous cyanide detoxification mechanisms. Sodium nitrite produces methaemoglobinaemia; methaemoglobin combines with cyanide to form cyanmethaemoglobin.

Oxygen should be administered and, if available, dicobalt edetate 300 mg should be administered i.v.; the dose is repeated in severe cases. An alternative but very expensive antidote is hydroxocobalamin 5 g i.v.; a second dose may be required in severe cases. If these two antidotes are not available, sodium nitrite 300 mg i.v. and sodium thiosulphate 12.5 g i.v. should be administered.

Digoxin

Toxicity occurring during chronic administration is common, though acute poisoning is infrequent.

Clinical features

These include nausea, vomiting, dizziness, anorexia and drowsiness. Rarely, confusion, visual disturbances and hallucinations occur. Sinus bradycardia is often marked and may be followed by supraventricular arrhythmias with or without heart block, ventricular premature beats and ventricular tachycardia. Hyperkalaemia occurs due to the inhibition of the sodium-potassium activated ATPase pump.

Treatment

Sinus bradycardia, atrioventricular block and sinoatrial standstill are often reduced or even abolished by atropine 1.2–2.4 mg i.v. If cardiac output is compromised, however, digoxin-specific antibody fragments (digoxin-Fab) should be administered. In both acute and chronic poisoning, only half the estimated dose (calculated from amount of drug taken or serum digoxin concentration) required for full neutralization need be given initially; a further dose may be given if clinically indicated.

Ethanol

Ethanol is commonly ingested in beverages and deliberately with other substances in overdose. It is also present in many cosmetic and antiseptic preparations. Following absorption, ethanol is oxidized to acetaldehyde and then to acetate. Ethanol is a CNS depressant and the features of ethanol intoxication are generally related to blood concentrations (Table 16.19).

Table 16.19 Clinical features of ethanol poisoning

Blood [ethanol] 500–1500 mg/L (11.0–32.5 mmol/L)
Emotional lability
Mild impairment of coordination

Blood [ethanol] 1500–3000 mg/L (32.5–65.0 mmol/L)
Visual impairment
Incoordination
Slowed reaction time
Slurred speech

Blood [ethanol] 3000–5000 mg/L (65.0–108.5 mmol/L)
Marked incoordination
Blurred or double vision
Stupor
Occasionally hypoglycaemia, hypothermia and convulsions

Blood [ethanol] >5000 mg/L (>108.5 mmol/L)
Depressed reflexes
Respiratory depression
Hypotension
Hypothermia
Death (from respiratory or circulatory failure or aspiration)

Clinical features

In children in particular severe hypoglycaemia may accompany alcohol intoxication due to inhibition of gluconeogenesis. Hypoglycaemia is also observed in those who are malnourished or who have fasted in the previous 24 hours. In severe cases of intoxication, coma and hypothermia are often present and lactic acidosis, ketoacidosis and acute renal failure have been reported.

Treatment

As ethanol-induced hypoglycaemia is not responsive to glucagon, i.v. glucose 25 g (50 mL of 50% dextrose) should be given. Haemodialysis is used if the blood ethanol concentration exceeds 7500 mg/L and if a severe metabolic acidosis is present.

Ethylene glycol

Ethylene glycol is a common constituent of antifreeze fluid used in car radiators. Ethylene glycol itself is non-toxic but is metabolized to toxic products (Fig. 16.3).

Clinical features

Initially the features of ethylene glycol poisoning are similar to ethanol intoxication (though there is no ethanol on the breath). Coma and convulsions follow and a variety of neurological abnormalities including nystagmus and ophthalmoplegias can be seen. Severe metabolic acidosis, hypocalcaemia and the presence of calcium oxalate crystalluria are well recognized complications.

Treatment

If the patient presents early after ingestion, the first priority is to inhibit metabolism using either intravenous ethanol or fomepizole. Secondly, following a substantial ingestion, haemodialysis or haemodialfiltration should be employed to

FURTHER READING

Bateman DN. Digoxin-specific antibody fragments: how much and when? *Toxicological Reviews* 2004; **23**: 135–143.

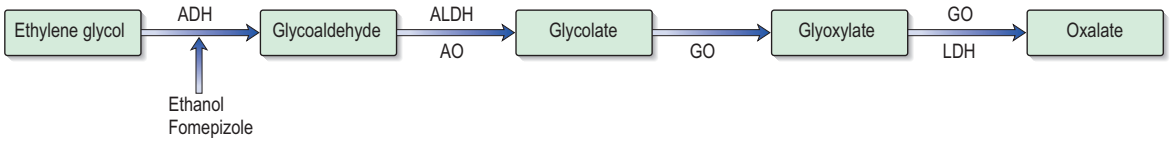


Fig. 16.3 **The metabolism of ethylene glycol.** ADH, alcohol dehydrogenase; ALDH, aldehyde dehydrogenase; AO, aldehyde oxidase; GO, glycolate oxidase; LDH, lactate dehydrogenase.

remove ethylene glycol, its aldehyde metabolites and glycolate. A loading dose of ethanol 50 g should be administered followed by an i.v. infusion of ethanol 10–12 g/h to produce blood ethanol concentrations of 500–1000 mg/L (11–22 mmol/L). The infusion is continued until ethylene glycol is no longer detectable in the blood. If haemodialysis is employed, the rate of ethanol administration will need to be increased to 17–22 g/h as ethanol is dialyzable. Alternatively, fomepizole 15 mg/kg body weight can be administered followed by four 12-hourly doses of 10 mg/kg, then 15 mg/kg every 12 hours until ethylene glycol concentrations are less than 200 mg/L. If dialysis is employed the frequency of fomepizole dosing should be increased to 4-hourly because fomepizole is dialyzable. Supportive measure to combat shock, hypocalcaemia and metabolic acidosis should be instituted.

FURTHER READING

Brent J, McMartin K, Phillips S et al. Fomepizole for the treatment of ethylene glycol poisoning. *New England Journal of Medicine* 1999; **340**: 832–838.

Gamma-hydroxybutyric acid (GHB)

Gamma-hydroxybutyric acid occurs naturally in mammalian brain where it is derived metabolically from gamma-aminobutyric acid (GABA). GHB has emerged as a major recreational drug for body building, weight loss and for producing a 'high'. Street names include cherry meth and liquid X. It is taken as a colourless liquid dissolved in water.

Clinical features

Poisoning with GHB is characterized by aggressive behaviour, ataxia, amnesia, vomiting, drowsiness, bradycardia, respiratory depression and apnoea, seizures and coma, which is often short lived.

Management

In a patient who is breathing spontaneously, the management of GHB poisoning is primarily supportive with oxygen supplementation and the administration of atropine for persistent bradycardia, as necessary. Those who are severely poisoned will require mechanical ventilation, though recovery is usually complete within 6–8 hours.

FURTHER READING

Tenenbein M. Benefits of parenteral deferoxamine for acute iron poisoning. *Clinical Toxicology* 1996; **34**: 485–489.

Household products

The agents most commonly involved are bleach, cosmetics, toiletries, detergents, disinfectants and petroleum distillates such as paraffin and white spirit. Ingestion of household products is usually accidental and is most common among children less than 5 years of age. If the ingestion is accidental, features very rarely occur except in the case of petroleum distillates where aspiration is a recognized complication because of their low surface tension. Powder detergents, sterilizing tablets, denture cleaning tablets and industrial bleaches (which contain high concentrations of sodium hypochlorite) are corrosive to the mouth and pharynx if ingested. Nail polish and nail polish remover contain acetone that may produce coma if ingested in substantial quantities. Inhalation by small children of substantial quantities of talcum powder has occasionally given rise to severe pulmonary oedema and death.

Iron

Unless more than 60 mg of elemental iron per kg of body weight is ingested (a ferrous sulphate tablet contains 60 mg of iron), features are unlikely to develop. As a result poisoning is seldom severe but deaths still occur. Iron salts have a direct corrosive effect on the upper gastrointestinal tract.

Clinical features

The initial features are characterized by nausea, vomiting (the vomit may be grey or black in colour), abdominal pain and diarrhoea. Severely poisoned patients develop haematemesis, hypotension, coma and shock at an early stage. Usually, however, most patients only suffer mild gastrointestinal symptoms. A small minority deteriorate 12–48 hours after ingestion and develop shock, metabolic acidosis, acute renal tubular necrosis and hepatocellular necrosis. Rarely, up to 6 weeks after ingestion, strictures due to corrosive damage occur. The serum iron concentration should be measured some 4 hours after ingestion and if the concentration exceeds the predicted normal iron binding capacity (usually more than 5 mg/L [90 μmol/L]), free iron is circulating and treatment with desferrioxamine may be required.

Treatment

The majority of patients ingesting iron do not require desferrioxamine therapy. If a patient develops coma or shock, desferrioxamine should be given without delay in a dose of 15 mg/kg/h i.v. (total amount of infusion not to exceed 80 mg/kg in 24 hours). If the recommended rate of administration is exceeded, or the therapy is continued for several days, adverse effects including pulmonary oedema and ARDS have been reported.

Lead

Exposure to lead occurs occupationally, children may eat lead-painted items in their homes (pica) and the use of lead-containing cosmetics or 'drugs' has also resulted in lead poisoning.

Clinical features

Mild intoxication may result in no more than lethargy and occasional abdominal discomfort, though abdominal pain, vomiting, constipation and encephalopathy (seizures, delirium, coma) may develop in more severe cases. Encephalopathy is more common in children than in adults but is now rare in the developed world. Typically, though very rarely,

lead poisoning results in foot drop attributable to peripheral motor neuropathy. A bluish discolouration of the gum margins due to the deposition of lead sulphide is observed occasionally.

The characteristic haematological features include:

- *sideroblastic anaemia*, due to inhibition by lead of several enzymes involved in haem synthesis, including ALA synthetase
- *haemolysis*, which is usually mild, resulting from damage to the red cell membrane
- *punctate basophilia* (or basophilic stippling: the blood film shows red cells with small, round, blue particles), due to aggregates of RNA in immature red cells owing to inhibition by lead of pyrimidine-5-nucleotidase, which normally disperses residual RNA to produce a diffuse blue staining seen in reticulocytes on blood films (polychromasia).

Treatment

The social and occupational dimensions of lead poisoning must be recognized. Simply giving children chelation therapy and then returning them to a contaminated home environment is of no value. Similarly, returning a worker to an environment where he was exposed previously and excessively to lead is inappropriate.

The decision to use chelation therapy is based not only on the blood lead concentration but on the presence of symptoms. Parenteral sodium calcium edetate 75 mg/kg/day has been the chelating agent of choice for over 50 years. However, there is accumulating evidence to suggest that oral DMSA 30 mg/kg/day is of similar efficacy.

Lithium

Lithium toxicity is usually the result of therapeutic overdosage (chronic toxicity) rather than deliberate self-poisoning (acute toxicity). However, single large doses are occasionally ingested by individuals on long-term treatment with the drug (acute on therapeutic toxicity).

Clinical features

Features of intoxication include thirst, polyuria, diarrhoea and vomiting and in more serious cases impairment of consciousness, hypertonia and convulsions; irreversible neurological damage may occur. Measurement of the serum lithium concentration confirms the diagnosis. Therapeutic toxicity is usually associated with concentrations above 1.5 mmol/L (10.4 mg/L). Acute massive overdose may produce concentrations of 5 mmol/L (34.7 mg/L) without causing toxic features.

Treatment

Forced diuresis with sodium chloride 0.9% is effective in increasing elimination of lithium, though haemodialysis is far superior and should be undertaken particularly if neurological features are present, if renal function is impaired and if chronic toxicity or acute on chronic toxicity are the modes of presentation.

Mercury

Mercury is the only metal that is liquid at room temperature. It exists in three oxidation states, (elemental/metallic Hg^0, mercurous Hg_2^{2+} and mercuric Hg^{2+}) and can form inorganic (e.g. mercuric chloride) and organic (e.g. methylmercury) compounds. Metallic mercury is very volatile and when spilled has a large surface area so that high atmospheric concentrations may be produced in enclosed spaces, particularly when environmental temperatures are high. Thus, great care should be taken in clearing up a spillage of mercury if a thermometer or sphygmomanometer is broken. If ingested, metallic mercury will usually be eliminated per rectum, though small amounts may be found in the appendix. Mercury salts are well absorbed following ingestion as are organometallic compounds where mercury is covalently bound to carbon.

Clinical features

Inhalation of acute mercury vapour causes headache, nausea, cough, chest pain, bronchitis and occasionally pneumonia. Proteinuria and nephrotic syndrome are observed rarely. In addition a fine tremor and neurobehavioural impairment occurs and peripheral nerve involvement has also been observed. Ingestion of inorganic and organic mercury compounds causes an irritant gastroenteritis with corrosive ulceration, bloody diarrhoea and abdominal cramps and may lead to circulatory collapse and shock. Mercurous compounds are less corrosive and toxic than mercuric salts.

Treatment

Dimercaptopropanesulphonate (DMPS) is the antidote of choice and is given orally in a dose of 30 mg/kg per day. At least 5 days' treatment is usually required.

Methanol

Methanol is used widely as a solvent and is found in antifreeze solutions. Methanol is metabolized to formaldehyde and formate (Fig. 16.4). The concentration of formate increases greatly and is accompanied by accumulation of hydrogen ions causing metabolic acidosis.

Clinical features

Methanol causes inebriation and drowsiness. After a latent period coma supervenes. Blurred vision and diminished visual acuity occur. The presence of dilated pupils that are unreactive to light suggests that permanent blindness is likely to ensue. A severe metabolic acidosis may develop and be accompanied by hyperglycaemia and a raised serum amylase activity. A blood methanol concentration of 500 mg/L (15.63 mmol/L) confirms serious poisoning. The mortality correlates well with the severity and duration of metabolic acidosis. Survivors may show permanent neuro-

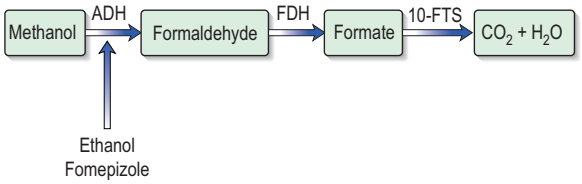


Fig. 16.4 **The metabolism of methanol.** ADH, alcohol dehydrogenase; FDH, formaldehyde dehydrogenase; 10-FTS, 10-formyl tetrahydrofolate synthetase.

FURTHER READING

Bradberry SM, Vale JA. Lead. *Medicine* 2007; **35**: 627–628.

More online

www.studentconsult.com

FURTHER READING

Clarkson TW, Magos L. The toxicology of mercury and its chemical compounds. *Critical Reviews in Toxicology* 2006; **36**: 609–662.

FURTHER READING

Jaeger A. Lithium. *Medicine* 2007; **35**: 629–630.

logical sequelae including parkinsonian-like signs as well as blindness.

Treatment

Treatment is similar to that of ethylene glycol poisoning (see p. 943). Metabolic acidosis should be corrected, methanol metabolism should be inhibited by the administration of ethanol or fomepizole, and haemodialysis is used to remove high circulating methanol concentrations. Folinic acid 30 mg i.v. 6-hourly may protect against ocular toxicity by accelerating formate metabolism.

FURTHER READING

Barceloux DG, Bond GR, Krenzelok EP et al. American Academy of Clinical Toxicology practice guidelines on the treatment of methanol poisoning. *Clinical Toxicology* 2002: **40**: 415–446.

Monoamine oxidase inhibitors

These are now used less frequently in the treatment of depression because of the dangers of dietary and drug interactions. Hence, poisoning with them is correspondingly uncommon. Clinical features after overdose may be delayed for 12–24 hours. They include excitement, restlessness, hyperpyrexia, hyper-reflexia, convulsions, opisthotonos, rhabdomyolysis and coma. Sinus tachycardia and either hypo- or hypertension have also been observed.

Treatment is supportive with control of convulsions and marked excitement; diazepam 10–20 mg i.v. in an adult should be given as necessary and repeated. Dantrolene 1 mg/kg i.v. should be administered if hyperpyrexia develops. Hypotension should be treated with plasma expansion and hypertension by the administration of an α-adrenoceptor blocker such as chlorpromazine.

Ask the authors
www.studentconsult.com

Nerve agents

Nerve agents are related chemically to organophosphorus insecticides and have a similar mechanism of toxicity, but a much higher mammalian acute toxicity, particularly via the dermal route. Two classes of nerve agent are recognized: G agents (named for Gerhardt Schrader who synthesized the first agents) and V agents (V allegedly stands for venomous). G agents include tabun, sarin, soman and cyclosarin. The V agents were introduced later, e.g. VX. The G agents are both dermal and respiratory hazards, whereas the V agents, unless aerosolized, are contact poisons.

Agents used in bioterrorism are described on page 959.

Clinical features

Systemic features include increased salivation and rhinorrhoea, miosis and eye pain, abdominal pain, nausea, vomiting and diarrhoea; involuntary micturition and defecation; muscle weakness and fasciculation, tremor, restlessness, ataxia and convulsions. Bradycardia, tachycardia and hypotension occur, dependent on whether muscarinic or nicotinic effects predominate. Death occurs from respiratory failure within minutes but mild or moderately exposed individuals usually recover completely.

Treatment

The parenteral administration of atropine 2 mg i.v. to patients presenting with rhinorrhoea and bronchorrhoea may be lifesaving. In addition, an oxime (pralidoxime, obidoxime or HI-6) should be given by slow i.v. injection; for example pralidoxime mesilate 30 mg/kg; the dose may be repeated 4- to 6-hourly. Outside hospital, atropine and oxime can be administered most conveniently by the use of an autoinjector. After admission, bolus doses of atropine and an infusion of an oxime (e.g. pralidoxime mesilate 8–10 mg/kg/h) should be given. Intravenous diazepam (adult 10–20 mg; child 1–5 mg) is useful in controlling apprehension, agitation, fasciculation and convulsions; the dose is repeated as required.

FURTHER READING

Burns MJ. The pharmacology and toxicology of atypical antipsychotic agents. *Clinical Toxicology* 2001; **39**: 1–14.

FURTHER READING

Marrs TC, Rice P, Vale JA. The role of oximes in the treatment of nerve agent poisoning in civilian casualties. *Toxicological Reviews* 2006; **25**: 297–323.

Neuroleptics and atypical neuroleptics

Neuroleptic (antipsychotic) drugs are thought to act predominantly by effects on dopamine D_2 receptors. Older neuroleptics include the phenothiazines, the butyrophenones and the substituted benzamides. More recently more selective 'atypical' antipsychotic drugs have been developed, including amisulpiride, aripiprazole, clozapine, olanzapine, quetiapine and risperidone.

Clinical features

These include impaired consciousness, hypotension, respiratory depression, hypothermia or hyperthermia, antimuscarinic effects such as tachycardia, dry mouth and blurred vision, occasionally seizures, rhabdomyolysis, cardiac arrhythmias (both atrial and ventricular) and acute respiratory distress syndrome. Extrapyramidal effects, including acute dystonic reactions, occur but are not dose related. Most 'atypical' antipsychotics have less profound sedative actions than the older neuroleptics. Q–T interval prolongation and subsequent ventricular arrhythmias (including torsades de pointes) have occurred following overdose with the atypical neuroleptics. Unpredictable fluctuations in conscious level, with variations between agitation and marked somnolence, have been particularly associated with olanzapine overdose.

Treatment

Gastric lavage or the administration of oral activated charcoal 50–100 g is used in an adult who presents within 1 hour of a substantial overdose. Benzatropine 2 mg i.v. in an adult is occasionally required for the treatment of dyskinesia and oculogyric crisis. If hypotension is severe and does not respond to intravenous fluids, a sympathomimetic amine such as noradrenaline (norepinephrine) may be required. After correcting acidosis with sodium bicarbonate, the preferred treatment for arrhythmias caused by antipsychotic drugs (usually torsade de pointes) is intravenous magnesium or cardiac pacing. Lidocaine or phenytoin may be effective if multifocal ventricular arrhythmias occur.

Non-steroidal anti-inflammatory agents (NSAIDs)

Self-poisoning with NSAIDs has increased, particularly now that ibuprofen is available without prescription in many countries.

Clinical features

In most cases minor gastrointestinal disturbance is the only feature but, in more severe cases, coma, convulsions and renal failure have occurred. Poisoning with mefenamic acid commonly results in convulsions though these are usually short-lived.

Treatment is symptomatic and supportive.

Opiates and opioids

Cardinal signs of opiate poisoning are pinpoint pupils, reduced respiratory rate and coma. Hypothermia, hypoglycaemia and convulsions are occasionally observed in severe cases. In severe heroin overdose, non-cardiogenic pulmonary oedema has been reported.

Treatment. Naloxone 1.2 mg i.v. in an adult (5–10 µg body weight in a child) will reverse severe respiratory depression and coma at least partially. In severe poisoning larger initial doses or repeat doses will be required. The duration of action of naloxone is often less than the drug taken in overdose, e.g. methadone, which has a very long half-life. For this reason an infusion of naloxone is often required. Non-cardiogenic pulmonary oedema should be treated with mechanical ventilation.

Organophosphorus insecticides

Organophosphorus insecticides are used widely throughout the world and are a common cause of poisoning, causing thousands of deaths, in the developing world. Intoxication may follow ingestion, inhalation or dermal absorption. Organophosphorus insecticides inhibit acetylcholinesterase causing accumulation of acetylcholine at central and peripheral cholinergic nerve endings, including neuromuscular junctions. Many OP insecticides require biotransformation before becoming active and so the features of intoxication may be delayed.

Clinical features

Poisoning is characterized by anxiety, restlessness, tiredness, headache, and muscarinic features such as nausea, vomiting, abdominal colic, diarrhoea, tenesmus, sweating, hypersalivation and chest tightness. Miosis may be present. Nicotinic effects include muscle fasciculation and flaccid paresis of limb muscles, respiratory muscles, and, occasionally of extraocular muscles. Respiratory failure will ensue in severe cases and is exacerbated by the development of bronchorrhoea and pulmonary oedema. Coma and convulsions occur in severe poisoning. Diagnosis is confirmed by measuring the erythrocyte cholinesterase activity; plasma cholinesterase activity is less specific but may also be depressed.

The *intermediate syndrome* usually becomes established 2–4 days after exposure when the symptoms and signs of the acute cholinergic syndrome are no longer obvious. The characteristic features of the syndrome are weakness of the muscles of respiration (diaphragm, intercostal muscles and accessory muscles including neck muscles) and of proximal limb muscles. Accompanying features often include weakness of muscles innervated by some cranial nerves.

Treatment

Mild cases require no specific treatment other then the removal of soiled clothing; contaminated skin should be washed with soap and water to prevent further absorption. Atropine 2 mg i.v. should be given to reduce bronchorrhoea. In addition, an oxime (e.g. pralidoxime mesilate 30 mg/kg by slow i.v. injection 4-hourly or an infusion of pralidoxime mesilate 8–10 mg/kg/h), which reactivates phosphorylated acetylcholinesterase, should be given in symptomatic patients where the diagnosis has been confirmed.

Paracetamol (acetaminophen)

Paracetamol is the most common form of poisoning encountered in the UK. In therapeutic dose, paracetamol is conjugated with glucuronide and sulphate. A small amount of paracetamol is metabolized by mixed function oxidase enzymes to form a highly reactive compound (*N*-acetyl-*p*-benzoquinoneimine, NAPQI), which is then immediately conjugated with glutathione and subsequently excreted as cysteine and mercapturic conjugates. In overdose, large amounts of paracetamol are metabolized by oxidation because of saturation of the sulphate conjugation pathway. Liver glutathione stores become depleted so that the liver is unable to deactivate the toxic metabolite. Paracetamol-induced renal damage probably results from a mechanism similar to that which is responsible for hepatotoxicity.

The severity of paracetamol poisoning is dose related. There is, however, some variation in individual susceptibility to paracetamol-induced hepatotoxicity. Patients with pre-existing liver disease, those with a high alcohol intake and poor nutrition, those receiving enzyme-inducing drugs, those suffering from anorexia nervosa and other eating disorders and HIV infection should be considered to be at greater risk and given treatment at plasma paracetamol concentrations lower than those normally used for interpretation (Fig. 16.5).

Clinical features

Following the ingestion of an overdose of paracetamol, patients usually remain asymptomatic for the first 24 hours or at the most develop anorexia, nausea and vomiting. Liver damage is not usually detectable by routine liver function tests until at least 18 hours after ingestion of the drug. Liver damage usually reaches a peak, as assessed by measurement of alanine transferase (ALT) activity and prothrombin time (INR), at 72–96 hours after ingestion. Without treatment, a small percentage of patients will develop fulminant hepatic

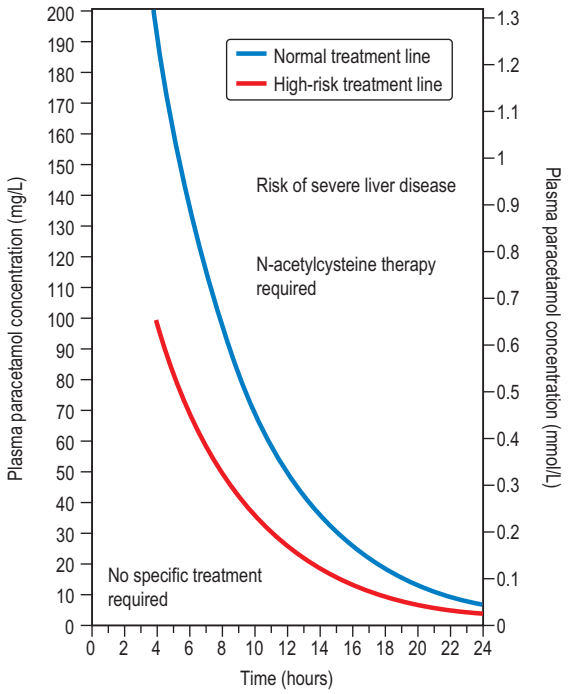


Fig. 16.5 **Nomogram of paracetamol.** For definition of 'high-risk patients', see text.

FURTHER READING

Marrs TC, Vale JA. Management of organophosphorus pesticide poisoning. In: Gupta RC (ed) *Toxicology of Organophosphate and Carbamate Compounds*. Amsterdam: Academic Press, 2006, 715–733.

failure. Renal failure due to acute tubular necrosis occurs in 25% of patients with severe hepatic damage and in a few without evidence of serious disturbance of liver function.

Treatment

The treatment protocol is dependent on the time of presentation and this is summarized in Table 16.20. *N*-acetylcysteine (NAC) has emerged as an effective protective agent provided that it is administered within 8–10 hours of ingestion of the overdose. NAC acts by replenishing cellular glutathione stores, though it may also repair oxidation damage caused by NAPQI. The treatment regimen is shown in Table 16.21.

Up to 10% of patients treated with intravenous NAC (20.25 h regimen) develop rash, angio-oedema, hypotension and bronchospasm. These reactions, which are related to the initial bolus, are seldom serious and discontinuing the infusion is usually all that is required. In more severe cases, chlorphenamine 10–20 mg i.v. in an adult should be given.

If liver or renal failure ensues, this should be treated conventionally though there is evidence that a continuing infusion of NAC (continue 16-h infusion until recovery) will improve the morbidity and mortality. Liver transplantation has been performed successfully in patients with paracetamol-induced fulminant hepatic failure (see p. 342).

FURTHER READING

Vale JA, Proudfoot AT. Paracetamol (acetaminophen) poisoning. *Lancet* 1995; **346**: 547–552.

Phosphides

Aluminium and zinc phosphides are used as rodenticides and insecticides. They react with moisture in the air (and the gastrointestinal tract) to produce phosphine, the active pesticide. Acute poisoning with these compounds may be direct, due to ingestion of the salts, or indirect from accidental inhalation of phosphine generated during their approved use. Both forms of poisoning are mediated by phosphine which inhibits cytochrome c oxidase.

Clinical features

Ingestion causes vomiting, epigastric pain, peripheral circulatory failure, severe metabolic acidosis, renal failure and disseminated intravascular coagulation, in addition to the features induced by phosphine. Exposure to phosphine causes lacrimation, rhinorrhoea, productive cough, breathlessness, chest tightness, dizziness, headache, nausea and drowsiness. Acute pulmonary oedema, hypertension, cardiac arrhythmias, convulsions and jaundice have been described in severe cases. Ataxia, intention tremor and diplopia are found on examination.

Treatment

Treatment is symptomatic and supportive. Gastric lavage should be avoided as it might increase the rate of disintegration of the product ingested and increase toxicity. Activated charcoal may bind metal phosphides. The mortality is very high despite supportive care.

Salicylates

Aspirin is metabolized to salicylic acid (salicylate) by esterases present in many tissues, especially the liver, and subsequently to salicyluric acid and salicyl phenolic glucuronide (Fig. 16.6); these two pathways become saturated with the consequence that the renal excretion of salicylic acid increases after overdose; this excretion pathway is extremely sensitive to changes in urinary pH.

Table 16.20 Management of patients with paracetamol poisoning

≤8 hours after ingestion

- Take blood for urgent estimation of the plasma paracetamol concentration as soon as 4 hours or more have elapsed since ingestion. Check INR, plasma creatinine and alanine transferase (ALT) activity
- Assess whether the patient is at increased risk of liver damage
- Give treatment if needed (Fig. 16.5)
- If the plasma paracetamol concentration is not available within 8 hours of the overdose and, if >150 mg/kg paracetamol has been ingested, treatment should be started at once and stopped if the plasma paracetamol concentration subsequently indicates that treatment is not required
- Check INR, plasma creatinine and ALT activity on the completion of treatment and before discharge

8–15 hours after ingestion

- Urgent action is required because the efficacy of treatment declines progressively from 8 hours after overdose. If >150 mg/kg paracetamol has been ingested, start treatment immediately
- Take blood for urgent estimation of the plasma paracetamol concentration, INR, plasma creatinine and ALT activity
- Assess whether the patient is at increased risk of liver damage
- The need for treatment should be determined by using Fig. 16.5
- In patients already receiving treatment, only discontinue if the plasma paracetamol concentration is below the relevant treatment line (Fig. 16.5) and there is no abnormality of the INR, plasma creatinine or ALT activity and the patient is asymptomatic. Do not discontinue the infusion if there is any doubt as to the timing of the overdose
- At the end of treatment, measure INR, plasma creatinine and ALT activity. If any test is abnormal or the patient is symptomatic, further monitoring is required and expert advice should be sought
- Patients with normal INR, plasma creatinine and ALT activity and who are asymptomatic may be discharged

15–24 hours after ingestion

- Urgent action is required because the efficacy of treatment is limited more than 15 hours after overdose. Start treatment immediately if >150 mg/kg paracetamol has been ingested
- In all patients, take blood for urgent estimation of the plasma paracetamol concentration, INR, plasma creatinine and ALT activity
- Assess whether the patient is at increased risk of liver damage
- Give treatment if needed (Fig. 16.5). The prognostic accuracy of the '200 mg/L line' after 15 hours is uncertain but a plasma paracetamol concentration above the extended treatment line should be regarded as carrying serious risk of severe liver damage
- At the end of treatment, check INR, plasma creatinine and ALT activity. If any test is abnormal or the patient is symptomatic, further monitoring is required and expert advice should be sought

Table 16.21 Regimen for *N*-acetylcysteine

- ***N*-acetylcysteine** 150 mg/kg in 200 mL 5% dextrose over 15 min, then 50 mg/kg in 500 mL of 5% dextrose over the next 4 hours and then 100 mg/kg in 1000 mL of 5% dextrose over the ensuing 16 hours
- Total dose 300 mg/kg over 20.25 hours

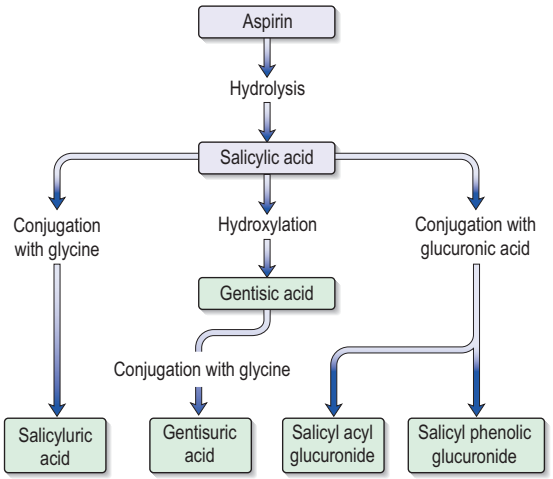


Fig. 16.6 **The metabolism of aspirin.**

Clinical features

Salicylates stimulate the respiratory centre, increase the depth and rate of respiration, and induce a respiratory alkalosis. Compensatory mechanisms, including renal excretion of bicarbonate and potassium, result in a metabolic acidosis. Salicylates also interfere with carbohydrate, fat and protein metabolism, disrupt oxidative phosphorylation, producing increased concentrations of lactate, pyruvate and ketone bodies, all of which contribute to the acidosis.

Thus, tachypnoea, sweating, vomiting, epigastric pain, tinnitus and deafness develop. Respiratory alkalosis and metabolic acidosis supervene and a mixed acid–base disturbance is observed commonly. Rarely, in severe poisoning, non-cardiogenic pulmonary oedema, coma and convulsions ensue.

Treatment

Fluid and electrolyte replacement is required and special attention should be paid to potassium supplementation. Severe metabolic acidosis requires at least partial correction with the administration of sodium bicarbonate intravenously. Mild cases of salicylate poisoning are managed with parenteral fluid and electrolyte replacement only. Patients whose plasma salicylate concentrations are in excess of 500 mg/L (3.62 mmol/L) should receive urine alkalinization (see p. 938). Haemodialysis is the treatment of choice for severely poisoned patients (plasma salicylate concentration >700 mg/L [>5.07 mmol/L]), particularly those with coma and metabolic acidosis.

Selective serotonin reuptake inhibitors (SSRIs)

Citalopram, fluoxetine, fluvoxamine, paroxetine and sertraline are antidepressants that inhibit serotonin reuptake. They lack the antimuscarinic actions of tricyclic antidepressants.

Clinical features

Even large overdoses appear to be relatively safe unless potentiated by ethanol. Most patients will show no signs of toxicity but drowsiness, nausea, diarrhoea and sinus tachycardia have been reported. Rarely, junctional bradycardia, seizures and hypertension have been encountered and influenza-like symptoms may develop.

Treatment

Supportive measures are all that are required.

FURTHER READING

Isbister GK, Bowe SJ, Dawson A et al. Relative toxicity of selective serotonin reuptake inhibitors (SSRIs) in overdose. *Clinical Toxicology* 2004; **42**: 277–285.

Theophylline

Poisoning may complicate therapeutic use as well as being the result of deliberate self-poisoning. If a slow-release preparation is involved, peak plasma concentrations are not attained until 6–12 hours after overdosage and the onset of toxic features is correspondingly delayed.

Clinical features

Nausea, vomiting, hyperventilation, haematemesis, abdominal pain, diarrhoea, sinus tachycardia, supraventricular and ventricular arrhythmias, hypotension, restlessness, irritability, headache, hyperreflexia, tremors and convulsions have been observed. Hypokalaemia probably results from activation of Na^+/K^+ ATPase. A mixed acid–base disturbance is common. Most symptomatic patients have plasma theophylline concentrations in excess of 25 mg/L (430 µmol/L). Convulsions are seen more commonly when concentrations are >50 mg/L (>860 µmol/L). Plasma potassium concentrations of <2.6 mmol/L, metabolic acidosis, hypotension, seizures and arrhythmias are indications of severe poisoning.

Treatment

There is good evidence that multiple-dose (12.5 g/h) activated charcoal enhances the elimination of theophylline. However, protracted theophylline-induced vomiting may mitigate the benefit of this therapy, unless vomiting is suppressed by a $5HT_3$ antagonist such as ondansetron 8 mg i.v. in an adult. Correction of hypokalaemia to prevent or treat tachyarrhythmias is of great importance. A non-selective β-adrenoceptor blocking drug, such as propranolol, is also useful in the treatment of tachyarrhythmias secondary to hypokalaemia, but should not be given if the patient has severe airways disease. Convulsions should be treated with diazepam 10–20 mg i.v. in an adult.

FURTHER READING

Minton NA, Henry JA. Acute and chronic human toxicity of theophylline. *Human and Experimental Toxicology* 1996; **15**: 471–481.

Tricyclic antidepressants

Tricyclic antidepressants block the reuptake of norepinephrine into peripheral and intracerebral neurones thereby increasing the concentration of monoamines in these areas. These drugs also have antimuscarinic actions and class 1 antiarrhythmic (quinidine-like) activity.

Clinical features

Features of poisoning usually appear within 60 minutes after ingestion. Drowsiness, sinus tachycardia, dry mouth, dilated pupils, urinary retention, increased reflexes and extensor plantar responses are the most common features of mild poisoning. Severe intoxication leads to coma, often with divergent strabismus and convulsions. Plantar, oculocephalic and oculovestibular reflexes may be abolished temporarily. An ECG will often show a wide QRS interval and there is a reasonable correlation between the width of the QRS complex and the severity of poisoning. Metabolic acidosis and cardiorespiratory depression are observed in severe cases.

Treatment

The majority of patients recover with supportive therapy alone (adequate oxygenation, control of convulsions and cor-

rection of acidosis), though a small percentage of patients will require mechanical ventilation for 24–48 hours. The onset of supraventricular tachycardia and ventricular tachycardia should be treated with sodium bicarbonate (8.4%) 50 mmol intravenously over 20 minutes, even if there is no acidosis present. If ventricular tachycardia is compromising cardiac output, lidocaine 50–100 mg i.v. should be administered in an adult.

FURTHER READING

Bradberry SM, Thanacoody HKR, Watt BE et al. Management of the cardiovascular complications of tricyclic antidepressant poisoning: role of sodium bicarbonate. *Toxicological Reviews* 2005; **24**: 195–204.

Thanacoody HK, Thomas SH. Tricyclic antidepressant poisoning: cardiovascular toxicity. *Toxicological Reviews* 2005; **24**: 205–214.

SPECIFIC POISONS: MARINE ANIMALS

Ciguatera fish poisoning

Over 400 fish species have been reported as ciguatoxic (*Cigua* is Spanish for poisonous snail), though barracuda, red snapper, amberjack and grouper are most commonly implicated. Ciguatera fish contain ciguatoxin, maitotoxin and scaritoxin, which are lipid-soluble, heat-stable compounds that are derived from dinoflagellates such as *Gambierdiscus toxicus*. Ciguatoxin opens voltage-sensitive sodium channels at the neuromuscular junction and maitotoxin opens calcium channels of the cell plasma membrane.

Clinical features

The onset of symptoms occurs from a few minutes to 30 hours after ingestion of toxic fish. Typically features appear between 1 and 6 hours and include abdominal cramps, nausea, vomiting and watery diarrhoea. In some cases, numbness and paraesthesiae of the lips, tongue and throat occur. Other features described include malaise, dry mouth, metallic taste, myalgia, arthralgia, blurred vision, photophobia and transient blindness. In more severe cases, hypotension, cranial nerve palsies and respiratory paralysis have been reported. The mortality in severe cases may be as high as 12%. Recovery takes from 48 hours to one week in the mild form and from one to several weeks in the severe form.

Treatment

Treatment is symptomatic, although in a few patients atropine has lessened some of the cardiovascular and gastrointestinal manifestations.

FURTHER READING

Sobel J, Painter J. Illnesses caused by marine toxins. *Clinical Infectious Diseases* 2005; **41**: 1290–1296.

Paralytic shell fish poisoning

This is uncommon and is caused by bivalve molluscs being contaminated with neurotoxins, including saxitoxin, produced by the toxic dinoflagellates *Gonyanlax catenella* and *Gonyanlax tamarensis.* Saxitoxin blocks voltage-gated sodium channels in nerve and muscle cell membranes, blocking nerve signal transmission.

Clinical features

Symptoms develop within 30 minutes of ingestion. The illness is characterized by paraesthesiae of the mouth, lips, face and extremities and is often accompanied by nausea, vomiting and diarrhoea. In more severe cases, dysphonia, dysphagia, muscle weakness, paralysis, ataxia and respiratory depression occur.

Treatment

Treatment is symptomatic and supportive.

FURTHER READING

Clark RF, Girard RH, Rao D et al. Stingray envenomation: a retrospective review of clinical presentation and treatment in 119 cases. *Journal of Emergency Medicine* 2007; **33**: 33–37.

Scombroid fish poisoning

This is due to the action of bacteria such as *Proteus morgani* and *Klebsiella pneumoniae* in decomposing the flesh of fish such as tuna, mackerel, mahi-mahi, bonito and skipjack if the fish are stored at insufficiently low temperatures. The spoiled fish can contain excessively high concentrations of histamine (muscle histidine is broken down by the bacteria to histamine), though the precise role of histamine in the pathogenesis of the clinical syndrome is uncertain.

Clinically the mean incubation period is 30 minutes. The illness is characterized by flushing, headache, sweating, dizziness, burning of the mouth and throat, abdominal cramps, nausea, vomiting and diarrhoea and is usually short lived; the mean duration is 4 hours.

Treatment is symptomatic and supportive. Antihistamines may alleviate the symptoms.

Stings from marine animals

Several species of fish have venomous spines in their fins. These include the weaver fish, short-spine cottus, spiny dogfish and the stingray. Bathers and fishermen may be stung if they tread on or handle these species.

The immediate result of a sting is intense local pain, swelling, bruising, blistering, necrosis and, if the poisoned spine is not removed, chronic sepsis (though this is uncommon). Occasionally systemic symptoms including, vomiting, diarrhoea, hypotension and tachycardia occur.

Treatment by immersing the affected part in hot water may relieve local symptoms as this denatures the thermolabile toxin.

Jellyfish

Most of the jellyfish found in North European coastal waters are non-toxic as their stings cannot penetrate human skin. A notable exception is the 'Portuguese man-o'-war' (*Physalia physalis*) whose sting contains a toxic peptide, phospholipase A, and a histamine-liberating factor. Toxic jellyfish are found more frequently in Australia and some cause the Irukandji syndrome.

Clinical features

Local pain occurs followed by myalgia, nausea, griping abdominal pain, dyspnoea and even death. The cluster of severe systemic symptoms that constitute the Irukandji syndrome occur some 30 minutes after the jellyfish sting. The symptoms include severe low back pain, excruciating muscle cramps in all four limbs, the abdomen and chest, sweating, anxiety, restlessness, nausea, vomiting, headache, palpitations, life-threatening hypertension, pulmonary oedema and toxic global heart dilatation.

Treatment

Adhesive tape may be used to remove any tentacles still adherent to the bather. Local application of 5% acetic acid is said to prevent stinging cells adherent to the skin discharging. Local analgesia and antihistamine creams provide symptomatic relief. Other features should be treated symptomatically and supportively.

FURTHER READING

Bailey PM, Little M, Jelinek GA et al. Jellyfish envenoming syndromes: unknown toxic mechanisms and unproven therapies. *Medical Journal of Australia* 2003; **178**: 34–37.

SPECIFIC POISONS: VENOMOUS ANIMALS

Insect stings and bites

Insect stings from wasps and bees, and bites from ants, produce pain and swelling at the puncture site. Following the sting or bite, patients should be observed for 2 hours for any signs of evolving urticaria, pruritus, bronchospasm or oropharyngeal oedema. The onset of anaphylaxis requires urgent treatment (see p. 73).

Scorpions

Scorpion stings are a serious problem in North Africa, the Middle East and the Americas. Scorpion venoms stimulate the release of acetylcholine and catecholamines causing both cholinergic and adrenergic symptoms.

Severe pain occurs immediately at the site of puncture, followed by swelling. Signs of systemic involvement, which may be delayed for 24 hours, include vomiting, sweating, piloerection, abdominal colic, diarrhoea. In some cases, depending on the species, shock, respiratory depression and pulmonary oedema may develop.

Treatment by local infiltration with anaesthetic or a ring block will usually alleviate local pain, though systemic analgesia may be required. Specific antivenom, if available, should be administered as soon possible.

Spiders

The black widow spider (*Latrodectus mactans*) is found in North America and the tropics and occasionally in Mediterranean countries.

The bite quickly becomes painful and generalized muscle pain, sweating, headache and shock may occur.

No systemic treatment is required except in cases of severe systemic toxicity, when specific antivenom should be given, if this is available.

Venomous snakes

Approximately 15% of the 3000 species of snake found world-wide are considered to be dangerous to humans. Snake bite is common in some tropical countries. For example, in Sri Lanka there are some 6 bites per 100 000 population and 900 deaths per year. In Nigeria there are some 500 bites per 100 000 population with a 12% mortality. In Myanmar there are 15 deaths per 100 000 population from snake bites. In the USA (population 301 million) there are some 45 000 bites per year, 8000 of which are by venomous species, with some six deaths annually. In the UK (population 60 million) approximately 100 people are admitted to hospital annually but no deaths have occurred since 1970. In Australia (population 20 million) there are two or three deaths annually.

There are three main groups of venomous snakes, representing some 200 species, which have in their upper jaws a pair of enlarged teeth (fangs) that inject venom into the tissues of their victim. These are Viperidae (with two subgroups: Viperinae – European adders and Russell's vipers; and Crotalinae – American rattlesnakes, moccasins, lance-headed vipers and Asian pit vipers), Elapidae (cobras, kraits, mambas, coral snakes, Australian venomous snakes) and Hydrophiidae (sea snakes). In addition, some members of the family Colubridae are mildly venomous (mongoose snake).

Fig. 16.7 **Snake bite showing swelling at site.**

More online
www.studentconsult.com

Clinical features

Viperidae (Viperinae and Crotalinae)

Russell's viper causes most of the snake-bite mortality in India, Pakistan and Burma. There is local swelling at the site of the bite (Fig. 16.7) which may become massive. Local tissue necrosis may occur, particularly with cobra bites. Evidence of systemic involvement occurs within 30 minutes, including vomiting, evidence of shock and hypotension. Haemorrhage due to incoagulable blood can be fatal.

Elapidae

There is not usually any swelling at the site of the bite, except with Asian cobras and African spitting cobras – here the bite is painful and is followed by local tissue necrosis. Vomiting occurs first followed by shock and then neurological symptoms and muscle weakness, with paralysis of the respiratory muscles in severe cases. Cardiac muscle can be involved.

Hydrophiidae

Systemic features are muscle involvement, myalgia and myoglobinuria, which can lead to acute renal failure. Cardiac and respiratory paralysis may occur.

Management

As a first aid measure, a firm pressure bandage should be placed over the bite and the limb immobilized. This may delay the spread of the venom. Arterial tourniquets should *not* be used, and incision or excision of the bite area should not be performed. Local wounds often require little treatment. If necrosis is present, antibiotics should be given. Skin grafting may be required later. Antitetanus prophylaxis must be given. The type of snake should be identified if possible.

In about 50% of cases no venom has been injected by the bite. Nevertheless, careful observation for 12–24 hours is necessary in case systemic features (envenomation) develop. General supportive measures should be given, as necessary. These include intravenous fluids with volume expanders for hypotension and diazepam for anxiety. Treatment of acute respiratory, cardiac and renal failure is instituted as necessary.

Antivenoms are not generally indicated unless envenomation is present, as they can cause severe allergic reactions. Antivenoms can rapidly neutralize venom, but only if an amount in excess of the amount of venom is given. Large quantities of antivenom may be required. As antivenoms cannot reverse the effects of the venom, they must be given early to minimize some of the local effects and may prevent necrosis at the site of the bite. Antivenoms should be administered intravenously by slow infusion, the same dose being given to children and adults.

Allergic reactions are frequent, and adrenaline (epinephrine) 1 in 1000 solution should be available. In severe cases, the antivenom infusion should be continued even if an allergic reaction occurs, with subcutaneous injections of adrenaline being given as necessary. Some forms of neurotoxicity, such as those induced by the death adder, respond to anticholinesterase therapy with neostigmine and atropine.

FURTHER READING

Warrell DA. Venomous animals. *Medicine* 2007; **35**: 659–662.

SPECIFIC POISONS: PLANTS

Many plants are known to be poisonous, but it is unusual for severe poisoning to occur. In some regions of the world, however, plant poisonings cause substantial morbidity and mortality. While deaths do occur after unintentional poisoning with plants such as *Atractylis gummifera* (bird-lime or blue thistle) and *Blighia sapida* (ackee tree), the majority of deaths globally occur following intentional self-poisoning with plants such as *Thevetia peruviana* (yellow oleander).

FURTHER READING

Bradberry S, Vale A. Plants. *Medicine* 2007; **35**: 649–651.

Atropa belladonna

Atropa belladonna (deadly nightshade) contains hyoscyamine and atropine and causes antimuscarinic effects – a dry mouth, nausea and vomiting – leading to blurred vision, hallucinations, confusion and hyperpyrexia.

Cicuta species

Cicuta spp. (water hemlock) and the related genus *Oenanthe* contain cicutoxin, a potent central nervous system (CNS) stimulant that produces violent seizure activity. The CNS effects of cicutoxin are similar to those of picrotoxin, a known inhibitor of GABA. Severe gastrointestinal symptoms, diaphoresis, salivation and skeletal muscle stimulation may precede the seizure activity.

Conium maculatum

Conium maculatum (poison hemlock) contains a variety of volatile pyridine alkaloids, including coniine, *N*-methylconiine and gamma-coniceine. Coniceine is significantly more toxic than coniine and is thought to be the precursor to coniine. The toxic activity of the alkaloids is similar to that of nicotine. Large doses produce non-polarizing neuromuscular blockade which may result in respiratory depression and death.

Datura stramonium

Datura stramonium (jimsonweed) and other *Datura* species contain L-hyoscyamine and atropine. These alkaloids are potent antagonists of acetylcholine at muscarinic receptors and produce the anticholinergic syndrome. While morbidity is significant, fatalities are rare and are the consequence of hyperthermia, seizures and/or arrhythmias.

Digitalis purpurea, Nerium oleander, Thevetia peruviana

Ingestion of *Digitalis purpurea,* or the common (*Nerium oleander*) or yellow (*Thevetia peruviana*) oleander can produce a syndrome similar to digoxin poisoning. A randomized controlled trial has shown that digoxin-specific antibody fragments rapidly and safely reverse yellow oleander-induced arrhythmias, restore sinus rhythm, and rapidly reverse bradycardia and hyperkalaemia. The administration of multiple doses of activated charcoal is used, but the effect on survival is debated.

FURTHER READING

Persson H. Mushrooms. *Medicine* 2007; **35**: 635–637.

SPECIFIC POISONS: MUSHROOMS

Most mushrooms and edible fungi are not poisonous, but transient nausea, vomiting and diarrhoea can occur after ingestion of some varieties. There are a few species which are toxic.

Amanita muscaria, Amanita pantherina

Amanita muscaria (fly agaric) and *Amanita pantherina* (false blusher) contain isoxazoles (γ-aminobutyric acid agonists) which cause inebriation, visual disturbances, hallucinations, myoclonic jerks, muscle fasciculation, convulsions and coma. Treatment is supportive.

Amanita phalloides

Amanita phalloides (death cap mushroom) contains amatoxins which inhibit transcription from DNA to mRNA by blockade of nuclear RNA polymerase II, resulting in impaired protein synthesis and cell death. Amatoxins are not inactivated by cooking. Hepatocellular necrosis follows 72 hours after ingestion and some varieties cause renal failure. Haemodialysis and liver/renal transplantation may be necessary. The value of silibinin and *N*-acetylcysteine is unproven.

Coprinus atramentarius

Coprinus atramentarius (ink cap) contains coprin, an acetaldehyde dehydrogenase inhibitor with a disulfiram-like effect, causing flushing and rash. Treatment is supportive.

Cortinarius orellanus, Cortinarius speciossimus

Cortinarius orellanus and *Cortinarius speciossimus* contain orellanin, a potent nephrotoxin. Renal function may recover only partially.

'Magic' mushrooms

'Magic' mushrooms, such as *Psilocybe*, *Panaeolus*, *Conocybe*, *Gymnopilus*, *Stropharia*, *Pluteus* and *Panaeolina* spp., contain the hallucinogen, psilocybin. Treatment is supportive.

SIGNIFICANT WEBSITES

www.toxbase.co.uk

Toxbase – Database of UK National Poisons Information Service Service

www.toxinz.com

Database of the New Zealand Poisons Centre

www.toxnet.nlm.nih.gov

National Library of Medicine's Toxnet

www.who.int/ipcs/poisons/centre/directory/en

Contact details of all poisons centres world-wide

www.wikitox.org

Home of the Clinical Toxicology Teaching Resource Project

17 Environmental medicine

Disease and the environment 953
Heat 953
Cold 954
High altitudes 955
Diving 956
Drowning and near-drowning 957
Ionizing radiation 958
Electric shock 959
Smoke 959
Noise 959
Bioterrorism/biowarfare 959
Travel 960
Building-related illnesses 960

DISEASE AND THE ENVIRONMENT

The incidence and prevalence of disease and causes of death within a community are a reflection of interrelated factors:

- Genetic predisposition.
- Nutrition, wealth and poverty.
- Purity of water sources, sanitation and atmospheric pollution.
- Environmental disasters and accidents.
- Background ionizing radiation and man-made radiation exposure, deliberate or accidental.
- Environmental temperature.
- Patterns of infective disease.
- Political forces determining levels of healthcare, preventative strategies and effects of war on civilian populations.

Some of these environmental effects have been clearly documented within the last decade, for example the civilian mortality during the current Iraq War, and loss of life and disease prevalence following the 2006 *tsunami*, the 2008 earthquake in Sechuan and cyclone *Nargris* in the Irawaddy delta. Flooding caused by *El Niño* in East Africa not only resulted in an increase in breeding sites for mosquito vectors but a major outbreak of rift valley fever due to the enforced close proximity of cattle with humans.

Environmental temperature

The effect of environmental temperature (T_{Env}) is paramount in infective diseases; changes as small as 1°C cause major changes in disease vectors. Climate change is an unquestionable phenomenon. Patterns of infective disease are likely to change radically within the next 20 years. There are suggestions that these effects are already becoming apparent, for example:

- Patterns of malaria in South East Asia.
- The occurrence of dengue fever in southern Italy.
- Outbreaks of cholera, and seasonal variation in diarrhoea and vomiting.

Research into the effect of climate change on the changing patterns of infective diseases will point to potential ways in which national and international efforts can be targeted.

HEAT

Body core temperature (T_{Core}) is maintained at 37°C by the thermoregulator centre in the hypothalamus.

Heat is produced by cellular metabolism and is dissipated through the skin by both vasodilatation and sweating and in expired air via the alveoli. When the environmental temperature (T_{Env}) is >32.5°C, profuse sweating occurs. *Sweat evaporation* is the principal mechanism for controlling T_{Core}, following exercise or in response to an increase in T_{Env}.

Heat acclimatization. *Acclimatization* to heat takes place over several weeks. The sweat volume increases and the sweat salt content falls. Increased evaporation of sweat reduces T_{Core}.

Heat cramps. Painful muscle cramps, usually in the legs, often occur in fit people when they exercise excessively, especially in hot weather. Cramps are probably due to low extracellular sodium caused by excess intake of water over salt. Cramps can be prevented by increasing dietary salt. They respond to combined salt and water replacement, and in the acute stage to stretching and muscle massage. T_{Core} remains normal.

Heat illness. At any environmental temperature (especially with T_{Env} of >25°C), and with a high humidity, exercise in clothes that inhibit sweating can cause an elevation in T_{Core} in less than 15 minutes. Weakness/exhaustion, cramps, dizziness and syncope, with T_{Core} > 37°C, define *heat illness* (heat exhaustion). Elevation of T_{Core} is more critical than water

and sodium loss. *Heat illness* may progress to *heat injury*, a serious emergency.

Management

Reduce (T_{Env}) if possible. Cool with sponging and fans. Give O_2 by mask. Other causes of high T_{Core}, e.g. malaria, should be ruled out if necessary.

Oral rehydration with *both* salt and water (25 g of salt per 5 L of water/day) is given in the first instance, with adequate replacement thereafter. In severe heat illness, i.v. fluids are needed. 0.9% saline is given; monitor serum sodium and correct secondary potassium loss.

Heat injury

Heat injury (heat stroke) is an acute life-threatening situation when T_{Core} rises above 41°C. There is headache, nausea, vomiting and weakness, progressing to confusion, coma and death. The skin *feels* intensely hot to the touch. Sweating is often absent, but not invariably.

Heat injury can develop in unacclimatized people in hot, humid windless climates, even without exercise. Sweating may be limited by prickly heat. Excessive exercise in inappropriate clothing, e.g. exercising on land in a wetsuit, can lead to heat injury in temperate climates. Diabetes and drugs, e.g. alcohol, antimuscarinics, diuretics and phenothiazines, can contribute. *Heat injury* can lead to a fall in cardiac output, lactic acidosis and intravascular coagulation.

Prevention

Acclimatization, fluids, avoiding inappropriate clothing and common sense.

Management

- Reduce T_{Env} if possible.
- Cool: sponging, icepacks, fanning. Give O_2 by mask.
- Manage in intensive care: monitor cardiac ouput, respiration, biochemistry, clotting and muscle enzymes.
- Give fluids intravenously with caution: intravascular volume can remain normal.

Prompt treatment is essential and can be curative even with T_{Core} of more than 41°C. Delay can be fatal. Complications are hypovolaemia, intravascular coagulation, cerebral oedema, rhabdomyolysis, and renal and hepatic failure.

Malignant hyperpyrexia

See p. 1182.

FURTHER READING

Simon H. Hyperthermia. *New England Journal of Medicine* 1993; **329:** 483–487.

COLD

Hypothermia is defined as a core temperature of less than 32°C. It is frequently lethal when T_{Core} falls below 30°C. Survival, with full recovery has however been recorded with T_{Core} of less than 16°C.

Cold injury includes:

- *Frostbite:* the local cold injury that follows freezing of tissue.
- *Non-freezing cold injury:* the damage – usually to feet – following prolonged exposure to T_{Env} between 0 and 5°C, usually in damp conditions.

Hypothermia

Hypothermia occurs in many settings.

At home. Hypothermia can occur when T_{Env} is below 8°C, when there is poor heating, inadequate clothing and poor nutrition. Depressant drugs, e.g. hypnotics, alcohol, hypothyroidism or intercurrent illness also contribute. Hypothermia is commonly seen in the poor, frail and elderly. The elderly have diminished ability to sense cold and also have little insulating fat. Neonates and infants become hypothermic rapidly because of a relatively large surface area in proportion to subcutaneous fat.

Outdoors on land. Hypothermia is a prominent cause of death in climbers, skiers, polar travellers and in wartime. Wet, cold conditions with wind chill, physical exhaustion, injuries and inadequate clothing are contributory. Children are especially susceptible.

Cold water immersion. Dangerous hypothermia can develop following immersion for more than one hour in water temperatures of 15–20°C. In T_{Water} below 12°C limbs rapidly become numb and weak. Recovery takes place gradually, several hours following rescue.

Clinical features

Mild hypothermia (T_{Core} < 32°C) causes shivering and initially intense discomfort. However, the hypothermic subject, though alert, may not act appropriately to rewarm, e.g. by huddling, wearing extra clothing or exercising. As the T_{Core} falls below 32°C, severe hypothermia causes impaired judgement – including lack of awareness of cold – drowsiness and coma. Death follows, usually from ventricular fibrillation.

Diagnosis

Diagnosis is straightforward, if a low-reading thermometer is available. If not, rapid clinical assessment is reliable. Someone who *feels* icy to touch – abdomen, groin, axillae – is *probably* hypothermic. If they are clammy, uncooperative or sleepy, T_{Core} is almost *certainly* <32°C.

Sequelae

Pulse rate and systemic BP fall. Cardiac output and cerebral blood flow are greatly affected by change in posture in hypothermia, and can fall further if the upright position is maintained, or the thorax restrained by a harness. This is why air-sea helicopter winch rescues are often carried out using a stretcher, rather than a chest harness. Respiration becomes shallow and slow. Muscle stiffness develops; tendon reflexes become sluggish and/or lost.

As coma ensues, pupillary and other brainstem reflexes are lost; pupils are fixed and may be dilated in severe hypothermia.

Metabolic changes are variable, with either metabolic acidosis or alkalosis. Arterial P_{O_2} may appear normal, i.e. falsely high.

Reduction in the temperature of haemoglobin shifts the oxygen dissociation curve to the left. Thus, if an arterial blood sample from a hypothermic patient is analysed at 37°C, the P_{O_2} will be falsely high. Within the range 37–33°C this factor is around 7% per °C. Many blood gas machines also calculate the arterial saturation; this too will be falsely high. When a patient is monitored using a pulse oximeter, the level of arterial oxygen saturation (S_aO_2) will however be correct – but if S_aO_2 is then converted by calculation to P_aO_2, a downwards correction must be applied – simply due to hypothermia.

Bradycardia with 'J' waves (rounded waves above the isoelectric line at the junction of the QRS complex and ST segment (Fig. 17.1)) are pathognomic of hypothermia. Prolongation of PR and QT intervals and QRS complex also occur. Ventricular dysrhythmias (tachycardia/fibrillation) or asystole are the usual causes of death.

Principles of management

- Maintain the patient horizontal, or slightly head down.
- Rewarm gradually.
- Correct metabolic abnormalities.
- Anticipate and treat dysrhythmias.
- Check for hypothyroidism (see p. 985).

If the patient is awake, with a core temperature of more than 32°C, place them in a warm room, use a foil wrap and give warm fluids orally. Outdoors, add extra dry clothing, huddle together and use a warmed sleeping bag. Rewarming may take several hours. Avoid alcohol: this adds to confusion, boosts confidence factitiously, causes peripheral vasodilatation and further heat loss, and can precipitate hypoglycaemia.

Severe hypothermia

In severe hypothermia, people look dead. Always exclude hypothermia before diagnosing brainstem death (see p. 920). Warm gradually, aiming at a 1°C per hour increase in T_{Core}. Direct mild surface heat from an electric blanket can be helpful. Treat any underlying condition promptly, e.g. sepsis. Monitor all vital functions. Correct dysrhythmias. Drug screening is essential.

Give warm i.v. fluids slowly. Correct metabolic abnormalities. Hypothyroidism, if present, should be treated with liothyronine. Various methods of artificial rewarming exist – inhaled warm humidified air, gastric or peritoneal lavage, and haemodialysis. These are rarely used.

Prevention

Hypothermia in the field can often be prevented by forethought and action. For the elderly, improved home heating and insulation, central heating in bedrooms and electric blankets are helpful in cold spells. Finance is often needed for this.

Cold injury

Frostbite

Ice crystals form within skin and superficial tissues when the temperature of the tissue (T_{Tissue}) falls to −3°C: T_{Env} generally must be below −6°C. Wind chill is frequently a factor. Typically, fingers, toes, nose and ears become frostbitten.

Frostbitten tissue is pale, greyish and initially doughy to touch. Later, tissue freezes hard, looking like meat from a freezer. Frostbite can easily occur when working or exercising in low temperatures and typically develops without the patient's knowledge. Below T_{Enu} −5°C, hands or feet that have lost their feeling are at risk of cold injury.

Management

Transport the patient – or if this is impossible, make them walk, even on frostbitten feet – to a place of safety before commencing warming. Warm the frozen part by immersion in hand hot water at 39–42°C, if feasible. Assess hypothermia. Continue warming until obvious thawing occurs; this can be painful. Vasodilator drugs have no part in management. Blisters form within several days and, depending on the depth of frostbite, a blackened shell – the *carapace* – develops as blisters regress or burst. Dry, non-adherent dressings and aseptic precautions are essential, though hard to achieve. Frostbitten tissues are anaesthetic and at risk from further trauma and infection. Recovery takes place over many weeks, and may be incomplete. Surgery may be needed, but should be avoided in the early stages.

Chillblains

These are small, purplish itchy inflammatory lesions, occurring on toes and fingers. They occur in cold, wet conditions. They are more common in women and heal in 7–14 days. Prevention is by keeping warm and wearing gloves and warm footwear.

Non-freezing cold injury (NFCI)

NFCI (trench foot) describes tissue damage following prolonged exposure, usually for several hours or more, at T_{Env} around or slightly above freezing, but without frostbite. Wet socks and boots are the usual cause. There is severe vasoconstriction, blotchiness of the lower limbs, with pain and oedema on rewarming. Recovery usually follows over several weeks. There may be prolonged late susceptibility to cold. NFCI is a prominent cause of morbidity in troops operating in low temperatures.

Prevention of frostbite and NFCI is largely by education and common sense: avoid damp feet and wet boots. Always carry spare dry socks, gloves and headgear.

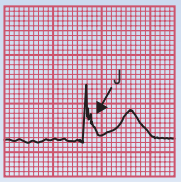


Fig. 17.1 **ECG showing J waves.**

FURTHER READING

Lazar HL. Editorial: The treatment of hypothermia. *New England Journal of Medicine* 1997; **337:** 1545–1547.

HIGH ALTITUDES

The partial pressure of atmospheric – and hence alveolar and arterial – oxygen falls in a near-linear relationship to altitude as barometric pressure falls (Fig. 17.2).

Commercial aircraft are pressurized to 2750 m (lowering the oxygen saturation by 3–4%). The incidence of deep venous thrombosis and pulmonary embolism is slightly greater in sedentary passengers on long-haul flights than in a similar population at sea level. Dehydration and alcohol probably contribute. Prophylactic aspirin is not recommended.

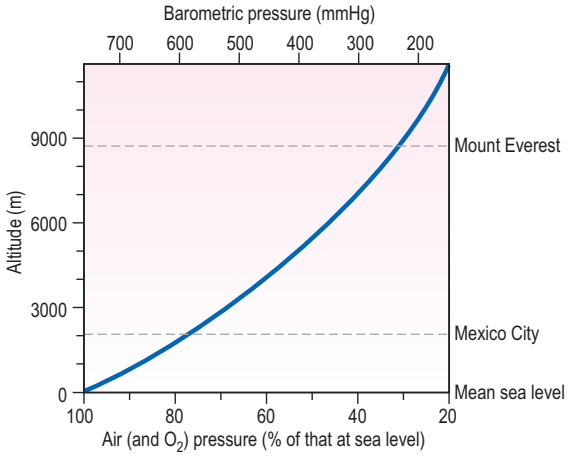


Fig. 17.2 **The decrease in oxygen and barometric pressure with increasing altitude.**

On land, below 3000 m there are few clinical effects. The resulting hypoxaemia causes breathlessness only in those with severe cardiorespiratory disease. Above 3000–3500 m hypoxia causes a spectrum of related syndromes that affect high-altitude visitors, principally climbers, trekkers, skiers and troops (Table 17.1). These conditions occur largely during acclimatization. Acclimatization takes several weeks and once completed can enable man to live – permanently if necessary – up to about 5600 m. At greater heights, although people can survive for days or weeks, deterioration due to chronic hypoxia is inevitable. The world's highest railway to Lhasa travels over 5000 m. Oxygen is provided in the carriages. Roads to similar altitudes are used extensively for transport in central Asia. Ascent of the world's highest 8000 m summits is just possible without supplementary oxygen. At the summit of Everest (8848 m) the barometric pressure is 34 kPa (253 mmHg). An acclimatized mountaineer has an alveolar P_{O_2} of 4.0–4.7 kPa (30–35 mmHg) – near man's absolute physiological limit.

Acute mountain sickness (AMS)

AMS describes malaise, nausea, headache and lassitude and affects the majority of people for a few days above 3500 m. Following arrival at this altitude there is usually a latent interval of 6–36 hours before symptoms begin. Treatment is rest, with analgesics if necessary. Recovery is usually spontaneous over several days.

Prophylactic treatment with acetazolamide, a carbonic anhydrase inhibitor and a respiratory stimulant, is of some value in preventing AMS. Acclimatizing, i.e. ascending gradually, provides better, and more natural prophylaxis.

In the minority, more serious sequelae – high-altitude pulmonary oedema and high-altitude cerebral oedema develop.

High-altitude pulmonary oedema

Predisposing factors include youth, rapidity of ascent, heavy exertion and severe AMS. Breathlessness, occasionally with frothy blood-stained sputum, indicates established oedema. Unless treated rapidly this leads to cardiorespiratory failure and death. Milder forms are common. Breathlessness at rest should raise the suspicion of pulmonary oedema.

High-altitude cerebral oedema

Cerebral oedema is the result of abrupt increase in cerebral blood flow that occurs even at modest altitudes of 3500–4000 m. It is unusual below 4500 m, and usually occurs in the first two weeks, during acclimatization. Cerebral oedema can also develop suddenly in well-acclimatized climbers above 7000 m. Headache is typical, and is accompanied by drowsiness, ataxia and papilloedema. Coma and death follow if brain oedema progresses.

Table 17.1 Conditions caused by sustained hypoxia

Condition	Incidence (%)	Usual altitude (m)
Acute mountain sickness	70	3500–4000
Acute pulmonary oedema	2	4000
Acute cerebral oedema	1	4500
Retinal haemorrhage	50	5000
Deterioration	100	6000
Chronic mountain sickness	Rare	4500

Treatment

Any but the milder forms of AMS require urgent treatment. Oxygen should be given by mask if available, and descent to a lower altitude should take place as quickly as possible. Nifedipine reduces pulmonary hypertension and is used in the treatment of pulmonary oedema. Dexamethasone is effective in reducing brain oedema. Portable pressure bags inflated by a foot pump are widely used; the patient is enclosed in the pressure bag.

Retinal haemorrhages

Small 'flame' haemorrhages within the retinal nerve fibre layer are common above 5000 m and usually symptomless. Rarely a haemorrhage will cover the macula, causing painless loss of central vision. Recovery is usual.

Deterioration

Prolonged residence between 6000 and 7000 m leads to weight loss, anorexia and listlessness after several weeks. Above 7500 m, the effects of deterioration become apparent more rapidly over several days, although it is possible to survive for a week or more at altitudes near 8000 m without supplementary oxygen.

Chronic mountain sickness

Chronic mountain sickness occurs in long-term residents at high altitudes, usually after several decades and is seen in the Andes and in central Asia.

Headache, polycythaemia, lassitude, cyanosis, finger clubbing, congested cheeks and ear lobes, and right ventricular enlargement develop. Chronic mountain sickness is gradually progressive.

By contrast, coronary artery disease and hypertension are rare in high-altitude native populations.

FURTHER READING

Clarke C. High altitude and mountaineering expeditions. In: Warrell D, Anderson S (eds) *Expedition Medicine,* 2nd edn. London: Royal Geographical Society, 2008.

Grocott MPW, Martin DS, Levett DZH et al. Arterial blood gases and oxygen content in climbers on Mount Everest. *New England Journal of Medicine* 2009; **360**: 140–149.

Information: Wilderness Medical Society, PO Box 2463, Indianapolis, Indiana 462206, USA. http://www.wms.org.

UIAA Mountain Medicine Data Centre leaflets, available from British Mountaineering Council, 177–179 Burton Road, Manchester M20 2BB, UK. http://www.thebmc.co.uk.

DIVING

Free diving by breath-holding is possible to around 5 m, or to greater depths with practice. Various methods are used to supply air to a diver. With the simplest, e.g. a snorkel, providing air at around 0.5 m, the limiting factor is inspiratory

Table 17.2 Depth and pressure

Water depth (m)	Pressure	
	atmospheres	mmHg
0	1	760
10	2	1520
50	6	4560
90	10	7600

effort. At depths >0.5 m, i.e. with a longer snorkel tube, forced negative-pressure ventilation can cause pulmonary capillary damage and haemorrhagic alveolar oedema. Scuba divers – the usual recreational sports diving down to 30 m – carry bottled compressed air, or a nitrogen–oxygen mixture. Divers who work at great depths commercially or in underwater exploration breathe helium–oxygen or nitrogen– oxygen mixtures, delivered by hose from the surface. Ambient pressures at various depths are shown in Table 17.2.

Problems during descent

Middle ear barotrauma (*squeeze)* is common and caused by inability to equalize pressure in the middle ear – usually the result of Eustachian tube blockage. Hearing loss occurs, sometimes with tympanic membrane rupture and acute vertigo. *Sinus squeeze* is due to blockage of the nasal and paranasal sinus ostia – there is intense local pain. *Treatment* is by holding the nostrils closed and swallowing, or similar manoeuvres – and decongestants. Avoid diving with a respiratory or sinus infection.

Oxygen narcosis

Pure oxygen is not used for diving because of oxygen toxicity. Lung damage – atelectasis, endothelial cell damage and pulmonary oedema – occurs when the alveolar oxygen pressure exceeds 1.5 atmospheres, a depth around 5 m. At around 10 m the CNS begins to be affected: apprehension, nausea and sweating are followed by muscle twitching and generalized convulsions.

Nitrogen narcosis

When compressed air is breathed below 30 m, narcotic effects of nitrogen begin to impair brain function. This problem sometimes also occurs with nitrogen–oxygen mixtures used for recreational diving. Changes of mood and performance may be hazardous. These resolve on ascent to the surface. Nitrogen narcosis is avoided by replacing air with helium–oxygen mixtures, enabling descent to 700 m. At these extreme depths direct effect of pressure on neurones can cause tremor, hemiparesis and cognitive impairment.

Problems during and following ascent

Free divers who breath-hold often hyperventilate deliberately prior to plunging in. This drives off CO_2 – reducing the stimulus to inspire. During the subsequent breath-hold P_aCO_2 rises; P_aO_2 falls. On surfacing, decompression lowers P_aO_2 further. This can lead to syncope, known as a shallow water blackout. Since loss of consciousness can take place in the water, this can lead to fatalities.

Decompression sickness

Decompression sickness (*the bends)* can follow returning too rapidly to the surface. Problems are caused by the release of bubbles of nitrogen or helium. Decompression tables indicate the duration required for safe return to the surface from a given depth.

The bends can be mild (Type 1, non-neurological bends), with skin irritation and mottling and/or joint pain. Type 2, neurological bends, are more serious – cortical blindness, hemiparesis, sensory disturbances or cord lesions develop.

If bubbles form in pulmonary vessels, divers experience retrosternal discomfort, dyspnoea and cough, known as *the chokes*. These develop within minutes or hours of a dive. Decompression problems do not only occur immediately on reaching the surface, they may take some hours to become apparent. Over the subsequent 24 hours, further ascent, e.g. air travel, can occasionally provoke the bends.

Other problems during ascent include severe paranasal sinus pain, nosebleeds and dental pain – gas bubbles within rotten fillings. Whilst medically minor, these can be dramatic, e.g. causing excruciating pain and a mask filled with blood-stained fluid.

Management: all but the mildest forms of decompression sickness, e.g. skin mottling alone, require recompression in a pressure chamber, following strict international guidelines. Recovery is usual. A long-term problem is aseptic necrosis of the hip due to nitrogen bubbles causing infarction of nutrient arteries. Distinct focal neurological damage may persist, but complaints of fatigue, poor concentration and difficulties with employment are issues compounded by the litigation that commonly follows diving accidents. Objective, evidence-based assessments are essential.

Lung rupture, pneumothorax and surgical emphysema

These emergencies occur particularly when divers breath-hold during emergency ascents after exhausting or losing their bottled gas supply. There is severe dyspnoea, cough and haemoptysis. Pneumothorax and emphysema usually resolve gradually with 100% oxygen. Air embolism can also occur and is treated with recompression/hyperbaric oxygen.

DROWNING AND NEAR-DROWNING

Drowning is the third commonest reason for accidental death in the UK and caused over 500 000 deaths world-wide in 2007. In the UK, some 40% of drownings occur in children under five. People can also drown following an epileptic seizure or a myocardial infarct whilst in water. Exhaustion, alcohol, drugs and hypothermia all contribute to deaths following immersion.

Dry drowning

Between 10% and 15% of drownings occur without water aspiration into the lungs. Laryngeal spasm develops acutely, followed by apnoea and cardiac arrest.

Wet drowning

Fresh or seawater aspiration destroys pulmonary surfactant, leading to alveolar collapse, ventilation/perfusion mismatch

FURTHER READING

Bennett P, Elliot D. *The Physiology and Medicine of Diving*, 5th edn. London: WB Saunders, 2002.

Melamed Y, Shupak A, Bitterman H. Medical problems associated with underwater diving. *New England Journal of Medicine* 1992; **326:** 30–36.

DIVING INFORMATION

Institute of Naval Medicine, Undersea Medicine Division, Alverstoke, Gosport, Hampshire PO12 2DL, UK. http://www.rnreference.mod.uk/09/inm/undersea.htm.

UK Diving, emergencies only: Ministry of Defence, Duty Diving Medical Officer. Tel: 07831 151523.

and hypoxaemia. Aspiration of hypertonic seawater (5% NaCl) pulls additional fluid into the alveoli with further ventilation/perfusion mismatch. In practice, there is little difference between saltwater and freshwater aspiration. In both, severe hypoxaemia develops rapidly. Severe metabolic acidosis develops in the majority of survivors.

Emergency treatment

CPR should be started immediately (see p. 710). Patients have survived for up to 30 minutes under water without suffering brain damage – and sometimes for longer periods if the temperature of the water (T_{Water}) is near 10°C. Survival is probably related to the protective role of the *diving reflex* – submersion causes reflex bradycardia and vasoconstriction. In addition, hypothermia decreases oxygen consumption.

Resuscitation should always be attempted, even with an absent pulse and fixed dilated pupils. Patients frequently make a dramatic recovery. All survivors should be admitted to hospital for intensive monitoring – they are liable to develop effects of acute lung injury during the subsequent 48 hours.

Prognosis

Recovery is frequently complete if the patient regains consciousness within several minutes following initiation of resuscitation but poor if they remain stuporose or in coma at 30 minutes.

FURTHER READING

Modell JH. Drowning. *New England Journal of Medicine* 1993; **328:** 253–256.

IONIZING RADIATION

Ionizing radiation is either penetrating (X-rays, γ-rays or neutrons) or non-penetrating (α- or β-particles). Penetrating radiation affects the skin and deeper tissues, while non-penetrating radiation affects the skin alone. All radiation effects depend on the type of radiation, the distribution of dose and the dose rate.

Dosage is measured in *joules per kilogram*, J/kg; 1 J/kg ≡ 1 gray (1 Gy) ≡ 100 rads.

Radioactivity is measured in *becquerels* (Bq). 1 Bq is defined as the activity of a quantity of radioactive material in which one nucleus decays per second. 3.7×10^{10} Bq = one c*urie* (Ci), the older, non SI unit.

Radiation differs in the density of ionization it causes. Therefore a dose-equivalent called a *sievert* (Sv) is used. This is the absorbed dose weighted for the damaging effect of the radiation. The annual background radiation is approximately 2.5 mSv. A chest X-ray delivers 0.02 mSv, and CT of the abdomen/pelvis about 10 mSv.

Excessive exposure to ionizing radiation follows accidents in industry, nuclear power plants and hospitals and deliberate nuclear explosions designed to eliminate populations – and exceptionally, by poisoning, e.g. with polonium.

Mild acute radiation sickness

Nausea, vomiting and malaise follow doses of approximately 1 Gy. Lymphopenia occurs within several days, followed 2–3 weeks later by a fall in all white cells and platelets.

Acute radiation sickness

Many systems are affected; the extent depends on the dose of radiation (Table 17.3).

Table 17.3 Systemic radiation effects

Acute effects	Delayed effects
Haemopoietic syndrome	Infertility
Gastrointestinal syndrome	Teratogenesis
CNS syndrome	Cataract
Radiation dermatitis	Neoplasia:
	Acute myeloid leukaemia
	Thyroid
	Salivary glands
	Skin
	Others

Haemopoietic syndrome

Absorption of 2–10 Gy is followed by transient vomiting in some individuals, followed by a period of improvement. Lymphocytes are particularly sensitive to radiation damage; severe lymphopenia develops over several days. A decrease in granulocytes and platelets follows 2–3 weeks later, since no new cells are formed in the marrow. Thrombocytopenia with bleeding develops and frequent overwhelming infections, with a high mortality.

Gastrointestinal syndrome

Doses >6 Gy cause vomiting several hours after exposure. This then stops, only to recur some 4 days later accompanied by diarrhoea. The villous lining of the intestine becomes denuded. Intractable bloody diarrhoea follows, with dehydration, secondary infection and sometimes death.

CNS syndrome

Exposures of >30 Gy are followed rapidly by nausea, vomiting, disorientation and coma. Death due to cerebral oedema can follow, usually within 36 hours.

Radiation dermatitis

Skin erythema, purpura, blistering and secondary infection occur. Total loss of body hair is a bad prognostic sign and usually follows an exposure >5 Gy.

Late effects of radiation exposure

Survivors of the US nuclear bombing of Hiroshima and Nagasaki provided data on long-term radiation effects. Risks of acute myeloid leukaemia and cancer, particularly of skin, thyroid and salivary glands, increase. Infertility, teratogenesis and cataract are also late sequelae, developing years after exposure.

The sequelae of *therapeutic* radiation – early, early-delayed and late-delayed radiation effects are discussed on page 466. Focussing techniques are used to target radiation towards the field being treated; radiosensitive structures such as the ovaries are protected by shielding.

Treatment

Acute radiation sickness is an emergency. Absorption of the initial radiation dose can be reduced by removing contaminated clothing.

Treatment is largely supportive – prevention and treatment of infection, haemorrhage and fluid loss. Harvesting of blood products is sometimes carried out.

Accidental ingestion of, or exposure to, bone-seeking radioisotopes (e.g. strontium-90 and caesium-137) are treated with chelating agents, e.g. EDTA and massive doses of oral calcium. Radio-iodine contamination should be treated

immediately with potassium iodide to block radioiodine absorption by the thyroid.

ELECTRIC SHOCK

Electric shock can produce:

- *Pain and psychological sequelae*. The common domestic electric shock is typically painful, rarely fatal or followed by serious sequelae. Nevertheless, it is an unpleasant and intensely frightening experience. A brief immediate jerking episode can occur, but not an epileptic seizure. There is typically no lasting neurological, cardiac or skin damage. More serious effects are distinctly rare following accidents in the home or in industry, but claims by survivors following industrial accidents are frequently made.
- *Cardiac, neurological and muscle damage*. Ventricular fibrillation, muscular contraction and spinal cord damage *can* follow a major shock. These are seen typically following lightning strikes with exceedingly high-voltage and amperage.
- *Electrical burns*. These are commonly restricted to the skin – non-fatal lightning strikes can cause fern-shaped burns. Deeper injuries can also occur, with muscle and spinal cord damage.
- *Electrocution*. This means death following ventricular fibrillation, either accidentally, or deliberately as a method of execution. In the US an initial voltage of >2000 volts was applied for some 15 seconds in the electric chair, causing loss of consciousness and ventricular fibrillation before the voltage was lowered. The T_{Core} during the execution process would sometimes increase to levels >50°C, leading to severe damage to internal organs.

SMOKE

Smoke consists of carbon particles in hot air and gases. The particles are coated with organic acids, aldehydes and synthetic materials. On combustion, other toxins such as carbon monoxide, sulphur dioxide, sulphuric and hydrochloric acids are released into smoke. The highly toxic polyvinyl chloride is no longer used in household goods. Air pollution is discussed on page 830.

Respiratory symptoms are typically immediate. Patients are dyspnoeic and tachypnoeic. Laryngeal stridor may require intubation. Hypoxia and pulmonary oedema can be fatal.

Patients should breathe through an aspirator or wet towel and be removed as rapidly as possible from the smoke. Give oxygen and arrange ITU support to monitor pulmonary damage.

Prevention: smoke alarms should be in every household.

NOISE

Sound *intensity* is expressed as the square of *sound pressure*. The *bel* is the ratio equivalent to a 10-fold increase in sound intensity; a *decibel* (dB) is one-tenth of a bel. Sound is made up of a number of frequencies ranging from 30 Hz to 20 kHz, with most being between 1 and 4 kHz. In practice a scale known as *A-weighted sound* is used; sound levels are reported as dB(A). A hazardous sound source is defined as one with an overall sound pressure of >90 dB(A).

Repeated prolonged exposure to loud noise, particularly between 2 and 6 kHz, causes first temporary and later permanent hearing loss, by physically destroying hair cells in the organ of Corti and, eventually, auditory neurones. This is a common occupational problem, not only in industry and the armed forces, but also in the home (drills and sanders), in sport (motor racing) and in entertainment (musicians, DJs and their audiences).

Serious noise-induced hearing loss is almost wholly preventable by personal protection (ear muffs, ear plugs). Little can be offered once hearing loss has become established.

Other effects of noise

Noise is intensely irritative, increasing or producing anxiety and anger. Excessive, repetitive noise is used in torture. Excess noise possibly affects child development and reading skills.

BIOTERRORISM/BIOWARFARE

Particular interest in biological warfare and bioterrorism intensified during the 1991 Iraq war and later following the destruction of the Twin Towers in New York in 2001. The potential of bacteria as weapons is illustrated by a suggestion that several kilograms of anthrax spores might kill as many people as a Hiroshima-sized nuclear weapon.

Potential pathogens

The US Center for Disease Control in Atlanta, Georgia, has developed a classification of potential biological agents (Table 17.4).

Smallpox

Smallpox is a highly infectious disease with a mortality >30%. There is no proven therapy, but there is an effective

Table 17.4 Critical biological agents

Category	Pathogens
A – Very infectious and/or readily disseminated organisms: high mortality with a major impact on public health	Smallpox, anthrax, botulinism, plague
B – Moderately easy to disseminate organisms causing moderate morbidity and mortality	Q fever, brucellosis, glanders, food-/water-borne pathogens, influenza
C – Emerging and possible genetically engineered pathogens	Viral haemorrhagic fevers, encephalitis viruses, drug-resistant TB

Adapted from Khan AS, Morse S, Lillibridge S. Public health preparedness for biological terrorism in the USA. *Lancet* 2000; **356**: 1179–1182, reprinted with permission from Elsevier.

Fig. 17.3 **Smallpox rash.**

vaccine. Universal vaccination was stopped in the early 1970s: the vast majority of the world's population is now unprotected against the *variola* virus (see p. 109). The potential exists for a world-wide epidemic of smallpox, possibly initiated by a bioterrorist act.

Smallpox has an incubation period of around 12 days, allowing any initial source of infection to go undetected until the rash (Fig. 17.3), similar to chicken pox, develops on the second or third day of the illness. Infection is transmitted by the airborne route; the patient becomes infectious to others 12–24 hours before the rash appears, thus allowing a potential infected volunteer to pass infection to others before being recognized as suffering from smallpox. If vaccines were to be administered widely to those potentially infected within 3 days of contact, an epidemic might well be prevented. Smallpox virus is stored in two secure laboratories – in Russia and in the USA. Supplies of vaccine are potentially available world-wide.

FURTHER READING

Buchler JW, Berkelman RL, Hartley DM. Syndromic surveillance and bioterrorism-related epidemics. *Emerging Infectious Diseases* 2003; **10:** 1197–1204.

Khan AS, Morse S, Lillibridge S. Public-health preparedness for biological terrorism in the USA. *Lancet* 2000; **356:** 1179–1182.

Anthrax (see p. 142)

In late 2001, anthrax organisms were sent through the US mail and infected 22 individuals. Eleven developed pulmonary anthrax, five of whom died; 11 suffered from cutaneous anthrax.

A simulated anthrax attack postulated release of anthrax powder from a truck passing a sports stadium with 74 000 spectators. 16 000 were estimated to become infected, with a death rate of 25%. In Russia, following accidental release of anthrax from a bioweapons factory, the death rate was substantial in those nearby and especially downwind of the factory.

Botulism

The toxin produced by *Clostridium botulinum* is one of the most potent poisons known to man (see p. 134).

As a bioweapon, botulinum toxin could be transmitted in food or by air, e.g. from a crop-spraying light aircraft. The toxin is inactivated by chlorine in domestic water supplies. There is no vaccine available.

Plague (see p. 146)

Plague could be transmitted as a bioweapon either by airborne dissemination or by infected rats. Immunization is of limited value.

Other potential infective agents are listed in Table 17.4.

Emergency planning

Many countries have plans to deal with bioterrorist attacks. These include training of healthcare staff and police. Such plans indicate the awareness of governments of the possibility of these threats. Stock piling of vaccines, antibiotics and protective clothing is essential.

FURTHER READING

Waterhouse J, Reilly T, Atkinson G. Jet-lag: trends and coping strategies. *Lancet* 2007; **369:** 1117–1129.

TRAVEL

Motion sickness

This common problem is caused by repetitive stimulation of the labyrinth. It occurs frequently at sea and in cars (especially in children), but also occurs with less usual forms of transport such as camels or elephants. Nowadays, motion sickness is rare during commercial flights, but it is a problem during space travel, and one reason why the airship industry has not flourished.

Nausea, sweating, dizziness, vertigo and profuse vomiting occur, accompanied by an irresistible desire either to stop moving or return to land. Prostration and intense incapacitating malaise can occur, for example in seasickness.

Prophylactic antihistamines, vestibular sedatives (hyoscine or cinnarizine) and stem ginger are of some value.

Jet-lag

Jet-lag (circadian dyschronism) is the well-known phenomenon that follows travelling through time-zones, particularly from West to East. Intense insomnia, fatigue, poor concentration, irritability, and loss of appetite are common. Headaches may occur. Symptoms last several days.

Mechanisms relate to the hypothalamic body clock within the suprachiasmatic nuclei. The clock is regulated by various *zeitgebers* (time-givers), e.g. light and melatonin.

Management of jet-lag includes its acceptance as a phenomenon causing poor performance – and thus waiting for 3–5 days to recover. Drink plenty of fluid, avoiding alcohol. Various hypnotics can help insomnia, but their value is disputed. Oral melatonin is widely used to reduce jet-lag. This probably hastens resetting of the body clock. Melatonin is not available on prescription in the UK.

BUILDING-RELATED ILLNESSES

Non-specific building-related illness

Multi-storey buildings typically have a controlled environment, often with automated heating and air-conditioning, and without ready access to external ventilation. More than half the adult workforce in developed countries work in such offices.

Headache, fatigue and difficulty in concentrating, sometimes in apparent epidemics, are the main complaints – and have become less frequent during the last decade. Psychological factors are very likely to have a substantial role. Temperature, humidity, dust, volatile organic compounds, e.g. paints and solvents, and even low level carbon monoxide toxicity have all been blamed, none with any scientific foundation. Maintenance of ready access to outdoor air is usually recommended, but changes in ventilation sometimes fail to improve matters.

Specific building-related illnesses

Legionnaires' disease (see p. 860) can certainly be caused by contamination of air-conditioning systems. *Humidifier fever* (see p. 877) is also due to contaminated systems, probably by fungi, bacteria and protozoa. Many common viruses are potentially transmissible in an enclosed environment, e.g. the common cold, influenza and rarely pulmonary TB. Allergic disorders, e.g. rhinitis, asthma and dermatitis, also occur following exposure to indoor allergens such as dust mites and plants. Office equipment, e.g. fumes from photocopiers, has also been implicated. Passive smoking (see p. 829) is no longer an issue in Europe and North America, following legislation against smoking.

Ask the authors
www.studentconsult.com

This page intentionally left blank

18 Endocrine disease

Introduction 963
Hormonal activity 965
Testing endocrine function 968
Disorders of pituitary and hypothalamus 969
Presentations of pituitary and hypothalamic disease 970
Growth and abnormal stature 977
Normal growth 977
Growth failure: short stature 978
Tall stature 979
Pituitary hypersecretion syndromes 979
Acromegaly and gigantism 979
Hyperprolactinaemia 980
Hypersecretion of other pituitary hormones 982
The thyroid axis 982
Hypothyroidism 985
Hyperthyroidism 987
Thyroid eye disease 991
Goitre (thyroid enlargement) 992
Thyroid carcinoma 993
Reproduction and sex 994
Clinical features of disorders of sex and reproduction 998
Disorders in the male 998
Disorders in the female 1001
Subfertility 1006
Disorders of sexual differentiation 1007
The glucocorticoid axis 1008
The thirst axis 1016
Disorders of calcium metabolism 1019
Endocrinology of blood pressure control 1022
Renin (and angiotensin) dependent hypertension 1023
Disorders of aldosterone secretion 1023
The adrenal medulla 1024
Other endocrine disorders 1026
Diseases of many glands 1026
Ectopic hormone secretion 1027

INTRODUCTION

Endocrine disease can involve all of the endocrine (or hormone) glands of the body as illustrated in Figure 18.1.

The most common endocrine disorders, excluding diabetes mellitus (Ch. 19) are:

- thyroid disorders, affecting 4–8 new patients per primary care physician each year (the most common problems are primary hypothyroidism, thyrotoxicosis and goitre)
- menstrual disorders and excessive hair growth in young women, most commonly polycystic ovary syndrome (PCOS)
- osteoporosis, especially in postmenopausal women largely owing to gonadal steroid deficiency
- primary hyperparathyroidism, affecting about 0.1% of the population
- subfertility, affecting 5–10% of all couples, often with an endocrine component
- disorders of growth or puberty.

While most other endocrine conditions are uncommon, they often affect young people and are usually curable or completely controllable with appropriate therapy.

Symptoms

Hormones produce widespread effects upon the body, and states of deficiency or excess typically present with symptoms that are generalized rather than focused on the anatomical location of the gland. Many endocrine symptoms are diffuse and vague, and the differential diagnosis is often wide. Symptoms of tiredness, weakness or lack of energy or drive and changes in appetite or thirst are common presentations of endocrine disease. Other typical 'hormonal' symptoms include changes in body size and shape, problems with libido and potency, periods or sexual development, and changes in the skin (dry, greasy, acne, bruising, thinning or thickening) and hair (loss or excess). Endocrine disorders should be always considered when assessing any patient with these common complaints.

History and examination

The past, family and social history is essential for making the diagnosis, planning appropriate management and interpreting results of borderline hormonal blood tests. The past history should include previous surgery or radiation involving endocrine glands, menstrual history, pregnancy and growth in childhood. A full drug history will exclude iatrogenic endocrine problems (Table 18.1). Family history of autoimmune disease, endocrine disease including tumours, diabetes and cardiovascular disease is frequently relevant, and knowledge of family members' height, weight, body habitus, hair growth and age of sexual development may aid interpretation of the patient's own symptoms.

Physical signs are listed under the relevant systems.

Aetiology of endocrine disease

Aetiological mechanisms common to many endocrine disorders include:

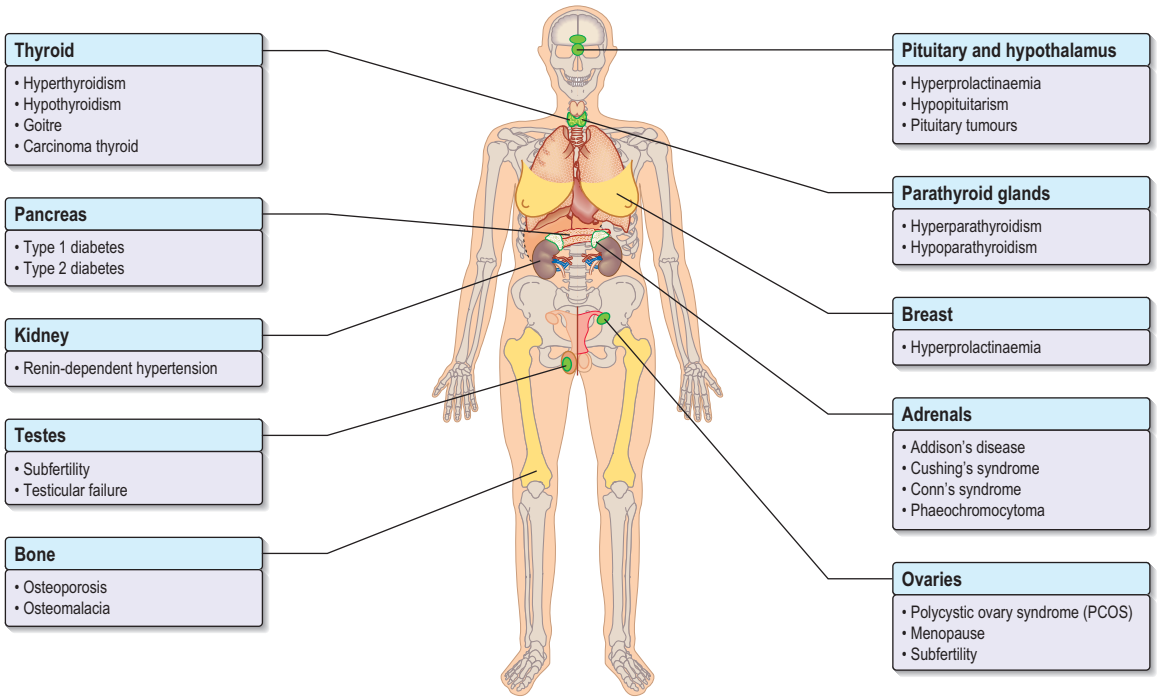


Fig. 18.1 **The major endocrine organs and common endocrine problems.**

Table 18.1 Drugs and endocrine disease

Drug*	Effect
Drugs inducing endocrine disease	
Chlorpromazine Metoclopramide (and all dopamine antagonists) Oestrogens	Increase prolactin, causing galactorrhoea
Iodine Amiodarone	Hyperthyroidism
Lithium Amiodarone	Hypothyroidism
Chlorpropamide	Inappropriate ADH secretion
Ketoconazole Metyrapone Aminoglutethimide	Hypoadrenalism
Chemotherapy	Ovarian and testicular failure
Drugs simulating endocrine disease	
Sympathomimetics Amfetamines	Mimic thyrotoxicosis or phaeochromocytoma
Liquorice Carbenoxolone	Increase mineralocorticoid activity; mimic aldosteronism
Purgatives	Hypokalaemia
Diuretics	Secondary aldosteronism
ACE inhibitors	Hypoaldosteronism
Drugs affecting hormone-binding proteins	
Anticonvulsants	Bind to TBG – decrease total T_4
Oestrogens	Raise TBG and CBG – increase total T_4/cortisol
Exogenous hormones or stimulating agents	
Use, abuse or misuse, by patient or doctor, of the following:	
Steroids	Cushing's syndrome Diabetes
Thyroxine	Thyrotoxicosis factitia
Vitamin D preparations Milk and alkali preparations	Hypercalcaemia
Insulin Sulfonylureas	Hypoglycaemia

*Drugs causing gynaecomastia are listed in Table 18.20.
Amiodarone may cause both hypo- or hyperthyroidism.
TBG, thyroxine-binding globulin; CBG, cortisol-binding globulin.

Table 18.2 Types of autoimmune disease affecting endocrine organs

Organ and frequency if known	Antibody	Antigen if known	Clinical syndrome
Stimulating			
Thyroid (1 in 100)	Thyroid-stimulating immunoglobulin (TSI, TSAb)	TSH receptor	Graves' disease, neonatal thyrotoxicosis
Destructive			
Thyroid (1 in 100)	Thyroid microsomal Thyroglobulin	Thyroid peroxidase enzyme (TPO)	Primary hypothyroidism (myxoedema)
Adrenal (1 in 20 000)	Adrenal cortex	21-Hydroxylase enzyme	Primary hypoadrenalism (Addison's disease)
Pancreas (1 in 500)	Islet cell	GAD (see p. 1033)	Type 1 diabetes
Stomach	Gastric parietal cell Intrinsic factor	Gastric parietal cell Intrinsic factor	Pernicious anaemia
Skin	Melanocyte	Melanocyte	Vitiligo
Ovary (1 in 500)	Ovary		Primary ovarian failure
Testis	Testis		Primary testicular failure
Parathyroid	Parathyroid chief cell	Parathyroid chief cell	Primary hypoparathyroidism
Pituitary	Pituitary-specific cells		Selective hypopituitarism (e.g. GH deficiency, diabetes insipidus)

Frequencies are approximate and refer to the population in Northern Europe.
GAD, glutamic acid dehydrogenase.
NB: Other related diseases include myasthenia gravis and autoimmune liver diseases.

Autoimmune disease

Organ-specific autoimmune diseases can affect every major endocrine organ (Table 18.2). They are characterized by the presence of specific antibodies in the serum, often present years before clinical symptoms are evident. The conditions are usually more common in women and have a strong genetic component, often with an identical-twin concordance rate of 50% and with HLA associations (see individual diseases). Several of the autoantigens have been identified.

Endocrine tumours

Hormone-secreting tumours occur in all endocrine organs, most commonly pituitary, thyroid and parathyroid. Fortunately, they are more commonly benign than malignant. While often considered to be 'autonomous' – that is, independent of the physiological control mechanisms – many do show evidence of feedback occurring at a higher 'set-point' than normal (e.g. ACTH secretion from a pituitary basophil adenoma). Non-functioning benign tumours of endocrine organs are even more common and often present as 'incidentalomas' found incidentally during imaging for another condition.

The molecular basis of some of these tumours is sometimes a very specific mutation of a single gene, such as the mutations of the *RET* proto-oncogene in MEN 2 (see p. 1026), but more commonly a wide variety of different mutations in tumour suppressor genes, growth factor receptors and other intracellular mediators have been identified.

Enzymatic defects

The biosynthesis of most hormones involves many stages. Deficient or abnormal enzymes can lead to absent or reduced production of the secreted hormone. In general, severe deficiencies present early in life with obvious signs; partial deficiencies usually present later with mild signs or are only evident under stress. An example of an enzyme deficiency is congenital adrenal hyperplasia (CAH) where the molecular basis has also been identified as mutations or deletions of the gene encoding the relevant enzymes.

Receptor abnormalities (see p. 971)

Hormones work by activating cellular receptors. There are rare conditions in which hormone secretion and control are normal but the receptors are defective; thus, if androgen receptors are defective, normal levels of androgen will not produce masculinization (e.g. testicular feminization). There are also a number of rare syndromes of diabetes and insulin resistance from receptor abnormalities (p. 1035); other examples include nephrogenic diabetes insipidus, thyroid hormone resistance and pseudohypoparathyroidism.

Hormones as therapy

Hormones are also widely used therapeutically:

- The oral contraceptive pill is the choice of perhaps 20–30% of women aged 18–35 years using contraception.
- Corticosteroid therapy is widely used in non-endocrine disease such as asthma (see p. 855).
- Hormone replacement therapy (HRT; oestrogens ± progestogens) is used for control of menopausal symptoms in postmenopausal women (see p. 997).

HORMONAL ACTIVITY

Hormones are chemical messengers produced by a variety of specialized secretory cells. They are transported in the blood to a distant site of action ('endocrine' activity) and sometimes act directly upon nearby cells ('paracrine' activity). Classic hormones are secreted by the main endocrine glands – pituitary, thyroid, adrenals, gonads, parathyroids and pancreas – and have effects elsewhere in the body. In addition, there are many other cells that secrete hormones, e.g. in the hypothalamus and elsewhere in the brain and in the gastrointestinal tract. These can have true endocrine or paracrine activity or behave more like neurotransmitters or neuromodulators. At the molecular level there is little difference in the way cellular activity is regulated between classical neurotransmitters that act across synaptic clefts,

intercellular factors acting across gap junctions, classic endocrine and paracrine activity and a variety of other chemical messengers involved in cell regulation – such as cytokines, growth factors and interleukins. Progress in basic cell biology has revealed the biochemical similarities in the messengers, receptors and intracellular post-receptor mechanisms underlying all these aspects of cell function.

Synthesis, storage and release of hormones

Hormones may be of several chemical structures: polypeptide, glycoprotein, steroid or amine. Hormone release is the end-product of a long cascade of intracellular events. In the case of polypeptide hormones, neural or endocrine stimulation of the cell leads to increased transcription from DNA to a specific mRNA, which is in turn translated to the peptide product. This is often in the form of a precursor molecule that may itself be biologically inactive. This 'prohormone' is then further processed before being packaged into granules, in the Golgi apparatus. These granules are then transported to the plasma membrane before release, which is itself regulated by a complex combination of intracellular regulators. Hormone release may be in a brief spurt caused by the sudden stimulation of granules, often induced by an intracellular Ca^{2+}-dependent process, or it is 'constitutive' (immediate and continuous secretion).

Plasma transport

Most classical hormones are secreted into the systemic circulation. In contrast, hypothalamic releasing hormones are released into the pituitary portal system so that much higher concentrations of the releasing hormones reach the pituitary than occur in the systemic circulation.

Many hormones are bound to proteins within the circulation. In most cases, only the free (unbound) hormone is available to the tissues and thus biologically active. This binding serves to buffer against very rapid changes in plasma levels of the hormone, and some binding protein interactions are also involved in the active regulation of hormone action. Many tests of endocrine function measure total rather than free hormone, which can give rise to difficulties in interpretation when binding proteins are altered in disease states or by drugs. Binding proteins comprise both specific, high-affinity proteins of limited capacity, such as thyroxine-binding globulin (TBG), cortisol-binding globulin (CBG), sex-hormone-binding globulin (SHBG) and IGF-binding proteins (e.g. IGF-BP3) and other less specific, low-affinity ones, such as prealbumin and albumin.

Hormone action and receptors

Hormones act by binding to specific receptors in the target cell. Most hormone receptors are proteins with complex tertiary structures. The structure of the hormone-binding domain of the receptor complements the tertiary structure of the hormone, while changes in other parts of the receptor in response to hormone binding are responsible for the effects of the activated receptor within the cell.

Hormone receptors may be broadly divided into:

- *cell surface or membrane receptors* – typically transmembrane receptors which contain hydrophobic sections spanning the lipid-rich plasma membrane and which trigger internal cellular messengers (see also p. 22)
- *nuclear receptors* which typically bind hormones and translocate them to the nucleus where they bind hormone response elements of nuclear DNA via characteristic amino-acid sequences (e.g. so-called 'zinc fingers', see p. 32).

Common structural mechanisms of hormone–receptor action are illustrated in Figure 2.4 (p. 23) and include:

G-protein coupled receptors (7-transmembrane or serpentine receptors). These bind hormones on their extracellular domain and activate the membrane G-protein complex with their intracellular domain. The activated complex may then:

- stimulate cyclic AMP (cAMP) generation by adenylate cyclase – activating further intracellular kinases and leading to phosphorylation
- activate phospholipase C (PLC) leading to generation of inositol 1,4,5-triphosphate (IP3) and release of intracellular calcium – in turn leading to calmodulin-dependent kinase activity and phosphorylation.
- lead to diacylglycerol (DAG) activation of C-kinase and subsequent protein phosphorylation.

Most peptide hormones act via G-protein coupled receptors.

Dimeric transmembrane receptors, from several receptor superfamilies, bind hormone in their extracellular components (sometimes causing the dimerization of the receptor monomer) and directly phosphorylate intracellular messengers via their intracellular components, leading to a variety of intracellular activation cascades. Growth hormone, prolactin and insulin-like growth factor-1 (IGF-1) act via this type of receptor.

Lipid-soluble molecules pass through the cell membrane and typically bind with their **nuclear receptors** in the cell cytoplasm before translocation of the activated hormone–receptor complex to the nucleus where it binds to nuclear DNA, often in combination with a multi-component complex of promoters, inhibitors and transcription factors. This interaction usually leads to increased transcription of the relevant gene product. Steroid and thyroid hormones act via this type of receptor.

The structure of common hormones and their receptors is described under individual hormone axes.

The activation of intracellular kinases, phosphorylation, release of intracellular calcium and other 'second messenger' pathways and the direct stimulation of DNA transcription results in some or all of:

- stimulation or release of pre-formed hormone from storage granules
- stimulation or synthesis of hormone and other cellular components
- opening or closing of ion or water channels in the cell membrane (e.g. calcium channels or aquaporin water channels)
- activation or deactivation of other DNA binding proteins leading to stimulation or inhibition of DNA transcription.

In each case, binding of the hormone to its receptor is the first step in a complex cascade of interrelated intracellular events which eventually lead to the overall effects of that hormone on cellular function.

The sensitivity and/or number of receptors for a hormone are often decreased after prolonged exposure to a high hormone concentration, the receptors thus becoming less sensitive ('downregulation', e.g. angiotensin II receptors, β-

adrenoceptors). The reverse is true when stimulation is absent or minimal, the receptors showing increased numbers or sensitivity ('upregulation').

Abnormal receptors are an occasional cause of endocrine disease (see p. 971), recognized and characterized more frequently owing to advances in molecular endocrinology.

Control and feedback

Most hormone systems are controlled by some form of feedback; an example is the hypothalamic–pituitary–thyroid axis (Fig. 18.2).

- TRH (thyrotrophin-releasing hormone) is secreted in the hypothalamus and travels via the portal system to the pituitary where it stimulates the thyrotrophs to produce thyroid-stimulating hormone (TSH).
- TSH is secreted into the systemic circulation where it stimulates increased thyroidal iodine uptake by the thyroid and the synthesis and release of thyroxine (T_4) and triiodothyronine (T_3).
- Serum levels of T_3 and T_4 are thus increased by TSH; in addition, the conversion of T_4 to T_3 (the more active hormone) in peripheral tissues is stimulated by TSH.
- T_3 and T_4 then enter cells where they bind to nuclear receptors and promote increased metabolic and cellular activity.
- Levels of T_3 (from the blood and from local conversion of T_4) are sensed by receptors in the pituitary and possibly the hypothalamus. If they rise above normal, TRH and TSH production is suppressed, leading to reduced T_3 and T_4 secretion.
- Peripheral T_3 and T_4 levels thus fall to normal.
- If, however, T_3 and T_4 levels are low (e.g. after thyroidectomy), increased amounts of TRH and thus TSH are secreted, stimulating the remaining thyroid to produce more T_3 and T_4; blood levels of T_3 and T_4 may be restored to normal, although at the expense of increased TSH drive, reflected by a high TSH level ('compensated euthyroidism'). Conversely, in thyrotoxicosis when factors other than TSH itself are maintaining high T_3 and T_4 levels, the same mechanisms lead to suppression of TSH secretion.

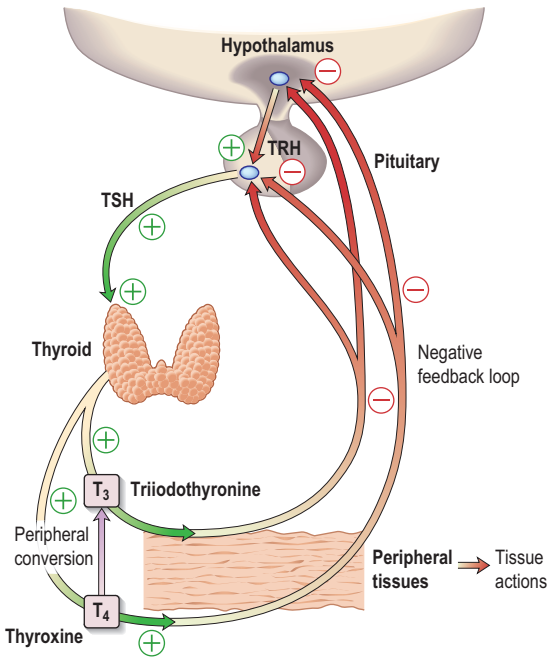


Fig. 18.2 **The hypothalamic–pituitary–thyroid axis.** Pituitary TSH is secreted in response to hypothalamic TRH and stimulates secretion of T_4 and T_3 from the thyroid. T_4 and T_3 have actions in peripheral tissues and exert negative feedback on pituitary and hypothalamus.

This is known as a 'negative feedback' system, referring to the effect of T_3 and T_4 on the pituitary and hypothalamus, which represents the most common mechanism for regulation of circulating hormone levels. There are also 'positive feedback' systems, classically seen in the regulation of the normal menstrual cycle (p. 996).

Patterns of secretion

Hormone secretion is continuous or intermittent. The former is shown by the thyroid hormones, where T_4 has a half-life of 7–10 days and T_3 of about 6–10 hours. Levels over the day, month and year show little variation.

In contrast, secretion of the gonadotrophins, LH and FSH, is normally pulsatile, with major pulses released every 1–2 hours depending on the phase of the menstrual cycle. Continuous infusion of LH to produce a steady equivalent level does not produce the same result (e.g. ovulation in the female) as the intermittent pulsatility, and may indeed produce downregulation. Thus a long-acting superactive gonadotrophin-releasing hormone (GnRH) analogue, such as goserelin, produces downregulation of the GnRH receptors and subsequent very low androgen or oestrogen levels, which is clinically valuable both in carcinoma of the prostate in men and in ovulation-induction regimens in infertile women. In contrast, pulsatile GnRH administration can produce normal menstrual cyclicity, ovulation and fertility in women with hypothalamic amenorrhoea but intact pituitary LH and FSH stores.

Growth hormone is also secreted in a pulsatile fashion, with undetectable levels in between pulses. A single measurement is therefore not helpful to diagnose GH deficiency or excess.

Biological rhythms

Circadian means changes over the 24 hours of the day–night cycle and is best shown for the pituitary–adrenal axis. Figure 18.3 shows plasma cortisol levels measured over 24 hours – levels are highest in the early morning and lowest overnight. Additionally, cortisol release is pulsatile, following the pulsatility of pituitary ACTH. Thus 'normal' cortisol levels vary during the day and great variations can be seen in samples taken only 30 minutes apart. The circadian (light–dark) rhythm is seen in reverse with the pineal hormone, melatonin, which shows high levels during darkness. Melatonin may be involved in entraining other hormonal rhythms and systems to the current light–dark cycle, but there is no known clinical syndrome related to abnormalities of this hormone, though a synthetic preparation has been used for 'jet-lag' (p. 960).

The menstrual cycle is the best example of a longer and more complex (28-day) biological rhythm (see p. 996).

Other regulatory factors

- *Stress*. Physiological 'stress' and acute illness produce rapid increases in ACTH and cortisol, growth hormone (GH), prolactin, epinephrine (adrenaline) and norepinephrine (noradrenaline). These can occur within seconds or minutes.

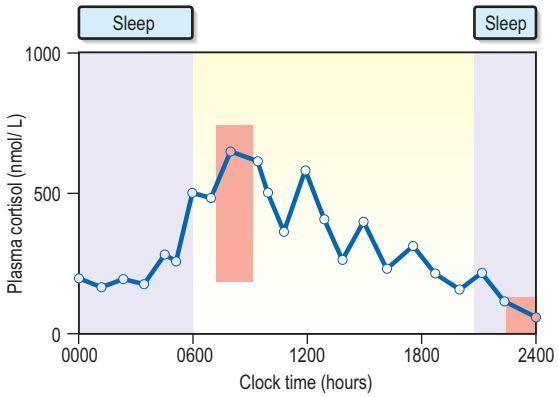


Fig. 18.3 **Plasma cortisol levels during a 24-hour period.** Note both the pulsatility and the shifting baseline. Normal ranges for 0900 h (180–700 nmol/L) and 2400 h (less than 100 nmol/L; must be taken when asleep) are shown in the orange boxes. Purple shading shows sleep.

- *Sleep*. Secretion of GH and prolactin is increased during sleep, especially the rapid eye movement (REM) phase.
- *Feeding and fasting*. Many hormones regulate the body's control of energy intake and expenditure and are therefore profoundly influenced by feeding and fasting. Secretion of insulin is increased and growth hormone decreased after ingestion of food, and secretion of a number of hormones is altered during prolonged food deprivation.

Measurement of hormone levels in normal individuals and in patients with disease must take these factors into account. For example, cortisol levels will often be high and fail to suppress during standard tests in a patient who is severely stressed by serious illness, and growth hormone will usually be low in postprandial individuals during the daytime.

TESTING ENDOCRINE FUNCTION

Endocrine function is assessed by measurement of hormone levels in blood (or more precisely in plasma or serum) and sometimes in other body fluids on samples obtained basally and in response to stimulation and suppression tests.

Basal blood levels

Assays for all clinically relevant pituitary and end-organ hormones are available. Hypothalamic hormones are not measured because of their low concentration and local action within the hypothalamo–pituitary axis.

The time, day and condition of measurement make great differences to hormone levels, and the method and timing of samples therefore depends upon the characteristics of the endocrine system involved. There are also sex, developmental and age differences.

Basal levels are especially useful for systems with long half-lives (e.g. T_4 and T_3). These vary little over the short term and random samples are therefore satisfactory.

Basal samples for other hormones may also be satisfactory if interpreted with respect to normal ranges for the time of day/month, diet or posture concerned. Examples are FSH, oestrogen and progesterone (varying with time of cycle) and renin/aldosterone (varying with sodium intake, posture and age). For these hormones, all relevant details must be recorded or the results may prove uninterpretable.

Stress-related hormones

Stress-related hormones (e.g. catecholamines, prolactin, GH, ACTH and cortisol) require samples to be taken via an indwelling needle some time after initial venepuncture; otherwise, high levels may be artefactual.

Urine collections

Collections over 24 hours have the advantage of providing an 'integrated mean' of a day's secretion but in practice are often incomplete or wrongly timed. They also vary with sex and body size or age. Written instructions should be provided for the patient to ensure accurate collection.

Saliva

Saliva is sometimes used for steroid estimations, especially in children or for samples taken at home. Midnight salivary cortisol levels are increasingly used for the diagnosis of Cushing's syndrome due to the practical difficulties in obtaining a midnight blood sample.

Stimulation and suppression tests

These tests are used when basal levels give equivocal information. In general, stimulation tests are used to confirm suspected deficiency, and suppression tests to confirm suspected excess of hormone secretion. These tests are valuable in many instances.

For example, where the secretory capacity of a gland is damaged, maximal stimulation by the trophic hormone will give a diminished output. Thus, in the short ACTH stimulation test for adrenal reserve (Box 18.1 and Figure 18.4a), the healthy subject shows a normal response while the subject with primary hypoadrenalism (Addison's disease) demonstrates an impaired cortisol response to tetracosactide (an ACTH analogue).

A patient with a hormone-producing tumour usually fails to show normal negative feedback. A patient with Cushing's disease (excess pituitary ACTH) will thus fail to suppress ACTH and cortisol production when given a dose of synthetic steroid, in contrast to normal subjects. Figure 18.4b shows the response of a normal subject given dexamethasone 1 mg at midnight; cortisol is suppressed the following morning. The subject with Cushing's disease shows inadequate suppression.

The detailed protocol for each test must be followed exactly, since differences in technique will produce variations in results.

Box 18.1 Short ACTH (tetracosactide) stimulation test

Indication

Diagnosis of Addison's disease
Screening test for ACTH deficiency

Procedure

Intravenous cannula for sampling
Any time of day, but best at 0900 h; non-fasting
Tetracosactide 250 µg, i.v. or i.m. at time 0
Measure serum cortisol at time 0 and time +30 min

Normal response

30 min cortisol >600 nmol/L*
(400–600 nmol/L borderline and may indicate deficiency)

*Precise cortisol normal ranges are variable between laboratories and assays – appropriate local reference ranges must be used.

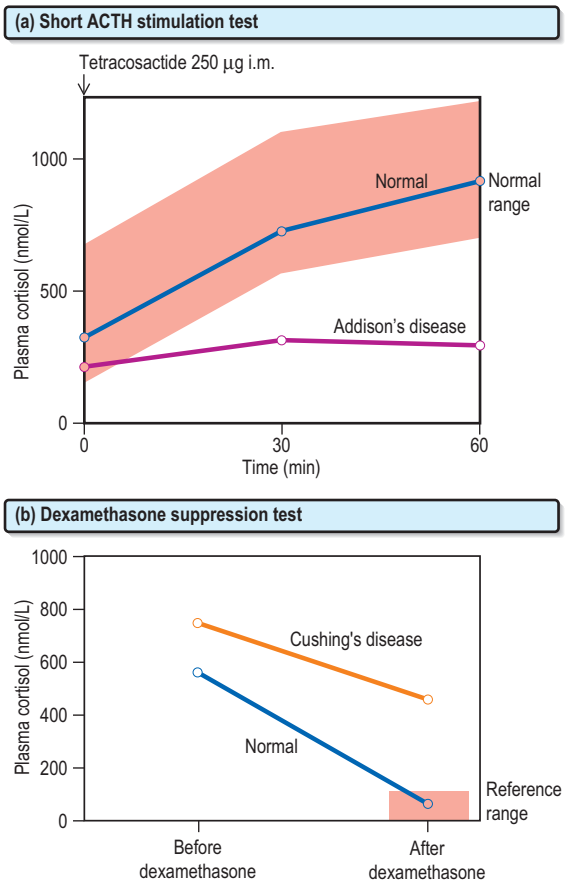


Fig. 18.4 **Short ACTH stimulation and dexamethasone tests. (a)** This test shows a normal response in a healthy subject and a decreased response in a patient with Addison's disease. **(b)** Dexamethasone suppression tests in a normal subject and in a patient with Cushing's disease showing inadequate suppression.

Measurement of hormone concentrations

Circulating levels of most hormones are very low (10^{-9}–10^{-12} mol/L) and cannot be measured by simple chemical techniques. Hormones are therefore usually measured by immunoassays which rely on highly specific antibodies (polyclonal or more usually monoclonal) which bind specifically to the hormone being measured during the assay incubation. This hormone–antibody interaction is measured by use of labelled hormone after separation of bound and free fractions (Fig. 18.5).

Immunoassay is sensitive but has limitations. In particular, the immunological activity of a hormone, as used in developing the antibody, may not necessarily correspond to biological activity. Other measurement techniques include high-pressure liquid chromatography (HPLC).

DISORDERS OF PITUITARY AND HYPOTHALAMUS

Anatomy

Most peripheral hormone systems are controlled by the hypothalamus and pituitary. The hypothalamus is sited at the base of the brain around the third ventricle and above the pituitary stalk, which leads down to the pituitary itself, carrying the hypophyseal–pituitary portal blood supply.

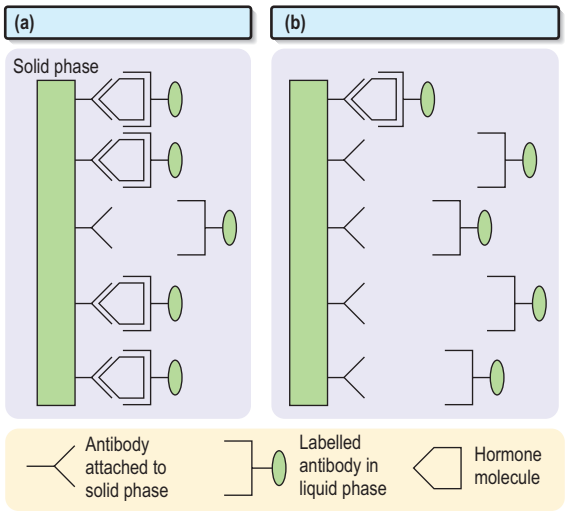


Fig. 18.5 **Principles of measurement of hormone levels in plasma by immunoassay** (precise details vary with different assays and manufacturers). Immunoassays use two antibodies specific to the hormone being measured – one typically attached to a solid phase and one labelled antibody in the liquid phase. **(a)** High hormone levels in plasma: large amount of hormone binds to antibody on solid phase – large amount of labelled antibody linked to solid phase via molecules of the hormone. **(b)** Low hormone levels in plasma: less hormone, and therefore less labelled antibody, is linked to the solid phase. Label (radioactive, chemiluminescent, enzymatic or fluorescent) can be measured in either solid or liquid phase after separation of phases; levels of label will be proportional to the amount of hormone in the sample.

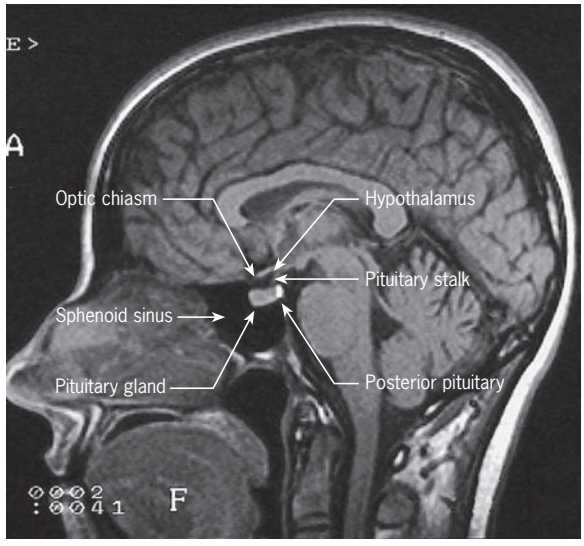


Fig. 18.6 **MR image of a sagittal section of the brain, showing the pituitary fossa and adjacent structures.** By kind permission of Dr Martin Jeffree.

The anatomical relations of the hypothalamus and pituitary (Fig. 18.6) include the optic chiasm just above the pituitary fossa; any expanding lesion from the pituitary or hypothalamus can thus produce visual field defects by pressure on the chiasm. Such upward expansion of the gland through the diaphragma sellae is termed 'suprasellar extension'. Lateral extension of pituitary lesions may involve the vascular and nervous structures in the cavernous sinus and

may rarely reach the temporal lobe of the brain. The pituitary is itself encased in a bony box; any lateral, anterior or posterior expansion must cause bony erosion.

Embryologically, the anterior pituitary is formed from an upgrowth of Rathke's pouch (ectodermal) which meets an outpouching of the third ventricular floor which becomes the posterior pituitary. This unique combination of primitive gut and neural tissue provides an essential link between the rapidly responsive central nervous system and the longer-acting endocrine system. Several transcription factors – LHX3, HESX1, PROP1, POUIF1 – are responsible for the differentiation and development of the pituitary cells. Mutation of these produces pituitary disease.

Physiology

Hypothalamus

This contains many vital centres for such functions as appetite, thirst, thermal regulation and sleeping/waking. It acts as an integrator of many neural and endocrine inputs to control the release of pituitary hormone-releasing factors. It plays a role in the circadian rhythm, menstrual cyclicity, and responses to stress, exercise and mood.

Hypothalamic neurones secrete pituitary hormone-releasing and -inhibiting factors and hormones (Table 18.3) into the portal system which runs down the stalk to the pituitary. As well as the classical hormones illustrated in Figure 18.7, the hypothalamus also contains large amounts of other neuropeptides and neurotransmitters such as neuropeptide Y, vasoactive intestinal peptide (VIP) and nitric oxide that can also alter pituitary hormone secretion.

Synthetic hypothalamic hormones and their antagonists are available for the testing of many aspects of endocrine function and for treatment.

Anterior pituitary

Hormone secretion is controlled by hypothalamic releasing or inhibitory hormones (Fig. 18.7). Many hormones are under dual control by both stimulatory and inhibitory hypothalamic factors. Examples are:

- Growth hormone release is stimulated by growth hormone-releasing hormone (GHRH) but inhibited by somatostatin (growth hormone release inhibitory hormone, GHRIH).
- TSH release is stimulated by TRH but partially inhibited by somatostatin.

Some hormones have a dual stimulatory control. For example, corticotrophin-releasing hormone (CRH) and vasopressin are both endogenous stimulators of ACTH release. Uniquely, prolactin is under predominant inhibitory dopaminergic control with some stimulatory TRH control. This means that disruption of the pituitary stalk leads to an elevation in serum prolactin but reduction in all other pituitary hormones.

Posterior pituitary

This, in contrast, acts merely as a storage organ. Antidiuretic hormone (ADH, also called vasopressin) and oxytocin, both nonapeptides, are synthesized in the supraoptic and paraventricular nuclei in the anterior hypothalamus. They are then transported along the axon and stored in the posterior pituitary. This means that damage to the stalk or pituitary alone does not prevent synthesis and release of ADH and oxytocin. ADH is discussed on page 1016; oxytocin produces milk ejection and uterine myometrial contraction.

PRESENTATIONS OF PITUITARY AND HYPOTHALAMIC DISEASE

Diseases of the pituitary can cause under- or over-activity of each of the hypothalamo–pituitary–end-organ axes which are under the control of this gland. The clinical features of the syndromes associated with such altered pituitary function, e.g. Cushing's syndrome, can be the presenting symptom of

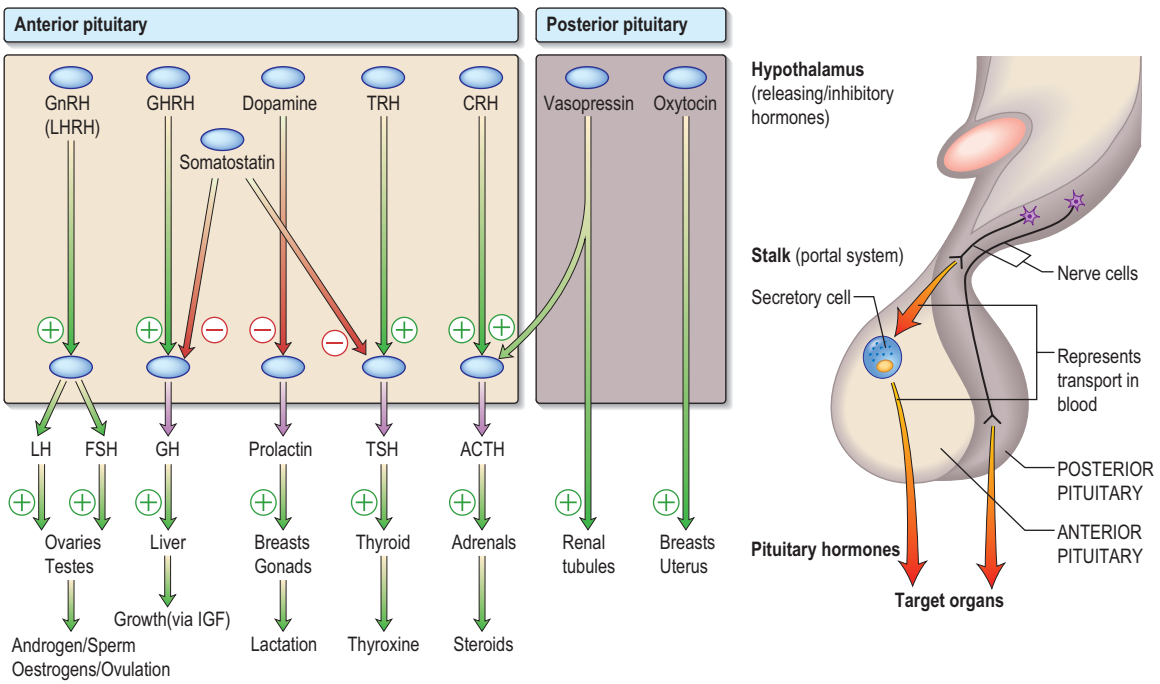


Fig. 18.7 **Hypothalamic releasing hormones and the pituitary trophic hormones.** See the text for abbreviations and an explanation. Activin, a peptide, is produced in the pituitary to stimulate FSH synthesis and is also produced in the ovary to facilitate the response of ovarian granuloma cells to FSH. It counteracts the effect of inhibin.

Table 18.3 Hormones and receptors of the hypothalamic–pituitary axis

Hormone	Source	Site of action	Hormone structure	Receptor	Post-receptor activity
Pituitary growth axis					
Growth hormone-releasing hormone (GHRH)	Hypothalamus	Pituitary	Peptide – 44 AA	Membrane – 7TM	G-proteins cAMP
Somatostatin (inhibitory GHRIH)	Hypothalamus	Pituitary	Cyclic peptide – 14 or 28 AA	Membrane – 7TM SST2 + SST5	G-proteins Inhibit cAMP
Growth hormone	Pituitary	Liver and other tissues	Peptide – 191 AA	Transmembrane Dimerized GHR	JAK2 STAT
Insulin-like growth factor 1 (IGF-1)	Liver + locally elsewhere	Many tissues	Peptide – 70 AA 3 S=S bridges	Transmembrane – IGFR 2α+ 2β subunits	Receptor tyrosine kinase
Pituitary–thyroid axis					
Thyrotrophin-releasing hormone (TRH)	Hypothalamus	Pituitary	Peptide – 3 AA	Membrane – 7TM TRHR	G-proteins PLC / IP_3
Thyroid stimulating hormone (TSH)	Pituitary	Thyroid	Glycoprotein – α and β subunits	Membrane – 7TM TSH-R	G-proteins cAMP / PLC/IP_3
Thyroxine and triiiodothyronine (T_4 and T_3)	Thyroid	All tissues	Thyronines – 4/3 iodine atoms	Nuclear TR- α and β	Transcription TRE/RXR
Pituitary–gonadal axis					
Gonadotrophin-releasing hormone (GnRH; LHRH)	Hypothalamus	Pituitary	Peptide – 10 AA	Membrane – 7TM GnRH-R1	G-proteins PLC / IP_3
Luteinizing hormone (LH)	Pituitary	Gonad	Glycoprotein – α and β subunits	Membrane – 7TM LHCGR	G-proteins cAMP
Follicle stimulating hormone (FSH)	Pituitary	Gonad	Glycoprotein – α and β subunits	Membrane – 7TM FSHR	G-proteins cAMP
Oestradiol	Ovary	Uterus, breast bone, vascular	Steroid ring	Nuclear ER α and β (ESR1/ESR2)	Homo-/hetero-Dimer ERE Transcription
Testosterone	Testis	Many tissues	Steroid ring	Nuclear AR (NR3C4)	Dimer ARE Transcription
Inhibin and activin	Gonad	Pituitary – hypothalamus	Peptide dimers α and β subunits	Transmembrane dimerized	Phosphorylation by receptor
Prolactin axis					
Dopamine	Hypothalamus	Pituitary	Amine	Membrane – 7TM D2 receptor	G-proteins Inhibit cAMP
Prolactin	Pituitary	Breast Other tissues	Peptide – 199 AA	Transmembrane PRLR Class 1 cytokine	JAK2 and other pathways
Pituitary–adrenal axis					
Corticotrophin releasing hormone (CRH)	Hypothalamus	Pituitary	Peptide – 41 AA	Membrane – 7TM CRF1	G-proteins cAMP
Adrenocorticotrophic hormone (ACTH)	Pituitary	Adrenal	Peptide – 39 AA	Membrane – 7TM ACTHR (MCR2)	G-proteins cAMP
Cortisol	Adrenal	All tissues	Steroid ring	Nuclear GRα	Transcription GRE
Vasopressin axis					
Vasopressin	Hypothalamus → pituitary	Kidney	Peptide – 9 AA	Membrane – 7TM AVPR2	G-proteins cAMP Aquaporin-2
	Hypothalamus → pituitary	Vascular	Peptide – 9 AA	Membrane – 7TM AVRP1A	G-proteins PLC/IP_3
	Hypothalamus → portal veins	Pituitary (ACTH secretion)	Peptide – 9 AA	Membrane – 7TM AVPR1B	G-proteins PLC/IP_3

7TM, 7 transmembrane (G-protein coupled receptor); AA, amino acids; ARE, androgen response element; cAMP, adenylate cyclase → cyclic AMP; ERE, oestrogen response element, GRE, glucocorticoid response element; JAK, Janus kinase; MCR, melanocortin receptor; PLC/IP_3, phospholipase C/inositol triphosphate; STAT, signal transducers and activators of transcriptions; GHRIH, growth hormone releasing inhibitory hormone; SST, somatostatin receptor subtypes.

Table 18.4 Characteristics of common pituitary and related tumours

Tumour or condition	Usual size	Most common clinical presentation
Prolactinoma	Most <10 mm (microprolactinoma)	Galactorrhoea, amenorrhoea, hypogonadism, erectile dysfunction
	Some >10 mm (macroprolactinoma)	As above plus headaches, visual field defects and hypopituitarism
Acromegaly	Few mm to several cm	Change in appearance, visual field defects and hypopituitarism
Cushing's disease	Most small – few mm (some cases are hyperplasia)	Central obesity, cushingoid appearance (local symptoms rare)
Nelson's syndrome	Often large – >10 mm	Post-adrenalectomy, pigmentation, sometimes local symptoms
Non-functioning tumours	Usually large – >10 mm	Visual field defects; hypopituitarism (microadenomas may be incidental finding)
Craniopharyngioma	Often very large and cystic (skull X-ray abnormal in >50%; calcification common)	Headaches, visual field defects, growth failure (50% occur below age 20; about 15% arise from within sella)

pituitary disease or of end-organ disease and are discussed later. First, however, we consider clinical features of pituitary disease which are common to all hormonal axes.

Pituitary space-occupying lesions and tumours

Pituitary tumours (Table 18.4) are the most common cause of pituitary disease, and the great majority of these are benign pituitary adenomas, usually monoclonal in origin. Problems are caused by:

- excess hormone secretion
- local effects of a tumour
- the result of inadequate production of hormone by the remaining normal pituitary, i.e. hypopituitarism.

Investigations (of a possible or proven mass)

Is there a tumour?

If there is, how big is it and what *local anatomical effects* is it exerting? Pituitary and hypothalamic space-occupying lesions, hormonally active or not, can cause symptoms by pressure on, or infiltration of:

- the visual pathways, with field defects and visual loss (most common)
- the cavernous sinus, with III, IV and VI cranial nerve lesions
- bony structures and the meninges surrounding the fossa, causing headache
- hypothalamic centres: altered appetite, obesity, thirst, somnolence/wakefulness or precocious puberty
- the ventricles, causing interruption of cerebrospinal fluid (CSF) flow leading to hydrocephalus
- the sphenoid sinus with invasion causing CSF rhinorrhoea.

Investigations:

- **MRI of the pituitary**. MRI is superior to CT scanning (Fig. 18.8) and will readily show any significant pituitary mass. Small lesions within the pituitary fossa on MRI consistent with small pituitary microadenomas are very common (10% of normal individuals in some studies). Such small lesions are sometimes detected during MRI scanning of the head for other reasons – so-called 'pituitary incidentalomas'.
- **Visual fields**. These should be plotted formally by automated computer perimetry or Goldmann perimetry, but clinical assessment by confrontation using a small red pin as target is also sensitive and valuable. Common defects are upper temporal quadrantanopia and bitemporal hemianopia (see p. 1101). Subtle defects may also be revealed by delay or attenuation of visual evoked potentials (VEPs).

Is there a hormonal excess?

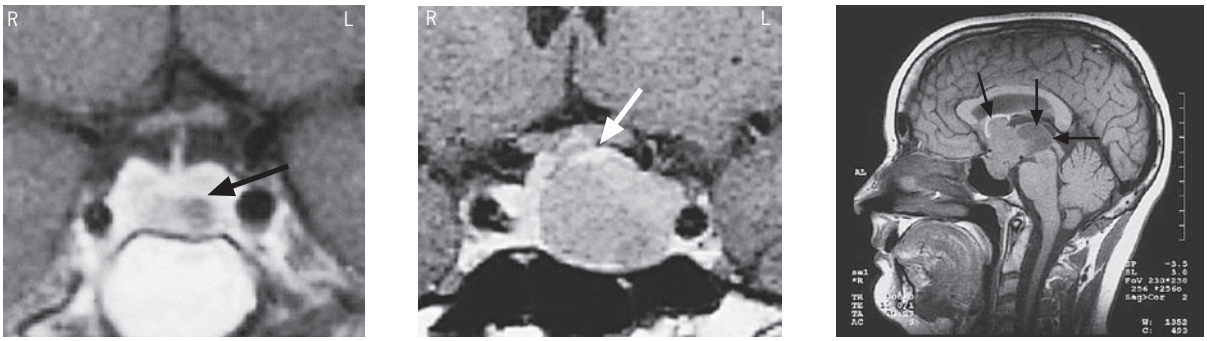
There are three major conditions usually caused by secretion from pituitary adenomas which will show positive immunostaining for the relevant hormone:

- prolactin excess (**prolactinoma or hyperprolactinaemia**) – histologically, prolactinomas are 'chromophobe' adenomas (a description of their appearance on classical histological staining)
- GH excess (**acromegaly or gigantism**) – somatotroph adenomas, usually 'acidophil', and sometimes due to specific G-protein mutations (see p. 971)
- excess ACTH secretion (**Cushing's disease and Nelson's syndrome**) – corticotroph adenomas, usually 'basophil'.

Many tumours are able to synthesize several pituitary hormones, and occasionally more than one hormone is secreted in clinically significant excess (e.g. both GH and prolactin).

The clinical features of acromegaly, Cushing's disease or hyperprolactinaemia are usually (but not always) obvious, and are discussed on pages 979, 1012, and 980. Hyperprolactinaemia may be clinically 'silent'. Tumours producing LH, FSH or TSH are well described but very rare.

Some common pituitary tumours, usually 'chromophobe' adenomas, cause no clinically apparent hormone excess and are referred to as 'non-functioning' tumours. Laboratory studies such as immunocytochemistry or in situ hybridization



(a) (b) (c)

Fig. 18.8 **(a) Coronal MRI of pituitary**, showing a left-sided lucent intrasellar microadenoma (arrowed). The pituitary stalk is deviated slightly to the right. **(b) Coronal MRI of pituitary**, showing macroadenoma with moderate suprasellar extension, and lateral extension compressing left cavernous sinus. The top of the adenoma is compressing the optic chiasm (arrowed). **(c) Sagittal MRI of head**, showing a pituitary macroadenoma with massive suprasellar extension (arrows).

Table 18.5 Comparisons of primary treatments for pituitary tumours

Treatment method	Advantages	Disadvantages
Surgical		
Trans-sphenoidal adenomectomy or hypophysectomy	Relatively minor procedure Potentially curative for microadenomas and smaller macroadenomas	Some extrasellar extensions may not be accessible Risk of CSF leakage and meningitis
Transfrontal	Good access to suprasellar region	Major procedure; danger of frontal lobe damage High chance of subsequent hypopituitarism
Radiotherapy		
External (40–50 Gy)	Non-invasive Reduces recurrence rate after surgery	Slow action, often over many years Not always effective Possible late risk of tumour induction
Stereotactic	Precise administration of high dose to lesion	Long-term follow-up data limited
Yttrium implantation	High local dose	Only ever used in a few centres
Medical		
Dopamine agonist therapy (e.g. bromocriptine, cabergoline)	Non-invasive; reversible	Usually not curative; significant side-effects in minority
Somatostatin analogue therapy (octreotide, lanreotide)	Non-invasive; reversible	Usually not curative; expensive
Growth hormone receptor antagonist (pegvisomant)	Highly selective	Usually not curative; very expensive

show that these tumours may often produce small amounts of LH and FSH or the α-subunit of LH, FSH and TSH, and occasionally ACTH.

Is there a deficiency of any hormone?

Clinical examination may give clues; thus, short stature in a child with a pituitary tumour is likely to be due to GH deficiency. A slow, lethargic adult with pale skin is likely to be deficient in TSH and/or ACTH. Milder deficiencies may not be obvious, and require specific testing (see Table 18.7).

Treatment

Treatment depends on the type and size of tumour (Table 18.5). In general, therapy has three aims:

Removal/control of tumour

- *Surgery* via the trans-sphenoidal route is usually the treatment of choice. Very large tumours are occasionally removed via the open transfrontal route.
- *Radiotherapy* – by an external three-beam technique, stereotactic or rarely via implant of yttrium needles – is usually employed when surgery is impracticable or incomplete, as it controls but rarely abolishes tumour mass. The standard regimen involves a dose of 45 Gy, given as 20–25 fractions via three fields.
- *Medical therapy* with somatostatin analogues and/or dopamine agonists sometimes cause shrinkage of specific types of tumour (see p. 980) and if successful can be used as primary therapy.

Reduction of excess hormone secretion

Reduction is usually obtained by surgical removal but sometimes by medical treatment. Useful control can be achieved with dopamine agonists for prolactinomas or somatostatin analogues for acromegaly, but ACTH secretion usually cannot be controlled by medical means. Growth hormone antagonists are also available for acromegaly (p. 980).

Replacement of hormone deficiencies

Replacement of hormone deficiencies, i.e. hypopituitarism, is discussed below (see Table 18.8).

Small tumours producing no significant symptoms, pressure or endocrine effects may be observed with appropriate clinical, visual field, imaging and endocrine assessments.

Differential diagnosis of pituitary or hypothalamic masses

Although pituitary adenomas are the most common mass lesion of the pituitary (90%), a variety of other conditions may also present as a pituitary or hypothalamic mass and form part of the differential diagnosis.

Other tumours

- *Craniopharyngioma* (1–2%), a usually cystic hypothalamic tumour, often calcified, arising from Rathke's pouch, often mimics an intrinsic pituitary lesion. It is the most common pituitary tumour in children but may present at any age.
- *Uncommon tumours* include meningiomas, gliomas, chondromas, germinomas and pinealomas. Secondary deposits occasionally present as apparent pituitary tumours, typically presenting with headache and diabetes insipidus.

Hypophysitis and other inflammatory masses

A variety of inflammatory masses occur in the pituitary or hypothalamus. These include rare pituitary-specific conditions (e.g. postpartum hypophysitis, lymphocytic hypophysitis, giant cell hypophysitis) or pituitary manifestations of more generalized disease processes (sarcoidosis, Langerhans' cell histiocytosis, Wegener's granulomatosis). These lesions may be associated with diabetes insipidus and/or an unusual pattern of hypopituitarism.

Other lesions

Carotid artery aneurysms may masquerade as pituitary tumours and must be diagnosed before surgery. Cystic lesions may also present as a pituitary mass, including arachnoid and Rathke cleft cysts.

Hypopituitarism

Pathophysiology

Deficiency of hypothalamic releasing hormones or of pituitary trophic hormones can be selective or multiple. Thus isolated deficiencies of GH, LH/FSH, ACTH, TSH and vasopressin are all seen, some cases of which are genetic and congenital and others sporadic and autoimmune or idiopathic in nature.

Multiple deficiencies usually result from tumour growth or other destructive lesions. There is generally a progressive loss of anterior pituitary function. GH and gonadotrophins are usually first affected. Hyperprolactinaemia, rather than prolactin deficiency, occurs relatively early because of loss of tonic inhibitory control by dopamine. TSH and ACTH are usually last to be affected.

Panhypopituitarism refers to deficiency of all anterior pituitary hormones; it is most commonly caused by pituitary tumours, surgery or radiotherapy. Vasopressin and oxytocin secretion will be significantly affected only if the hypothalamus is involved by a hypothalamic tumour or major suprasellar extension of a pituitary lesion, or if there is an infiltrative/inflammatory process. Posterior pituitary deficiency is rare in an uncomplicated pituitary adenoma.

Genetics of hypopituitarism

Specific genes are responsible for the development of the anterior pituitary involving interaction between signalling molecules and transcription factors. For example, mutations in PROP1 and POUIF1 (previously PIT-1) prevent the differentiation of anterior pituitary cells (precursors to somatotroph, lactotroph, thyrotroph and gonadotroph cells), leading to deficiencies of GH, prolactin, TSH and GnRH. In addition, novel mutations within GH and GHRH receptor genes have been identified which may explain the pathogenesis of isolated GH deficiency in children. Despite these advances, most cases of hypopituitarism do not have specific identifiable genetic causes.

Causes

Disorders causing hypopituitarism are listed in Table 18.6. Pituitary and hypothalamic tumours, and surgical or radiotherapy treatment, are the most common.

Clinical features

Symptoms and signs depend upon the extent of hypothalamic and/or pituitary deficiencies, and mild deficiencies may not lead to any complaint by the patient. In general, symptoms of deficiency of a pituitary-stimulating hormone are the same as primary deficiency of the peripheral endocrine gland (e.g. TSH deficiency and primary hypothyroidism cause similar symptoms due to lack of thyroid hormone secretion).

- *Secondary hypothyroidism* and *adrenal failure* both lead to tiredness and general malaise.
- *Hypothyroidism* causes weight gain, slowness of thought and action, dry skin and cold intolerance.
- *Hypoadrenalism* causes mild hypotension, hyponatraemia and ultimately cardiovascular collapse during severe intercurrent stressful illness.
- *Gonadotrophin* and thus *gonadal deficiencies* lead to loss of libido, loss of secondary sexual hair, amenorrhoea and erectile dysfunction.
- *Hyperprolactinaemia* may cause galactorrhoea and hypogonadism.
- *GH deficiency* causes growth failure in children and impaired well-being in some adults.
- Weight may increase (due to hypothyroidism, see above) or decrease in severe combined deficiency (pituitary cachexia).
- Long-standing *panhypopituitarism* gives the classic picture of pallor with hairlessness ('alabaster skin').

Particular syndromes related to hypopituitarism are:

Kallmann's syndrome. This syndrome is isolated gonadotrophin (GnRH) deficiency (p. 1000).

Septo-optic dysplasia. This is a rare congenital syndrome (associated with mutations in the *HESX-1* gene) presenting in childhood with a clinical triad of midline forebrain abnormalities, optic nerve hypoplasia and hypopituitarism.

Sheehan's syndrome is due to pituitary infarction following postpartum haemorrhage and is rare in developed countries.

Table 18.6 Causes of hypopituitarism

Congenital
Isolated deficiency of pituitary hormones (e.g. Kallmann's syndrome)
POU1F1 (Pit-1), Prop1, HESX1 mutations
Infective
Basal meningitis (e.g. tuberculosis)
Encephalitis
Syphilis
Vascular
Pituitary apoplexy
Sheehan's syndrome (postpartum necrosis)
Carotid artery aneurysms
Immunological
Pituitary antibodies
Neoplastic
Pituitary or hypothalamic tumours
Craniopharyngioma
Meningiomas
Gliomas
Pinealoma
Secondary deposits, especially breast
Lymphoma
Traumatic
Skull fracture through base
Surgery, especially transfrontal
Perinatal trauma
Infiltrations
Sarcoidosis
Langerhans' cell histiocytosis
Hereditary haemochromatosis
Hypophysitis
Postpartum
Lymphocytic
Giant cell
Others
Radiation damage
Fibrosis
Chemotherapy
Empty sella syndrome
'**Functional**'
Anorexia nervosa
Starvation
Emotional deprivation

Pituitary apoplexy. A pituitary tumour occasionally enlarges rapidly owing to infarction or haemorrhage. This may produce severe headache and sudden severe visual loss sometimes followed by acute life-threatening hypopituitarism.

The 'empty sella' syndrome. An 'empty sella' is sometimes reported on pituitary imaging. This is sometimes due to a defect in the diaphragma and extension of the subarachnoid space (cisternal herniation) or may follow spontaneous infarction or regression of a pituitary tumour. All or most of the sella turcica is devoid of apparent pituitary tissue, but, despite this, pituitary function is usually normal, the pituitary being eccentrically placed and flattened against the floor or roof of the fossa.

Investigations

Each axis of the hypothalamic–pituitary system requires separate investigation. However, the presence of normal gonadal function (ovulation/menstruation or normal libido/erections) suggests that multiple defects of anterior pituitary function are unlikely.

Tests range from the simple basal levels (e.g. free T_4 for the thyroid axis), to stimulatory tests for the pituitary, and tests of feedback for the hypothalamus (Table 18.7). Assessment of the hypothalamic–pituitary–adrenal axis remains controversial: basal 0900 h cortisol levels above 400 nmol/L usually indicate an adequate reserve, while levels below 100 nmol/L predict an inadequate stress response. In many cases basal levels are equivocal and a dynamic test is essential: the insulin tolerance test (Box 18.2) is widely regarded as the 'gold standard' but the short ACTH stimulation test (see Box 18.1), though an indirect measure, is used by many as a routine test of hypothalamic–pituitary–adrenal status. Occasionally, the difference between ACTH deficiency and normal HPA axis can be subtle, and the assessment of adrenal reserve is best left to an experienced endocrinologist.

Treatment

- Steroid and thyroid hormones are essential for life. Both are given as oral replacement drugs, as in primary thyroid and adrenal deficiency, aiming to restore the patient to clinical and biochemical normality (Table 18.8) and levels are monitored by routine hormone assays. **NB** Thyroid replacement should not commence until normal glucocorticoid function has been demonstrated or replacement steroid therapy initiated, as an adrenal 'crisis' may otherwise be precipitated.
- Sex hormones are replaced with androgens and oestrogens, both for symptomatic control and to prevent long-term problems related to deficiency (e.g. osteoporosis).
- When fertility is desired, gonadal function is stimulated directly by human chorionic gonadotrophin (HCG, mainly acting as LH), purified or biosynthetic gonadotrophins, or indirectly by pulsatile gonadotrophin-releasing hormone (GnRH – also known as luteinizing hormone-releasing hormone, LHRH); all are expensive and time-consuming and should be restricted to specialist units.
- GH therapy is given in the growing child, under the care of a paediatric endocrinologist. In adult GH deficiency, GH therapy also produces improvements in body composition, work capacity and psychological well-being, together with reversal of lipid abnormalities associated with a high cardiovascular risk, and this may result in significant symptomatic benefit in some cases. The long-term safety and efficacy of GH therapy is the subject of ongoing multinational studies. It is expensive and in the UK costs £2500–6000 per annum.
- Glucocorticoid deficiency may mask impaired urine concentrating ability, diabetes insipidus only becoming apparent after steroid replacement.

Ask the authors
www.studentconsult.com

Table 18.7 Tests for hypothalamic–pituitary (HP) function

- All hormone levels are measured in plasma unless otherwise stated.
- Tests **shown in bold** are those normally measured on a single basal 0900 h sample in the initial assessment of pituitary function.

Axis	Basal investigations			
	Pituitary hormone	End-organ product/function	Common dynamic tests	Other tests
Anterior pituitary				
HP-ovarian	**LH** **FSH**	**Oestradiol** Progesterone (day 21 of cycle)		Ovarian ultrasound LHRH test*
HP-testicular	**LH** **FSH**	**Testosterone**		Sperm count LHRH test*
Growth	GH	IGF-1 IGF-BP3	Insulin tolerance test Glucagon test	GH response to sleep, exercise or arginine infusion GHRH test*
Prolactin	**Prolactin**	**Prolactin**	–	–
HP-thyroid	**TSH**	**Free T_4**, T_3		TRH test*
HP-adrenal	ACTH	**Cortisol**	Insulin tolerance test Short ACTH (tetracosactide) stimulation test	Glucagon test CRH test* Metyrapone test
Posterior pituitary				
Thirst and osmoregulation		**Plasma/urine osmolality**	Water deprivation test	Hypertonic saline infusion

*Releasing hormone tests were a traditional part of pituitary function testing, but have been largely replaced by the advent of more reliable assays for basal hormones. They test only the 'readily releasable pool' of pituitary hormones and normal responses may be seen in hypopituitarism

Box 18.2 Insulin tolerance test

Indication

Diagnosis or exclusion of ACTH and growth hormone deficiency

Procedure

Test explained to patient and consent obtained
Should only be performed in experienced, specialist units
Exclude cardiovascular disease (ECG), epilepsy or unexplained blackouts; exclude severe untreated hypopituitarism (basal cortisol must be >100 nmol/L; normal free T_4)
Intravenous hydrocortisone and glucose available for emergency
Overnight fast, begin at 0800–0900 h
Soluble insulin, 0.15 Units/kg, i.v. at time 0
Glucose, cortisol and GH levels at 0, 30, 45, 60, 90, 120 min

Normal response

Cortisol rises above 550 nmol/L*
GH rises above 7 ng/L (severe deficiency = <3 ng/L (<9 mU/L))
Glucose must be <2.2 mmol/L to achieve adequate stress response

*Precise cortisol normal ranges are variable between laboratories and assays – appropriate local reference ranges must be used.

Table 18.8 Replacement therapy for hypopituitarism

Axis	Usual replacement therapies
Adrenal	Hydrocortisone 15–40 mg daily (starting dose 10 mg on rising/5 mg lunchtime/5 mg evening) (Normally no need for mineralocorticoid replacement)
Thyroid	Levothyroxine 100–150 μg daily
Gonadal	
Male	Testosterone intramuscularly, orally, transdermally or implant
Female	Cyclical oestrogen/progestogen orally or as patch
Fertility	HCG plus FSH (purified or recombinant) or pulsatile GnRH to produce testicular development, spermatogenesis or ovulation
Growth	Recombinant human GH used routinely to achieve normal growth in children Also advocated for replacement therapy in adults where GH has effects on muscle mass and well-being
Thirst	Desmopressin 10–20 μg one to three times daily by nasal spray or orally 100–200 μg three times daily Carbamazepine, thiazides and chlorpropamide are very occasionally used in mild diabetes insipidus
Breast (prolactin inhibition)	Dopamine agonist (e.g. cabergoline, 500 μg weekly)

FURTHER READING

Dattani MT. Growth hormone deficiency and combined pituitary deficiency: does the genotype matter? *Clinical Endocrinology* 2005; **63**: 121–130.

National Institute for Clinical Excellence. *Human growth hormone (somatotropin) in adults with growth hormone deficiency: Technology appraisal 64,* 2003. http://www.nice.org.uk/pdf/TA64_HGHadults_fullguidance.pdf.

Schneider H J, Aimaretti G, Kreitschmann-Andermahr I et al. Hypopituitarism. *Lancet* 2007; **369**: 1151–1470.

GROWTH AND ABNORMAL STATURE

Physiology and control of growth hormone (GH) (Fig. 18.9)

GH is the pituitary factor responsible for stimulation of body growth in humans. Its secretion is stimulated by GHRH, released into the portal system from the hypothalamus; it is also under inhibitory control by somatostatin. A separate GH stimulating system involves a distinct receptor (GH secretogogue receptor), which interacts with ghrelin (see p. 229). It is not known how these two systems interact but because ghrelin is synthesized in the stomach, it suggests a nutritional role for GH.

- *GH* acts by binding to a specific (single transmembrane) receptor located mainly in the liver (Table 18.3). This induces an intracellular phosphorylation cascade involving the JAK/STAT (signal transducing activators of transcription) pathway (p. 24). STAT proteins are translocated from the cytoplasm into the cell nucleus and cause GH-specific effects by binding to nuclear DNA.
- *IGF-1 (insulin-like growth factor-1)* stimulates growth and its hepatic secretion is stimulated by a tissue-specific effect of GH on the liver. Plasma levels of IGF-1, however, reflect local growth activity poorly, partly as there are multiple IGF-binding proteins (IGF-BP) – mainly IGF-BP3.

The metabolic actions of the system are:

- increasing collagen and protein synthesis
- promoting retention of calcium, phosphorus and nitrogen, necessary substrates for anabolism
- opposing the action of insulin.

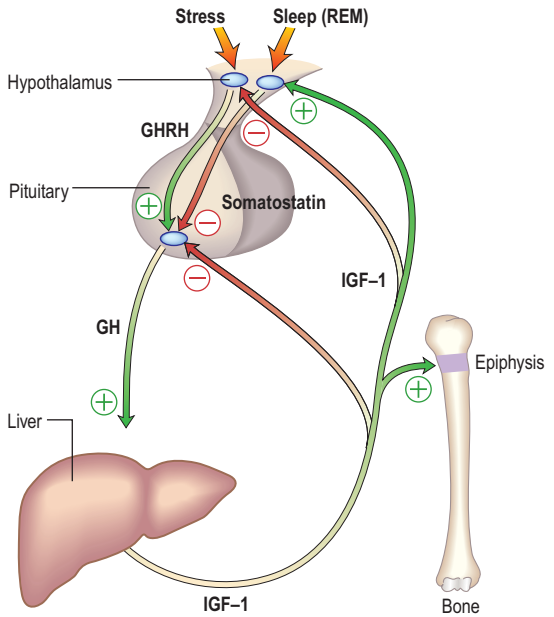


Fig. 18.9 **The control of growth hormone (GH) and insulin-like growth factor-1 (IGF-1).** Pituitary GH is secreted under dual control of GHRH and somatostatin and stimulates release of IGF-1 in liver and elsewhere. IGF-1 has peripheral actions including bone growth and exerts negative feedback to hypothalamus and pituitary.

GH release is intermittent and mainly nocturnal, especially during REM sleep. The frequency and size of GH pulses increase during the growth spurt of adolescence and decline thereafter. Acute stress and exercise both stimulate GH release while, in the normal subject, hyperglycaemia suppresses it.

IGF-1 may, in addition, play a major role in maintaining neoplastic growth. A relationship has been shown between circulating IGF-1 concentrations and breast cancer in premenopausal women and prostate cancer in men.

NORMAL GROWTH

There are factors other than GH involved in linear growth in the human.

- *Genetic factors*. Children of two short parents will probably be short and vice versa.
- *Nutritional factors*. Adequate nutrients must be available. Impaired growth can result from inadequate dietary intake or small bowel disease (e.g. coeliac disease).
- *General health*. Any serious systemic disease in childhood is likely to reduce growth (e.g. chronic kidney disease or chronic infection).
- *Intrauterine growth retardation*. These infants often grow poorly in the long term, while infants with simple prematurity usually catch up. There is some evidence that low birthweight may predispose to hypertension, diabetes and other health problems in later adult life (p. 208).
- *Emotional deprivation and psychological factors*. These can impair growth by complex, poorly understood mechanisms, probably involving temporarily decreased GH secretion.

The relevant aspects of history and examination in the assessment of problems are shown in Box 18.3.

Assessment of growth

Charts showing ranges of height and weight for normal British children are available (Fig. 18.10), and other national data are available. Height must be measured, ideally at the same time of day on the same instrument by the same observer.

Box 18.3 Assessment of problems of growth and development

History

Pregnancy records
Rate of growth (home/school records, e.g. heights on kitchen door)
Comparison with peers at school and siblings
Change in appearance (old photographs)
Change in shoe/glove/hat size or frequency of 'growing out'
Age of appearance of pubic hair, breasts, menarche

Physical signs

Evidence of systemic disease
Body habitus, size, relative weight, proportions (span versus height)
Skin thickness, interdental separation
Facial features
Spade hands/feet
Grading of secondary sexual characteristics

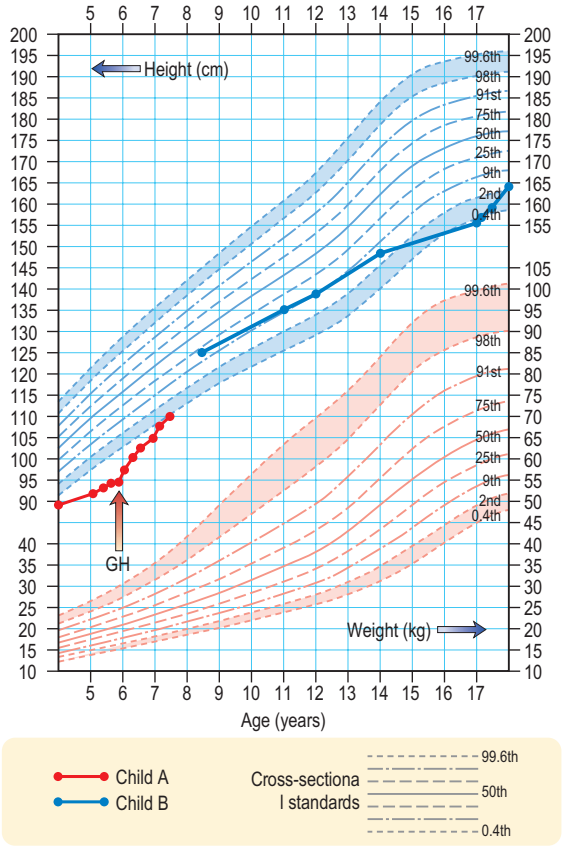


Fig. 18.10 **A height chart for boys. Child A illustrates the course of a child with hypopituitarism**, initially treated with cortisol and thyroxine, but showing growth only after growth hormone treatment. **Child B shows the course of a child with constitutional growth delay without treatment.** Chart © Child Growth Foundation, further information www.healthforallchildren.co.uk.

In general, there are three overlapping phases of growth: infantile (0–2 years), which appears largely substrate (food) dependent; childhood (age 2 years to puberty), which is largely GH dependent; and the adolescent 'growth spurt', dependent on GH and sex hormones.

Height velocity is more helpful than current height. It requires at least two measurements some months apart and, ideally, multiple serial measurements. Height velocity is the rate of current growth (cm per year), while the current attained height is largely dependent upon previous growth.

Standard deviation scores (SDS) based on the degree of deviation from age–sex norms are widely used – these and growth velocities are far more sensitive than simple charts in assessing growth. Computer programs also allow calculation of many of these indices.

The approximate future height of a child ('mid-parental height') can be simply predicted from the parental heights. For a boy, this is:

[(Maternal height + 14 cm (5.5 inches) + Paternal height)/2]

and for a girl:

[(Paternal height – 14 cm (5.5 inches) + Maternal height)/2].

Thus, with a father of 180 cm and mother of 154 cm, the predicted heights are 174 cm for a son and 160 cm for a daughter.

GROWTH FAILURE: SHORT STATURE

When children or their parents complain of short stature, particular attention should focus on:

- intrauterine growth retardation, weight and gestation at birth
- possible systemic disorders – any system, but especially small bowel disease
- evidence of skeletal, chromosomal or other congenital abnormalities
- endocrine status – particularly thyroid
- dietary intake and use of drugs, especially steroids for asthma
- emotional, psychological, family and school problems.

School, general practitioner, clinic and home records of height and weight should be obtained if possible to allow growth-velocity calculation. If unavailable, such data must be obtained prospectively.

A child with normal growth velocity is unlikely to have significant endocrine disease and the commonest cause of short stature in this situation is pubertal or 'constitutional' delay. However, low growth velocity without apparent systemic cause requires further investigation. Sudden cessation of growth suggests major physical disease; if no gastrointestinal, respiratory, renal or skeletal abnormality is apparent, then a cerebral tumour or hypothyroidism is likeliest.

Consistently slow-growing children require full endocrine assessment. Features of the more common causes of growth failure are given in Table 18.9.

Around the time of puberty, where constitutional delay is clearly shown and symptoms require intervention, then very-low-dose sex steroids in 3- to 6-month courses will usually induce acceleration of growth.

Investigations

Systemic disease having been excluded, perform:

- **Thyroid function tests** – serum TSH and free T_4 to exclude hypothyroidism.
- **GH status**. Basal levels are of little value, though urinary GH measurements may prove to be of some value in screening. Dynamic tests include the GH response to insulin (the 'gold standard'; Box 18.2), glucagon, arginine, exercise and clonidine. Tests should only be performed in centres experienced in their use and interpretation. Normal responses depend on test and GH assay used.
- **Blood levels of IGF-1 (insulin-like growth factor-1) and IGF-BP3 (binding protein 3)** may provide evidence of GH undersecretion.
- **Assessment of bone age**. Non-dominant hand and wrist X-rays allow assessment of bone age by comparison with standard charts.
- **Karyotyping in females**. Turner's syndrome (p. 1007) is associated with short stature. It is thought that this is due to a defect in the short stature homeobox (*SHOX*) gene which has a role in non-GH mediated growth.

Treatment

Systemic illness should be treated and primary hypothyroidism treated with levothyroxine.

For GH insufficiency, recombinant GH (somatropin) is given as nightly injections in doses of 0.17–0.35 mg/kg per

Table 18.9 Clinical features of common causes of short stature

Cause	Family history	Growth pattern, clinical features and puberty	Bone age	Remarks
Constitutional delay	Often present	Slow from birth, immature but appropriate with late but spontaneous puberty	Moderate delay	Often difficult to differentiate from GH deficiency Growth velocity measurement vital
Familial short stature	Positive	Slow from birth, clinically normal with normal puberty	Normal	Need heights of family members Growth velocity normal
GH insufficiency	Rare	Slow growth, immature, often overweight, delayed puberty	Moderate delay, increasing with time	Early investigation and treatment vital Increased suspicion if child is plump
Primary hypothyroidism	Rare	Slow growth, immature and delayed puberty	Marked delay	Measure TSH, T_4 in all cases of short stature Clear clinical signs not obvious
Small bowel disease	Sometimes	Slow, immature, usually thin for height, delayed puberty	Delayed	Diarrhoea and/or macrocytosis/ anaemia Occasionally no GI symptoms

week. Treatment is expensive and should be supervised in expert centres. Human GH (collected from pituitaries) was previously used but was withdrawn as cases of Creutzfeldt–Jakob disease were reported.

GH treatment in so-called 'short normal' children has not been shown to produce any worthwhile increase in final height. In Turner's syndrome (see p. 1007) large doses of GH are effective in increasing final height, especially in combination with appropriate very-low-dose oestrogen replacement. Familial cases of resistance to GH owing to an abnormal GH receptor (Laron-type dwarfism) are well described. They are very rare but may respond to therapy with synthetic IGF-1.

TALL STATURE

The most common causes are hereditary (two tall parents!), idiopathic (constitutional) or early development. It can occasionally be due to hyperthyroidism. Other causes include chromosomal abnormalities (e.g. Klinefelter's syndrome, Marfan's syndrome) or metabolic abnormalities. GH excess is a very rare cause and is usually clinically obvious).

PITUITARY HYPERSECRETION SYNDROMES

ACROMEGALY AND GIGANTISM

Growth hormone stimulates skeletal and soft tissue growth. GH excess therefore produces gigantism in children (if acquired before epiphyseal fusion) and acromegaly in adults.

Both are due to a pituitary tumour in almost all cases. Hyperplasia due to GHRH excess is very rare. Overall incidence is approximately 3–4/million per year and the prevalence is 50–80/million world-wide.

Clinical features

Symptoms and signs of acromegaly are shown in Figure 18.11. One-third of patients present with changes in appearance, one-quarter with visual field defects or headaches; in the remainder the diagnosis is made by an alert observer in another clinic, e.g. GP, diabetic, hypertension, dental, dermatology. Sleep apnoea is common and requires investigation and treatment if there are suggestive symptoms (see p. 842). Sweating, headaches and soft tissue swelling are particularly useful symptoms of persistent growth hormone secretion. Headache is very common in acromegaly and may be severe even with small tumours; it is often improved after surgical cure or with somatostatin analogues.

Investigations

- **GH levels** may exclude acromegaly if undetectable but a detectable value is non-diagnostic taken alone. Normal adult levels are <1 mU/L for most of the day except during stress or a 'GH pulse'.
- **A glucose tolerance test** is diagnostic if there is no suppression of GH. Acromegalics fail to suppress GH below 1 mU/L and some show a paradoxical rise; about 25% of acromegalics have a diabetic glucose tolerance test.
- **IGF-1 levels** are almost always raised in acromegaly – a single plasma level of IGF-1 reflects mean 24-hour GH levels and is useful in diagnosis. A normal IGF-1 together with GH <5 mU/L (2.5 ng/L) may be taken to exclude acromegaly if the diagnosis is clinically unlikely.
- **Visual field examination** – defects are common, e.g. bitemporal hemianopia.
- **MRI scan** of pituitary if above tests abnormal. This will almost always reveal the pituitary adenoma.
- **Pituitary function** – partial or complete anterior hypopituitarism is common.
- **Prolactin** – mild to moderate hyperprolactinaemia occurs in 30% of patients (see Fig. 18.13). In some, the adenoma secretes both GH and prolactin.

Management and treatment

Untreated acromegaly results in markedly reduced survival. Most deaths occur from heart failure, coronary artery disease and hypertension-related causes. In addition, there is an increase in deaths due to neoplasia, particularly large bowel tumours. Treatment is therefore indicated in all except the elderly or those with minimal abnormalities. The aim of therapy is to achieve a mean growth hormone level below 5 mU/L (or 2.5 ng/L); this has been shown to reduce mortality to normal levels. A normal IGF-1 is also a goal of therapy.

Symptoms

Change in appearance
Increased size of hands/feet
Headaches
Excessive sweating
Visual deterioration
Tiredness
Weight gain
Amenorrhoea
oligomenorrhoea
in women
Galactorrhoea
Impotence or poor libido
Deep voice
Goitre
Breathlessness
Pain/tingling in hands
Polyuria/polydipsia
Muscular weakness
Joint pains

Old photographs are frequently useful
Symptoms of hypopituitarism may also be present

Signs

Prominent supraorbital ridge
Prognathism
Interdental separation
Large tongue
Hirsutism
Thick greasy skin
Spade-like hands and feet
Tight rings
Carpal tunnel syndrome
Colonic polyps
Visual field defects
Galactorrhoea
Hypertension
Oedema
Heart failure
Arthropathy
Proximal myopathy
Glycosuria
(plus possible signs of hypopituitarism)

Fig. 18.11 **Acromegaly – symptoms and signs.** Bold type indicates signs of greater discriminant value.

Complete cure is often slow to achieve, if possible at all. The general pros and cons of surgery, radiotherapy and medical treatment are discussed on page 973. Progress can be assessed by monitoring GH and IGF-1 levels.

When present, hypopituitarism should be corrected (see p. 975) and concurrent diabetes and/or hypertension should be treated conventionally; both usually improve with treatment of the acromegaly. Some specialist centres advocate regular colonoscopy to detect and remove colonic polyps to reduce the risk of colonic cancer.

- **Surgery**. Trans-sphenoidal surgery is the appropriate first-line therapy. It will result in clinical remission in a majority of cases (60–90%) with pituitary microadenoma, but in only 50% of those with macroadenoma. Very high pre-operative GH and IGF-1 levels are also poor prognostic markers of surgical cure. Surgical success rates are variable and highly dependent upon experience, and a specialist pituitary surgeon is essential. Transfrontal surgery is rarely required except for massive macroadenomas. There is approximately a 10% recurrence rate.
- **Pituitary radiotherapy**. External radiotherapy is normally used after pituitary surgery fails to normalize GH levels rather than as primary therapy. It is often combined with medium-term treatment with a somatostatin analogue or a dopamine agonist because of the slow biochemical response to radiotherapy, which may take 10 years or more and is often associated with hypopituitarism which makes it unattractive in patients of reproductive age. Stereotactic radiotherapy is used in some centres.
- **Medical therapy**. There are three receptor targets for the treatment of acromegaly, viz. pituitary somatostatin receptors, dopamine (D_2) receptors and growth hormone receptors in the periphery.
  - **Somatostatin receptor agonists**. *Octreotide* and *lanreotide* are synthetic analogues of somatostatin (p. 977) that selectively act on somatostatin receptor subtypes (SST2 and SST5), which are highly expressed in growth-hormone-secreting tumours. These drugs were used as a short-term treatment whilst other modalities become effective, but now are sometimes used as primary therapy. They reduce GH and IGF levels in most patients. Both drugs are typically administered as monthly depot injections and are generally well tolerated but are associated with an increased incidence of gallstones and are expensive.
  - **Dopamine agonists**. Dopamine agonists act on D_2 receptors (p. 907) and can be given to shrink tumours prior to definitive therapy or to control symptoms and persisting GH secretion; they are probably most effective in mixed growth-hormone-producing (somatotroph) and prolactin-producing (mammotroph) tumours. The doses are bromocriptine 10–60 mg daily or cabergoline 0.5 mg daily (higher than for prolactinomas) which should be started slowly (see p. 982). Given alone they reduce GH to 'safe' levels in only a minority of cases – but they are useful for mild residual disease or in combination with somatostatin analogues. Drugs with combined somatostatin and dopamine receptor activity are under development.
  - **Growth hormone antagonists**. Pegvisomant (a genetically modified analogue of GH) is a GH receptor antagonist which has its effect by binding to and preventing dimerization of the GH receptor. It does not lower growth hormone levels or reduce tumour size but has been shown to normalize IGF-1 levels in 90% of patients. Its main role at the present time is treatment of patients in whom GH and IGF levels cannot be reduced to safe levels with somatostatin analogues alone, surgery or radiotherapy.

HYPERPROLACTINAEMIA

The hypothalamic–pituitary control of prolactin secretion is illustrated in Figure 18.12. Prolactin is a large peptide

FURTHER READING

Dattani M, Preece M. Growth hormone deficiency and related disorders. *Lancet* 2004; **363**: 1977–1987.

Melmed S. Medical progress: Acromegaly. *New England Journal of Medicine* 2006; **355**: 2558–2573.

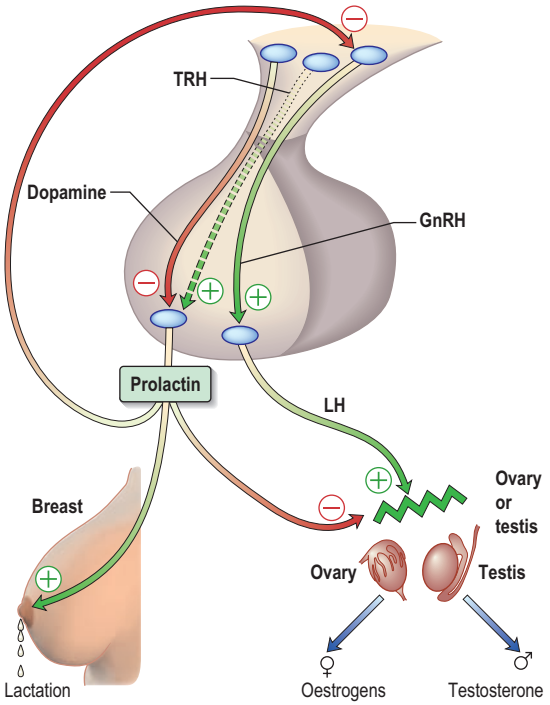


Fig. 18.12 **The control and actions of prolactin.** Prolactin is mainly controlled by tonic inhibition by hypothalamic dopamine. Prolactin stimulates lactation but also inhibits both hypothalamic GnRH secretion and the gonadal actions of LH.

secreted in the pituitary and acts via a transmembrane receptor stimulating JAK 2 and other pathways (Table 8.3).

Prolactin is under tonic dopamine inhibition: factors known to increase prolactin secretion (e.g. TRH) are probably of less relevance. Prolactin stimulates milk secretion (but not breast tissue development) but also reduces gonadal activity. It decreases GnRH pulsatility at the hypothalamic level and, to a lesser extent, blocks the action of LH on the ovary or testis, producing hypogonadism even when the pituitary gonadal axis itself is intact.

The role of prolactin outside pregnancy and lactation is not well defined, although there is some epidemiological evidence of a link between high prolactin levels and breast cancer which has led to an interest in the development of prolactin receptor antagonists.

Physiological hyperprolactinaemia occurs in pregnancy, lactation and severe stress, as well as during sleep and coitus. The range of serum prolactin seen in common causes of hyperprolactinaemia is illustrated in Figure 18.13. Mildly increased prolactin levels (400–600 mU/L) may be physiological and asymptomatic but higher levels require a diagnosis. Levels above 5000 mU/L always imply a prolactin-secreting pituitary tumour.

Causes

Hyperprolactinaemia has many causes. Common pathological causes include prolactinoma, co-secretion of prolactin in tumours causing acromegaly, stalk compression due to pituitary adenomas and other pituitary masses, polycystic ovary syndrome, primary hypothyroidism and 'idiopathic' hyperprolactinaemia. Rarer causes are oestrogen therapy (e.g. the 'pill'), renal failure, liver failure, post-ictal state and chest wall injury. Dopamine antagonist drugs (metoclopramide, domperidone,

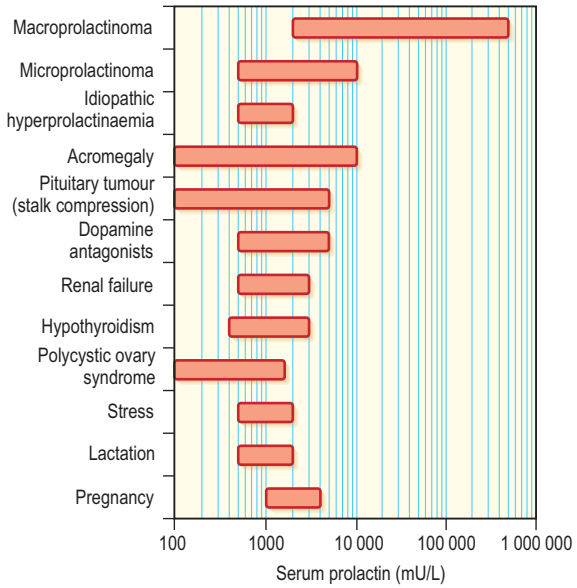


Fig. 18.13 **Range of serum prolactin seen in common causes of hyperprolactinaemia.**

most antipsychotics) are a common iatrogenic cause, as well as most other antiemetics (except cyclizine) and opiates.

Clinical features

Hyperprolactinaemia stimulates milk production in the breast and inhibits GnRH and gonadotrophin secretion per se. It usually presents with:

- galactorrhoea, spontaneous or expressible (60% of cases)
- oligomenorrhoea or amenorrhoea
- decreased libido in both sexes
- decreased potency in men
- subfertility
- symptoms or signs of oestrogen or androgen deficiency – in the long term osteoporosis may result, especially in women
- delayed or arrested puberty in the peripubertal patient
- mild gynaecomastia is often seen in men due to the associated hypogonadism rather than a direct effect of prolactin.

Additionally, headaches and/or visual field defects occur if there is a pituitary tumour (more common in men). Not all patients with galactorrhoea have hyperprolactinaemia, but the other causes are poorly understood – 'normoprolactinaemic galactorrhoea' and duct ectasia.

Investigations

Hyperprolactinaemia should be confirmed by repeat measurement. Further tests are appropriate after physiological and drug causes have been excluded:

- **Visual fields** should be checked.
- **Hypothyroidism** must be excluded since this is a cause of hyperprolactinaemia.
- **Anterior pituitary function** should be assessed if there is any clinical evidence of hypopituitarism or radiological evidence of a pituitary tumour (Table 18.7; Box 18.1 and Box 18.2).
- **MRI of the pituitary** is necessary if there are any clinical features suggestive of a pituitary tumour, and

desirable in all cases when prolactin is significantly elevated (above 1000 mU/L).

In the presence of a pituitary mass on MRI, the level of prolactin helps determine whether the mass is a prolactinoma or a non-functioning pituitary tumour causing stalk-disconnection hyperprolactinaemia: levels of above 5000 mU/L in the presence of a macroadenoma, or above 2000 mU/L in the presence of a microadenoma (or with no radiological abnormality), strongly suggest a prolactinoma (see p. 972). Macroprolactinoma refers to tumours above 10 mm diameter, microprolactinoma to smaller ones.

Occasionally very large prolactinomas can be associated with such high serum prolactin levels that some assays give an artefactual falsely low result (known as the 'hook effect'). If suspected, this can be excluded by serial dilutions of the serum sample.

If patients have no clinical features of hyperprolactinaemia but have an apparently elevated prolactin level, the possibility of *'macroprolactinaemia'* should be considered. This is a higher molecular weight complex of prolactin bound to IgG which is physiologically inactive but occurs in a small proportion of normal people and can therefore lead to unnecessary treatment. Macroprolactinaemia can be diagnosed in the laboratory by precipitation of IgG with polyethylene glycol, after which prolactin levels will be normal on testing.

Treatment

Hyperprolactinaemia is usually treated to avoid the long-term effects of oestrogen deficiency (even if the patient would otherwise welcome the lack of periods!) or testosterone deficiency in the male. Exceptions include minor elevations (400–1000 mU/L) with preservation of normal regular menstruation (or normal male testosterone levels) and postmenopausal patients with microprolactinomas who are not taking oestrogen replacement.

Medical treatment. Hyperprolactinaemia is controlled with a dopamine agonist.

- *Cabergoline* (500 μg once or twice a week judged on clinical response and prolactin levels) is the best tolerated and longest-acting drug.
- *Bromocriptine* is the longest-established therapy and therefore preferred if pregnancy is planned: initial doses should be small (e.g. 1 mg), taken with food and gradually increased to 2.5 mg two or three times daily. Side-effects, which prevent effective therapy in a minority of cases, include nausea and vomiting, dizziness and syncope, constipation and cold peripheries.
- *Quinagolide* (75–150 μg once daily) is an alternative.

Complications include pulmonary, retroperitoneal and pericardial fibrotic reactions and cardiac valve lesions (when used in higher doses in Parkinson's disease). Patients need careful monitoring although such adverse effects appear to be very rare in patients on 'endocrine' doses.

In most cases a dopamine agonist will be the first and only therapy and can be used in the long term. Prolactinomas usually shrink in size on a dopamine agonist, and in macroadenomas any pituitary mass effects commonly resolve. Microprolactinomas may not recur after several years of dopamine agonist therapy in a minority of cases, but in the majority hyperprolactinaemia will recur if treatment is stopped.

Trans-sphenoidal surgery may restore normoprolactinaemia in patients with microadenoma, but is rarely completely successful with macroadenomas and risks damage to normal pituitary function. Therefore most patients and physicians elect to continue medical therapy rather than proceed to surgery. Prolactin should therefore always be measured before surgery on any mass in the pituitary region. Some surgeons believe that long-term bromocriptine increases the hardness of the adenoma and makes resection more difficult, but others dissent from this view.

Radiotherapy usually controls adenoma growth and is slowly effective in lowering prolactin but causes progressive hypopituitarism. It may be advocated after medical tumour shrinkage or after surgery in larger tumours, especially where families are complete, but most workers simply advocate continuation of dopamine agonist therapy in responsive cases.

In patients planning *pregnancy*, it is useful to know the size of the pituitary lesion before starting dopamine agonist therapy. Rarely, tumours enlarge during pregnancy to produce headaches and visual field defects. Dopamine agonists, which are traditionally stopped after conception, should be restarted if there are any signs of tumour growth during pregnancy.

HYPERSECRETION OF OTHER PITUITARY HORMONES

Pituitary tumours may rarely secrete TSH (and cause thyrotoxicosis). Such 'TSHomas' lead to the unusual biochemical pattern of elevated fT_4 levels with normal or high circulating TSH levels. FSH and LH secreting pituitary tumours may cause elevated sex steroids but are extremely uncommon.

ACTH Hypersecretion

Cushing's disease and Nelson's syndrome are discussed in the adrenal disease section (p. 1012).

THE THYROID AXIS

The metabolism of virtually all nucleated cells of many tissues is controlled by the thyroid hormones. Overactivity or underactivity of the gland are the most common of all endocrine problems.

Anatomy

The thyroid gland consists of two lateral lobes connected by an isthmus. It is closely attached to the thyroid cartilage and to the upper end of the trachea, and thus moves on swallowing. It is often palpable in normal women.

Embryologically it originates from the base of the tongue and descends to the middle of the neck. Remnants of thyroid tissue can sometimes be found at the base of the tongue (lingual thyroid) and along the line of descent. The gland has a rich blood supply from superior and inferior thyroid arteries.

The thyroid gland consists of follicles lined by cuboidal epithelioid cells. Inside is the colloid (the iodinated glycoprotein thyroglobulin) which is synthesized by the follicular cells. Each follicle is surrounded by basement membrane, between which are parafollicular cells containing calcitonin-secreting C cells.

Physiology

Synthesis. The thyroid synthesizes two hormones, L-thyroxine (T_4) and triiodothyronine (T_3), of which T_3 acts at the cellular level and T_4 is the prohormone. Inorganic iodide is trapped by the gland by an enzyme dependent system, oxidized and incorporated into the glycoprotein thyroglobulin to form mono- and diiodotyrosine and then T_4 and T_3 (Fig. 18.14).

More T_4 than T_3 is produced, but T_4 is converted in some peripheral tissues (liver, kidney and muscle) to the more active T_3 by 5′-monodeiodination; an alternative 3′-monodeiodination yields the inactive reverse T_3 (rT_3). The latter step occurs particularly in severe non-thyroidal illness (see below).

In plasma, more than 99% of all T_4 and T_3 is bound to hormone-binding proteins (thyroxine-binding globulin, TBG; thyroid-binding prealbumin, TBPA; and albumin). Only free hormone is available for tissue action, where T_3 binds to specific nuclear receptors within the cell. Many drugs and other factors affect TBG; all may result in confusing total T_4 levels in blood, and most laboratories therefore now measure free T_4 levels.

Control of the hypothalamic–pituitary–thyroid axis. The control and feedback of the thyroidal axis is discussed in Figure 18.2 and page 967. Thyrotrophin-releasing hormone (TRH), a peptide produced in the hypothalamus, stimulates TSH from the pituitary. TSH stimulates growth and activity of the thyroid follicular cells via the G-protein coupled TSH membrane receptor (Table 18.3). Circulating T_4 is peripherally de-iodinated to T_3 which binds to the thyroid hormone nuclear receptor (TR) on target organ cells to cause modified gene transcription. Two TR receptors exist (TRα and TRβ) and the tissue-specific effects of T_3 are dependent upon the local expression of these TR receptors. TRα knockout mice show poor growth, bradycardia and hypothermia, whilst TRβ knockout mice show thyroid hyperplasia and high T_4 levels in the presence of inappropriately normal circulating TSH, suggesting a role for the latter receptors in thyroid hormone resistance (see p. 995).

Physiological effects of thyroid hormones:

- Cardiovascular system – increased heart rate and cardiac output.
- Skeletal – increased bone turnover and resorption.
- Respiratory – maintains normal hypoxic and hypercapnic drive in respiratory centre.
- Gastrointestinal – increases gut motility.
- Blood – increases red blood cell 2,3-BPG facilitating oxygen release to tissues.
- Neuromuscular – increases speed of muscle contraction and relaxation and muscle protein turnover.

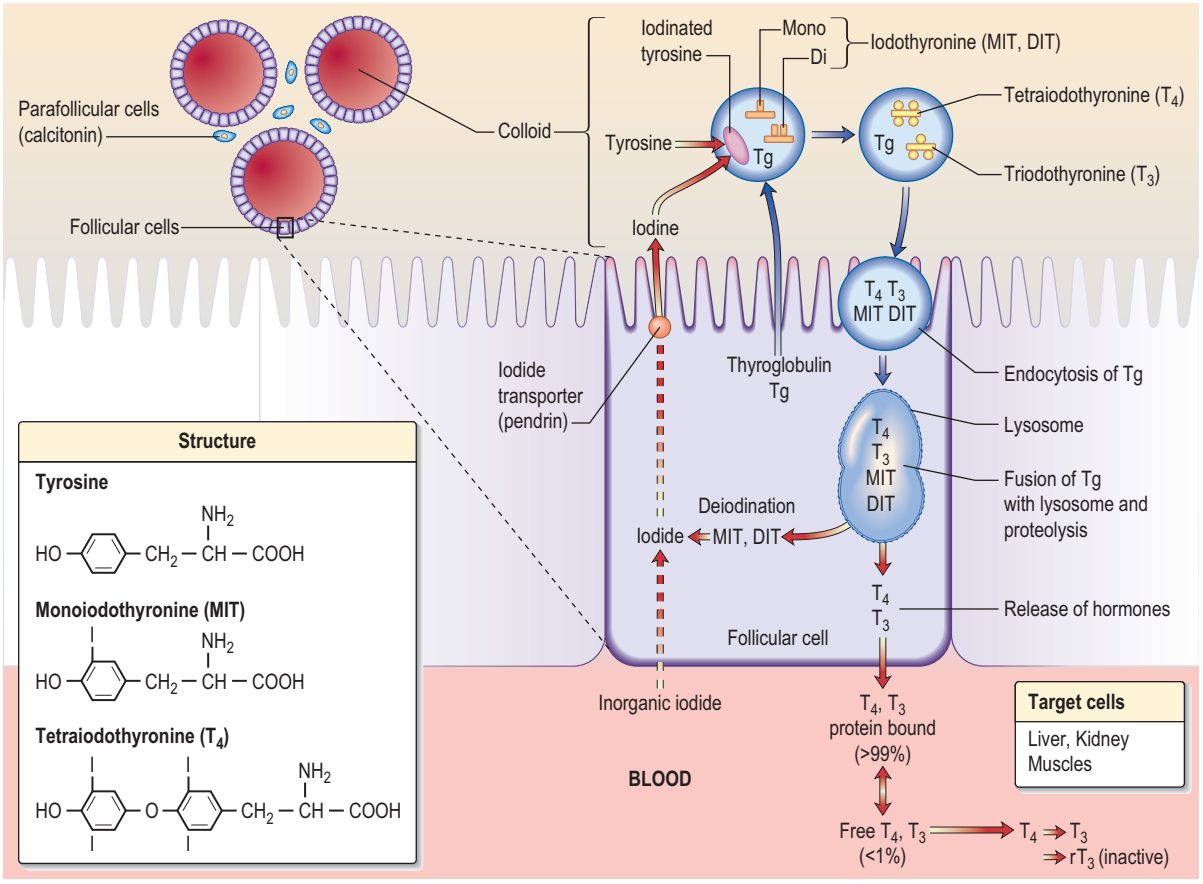


Fig. 18.14 **Synthesis and metabolism of the thyroid hormones.**

- Metabolism of carbohydrates – increases hepatic gluconeogenesis/glycolysis and intestinal glucose absorption.
- Metabolism of lipids – increased lipolysis and cholesterol synthesis and degradation.
- Sympathetic nervous tissue – increases catecholamine sensitivity and β-adrenergic receptor numbers in heart, skeletal muscle, adipose cells and lymphocytes. Decreases cardiac α-adrenergic receptors.

Iodine

Globally, dietary iodine deficiency is a major cause of thyroid disease, as iodine is an essential requirement for thyroid hormone synthesis. The recommended daily intake of iodine should be at least 140 μg, and dietary supplementation of salt and bread has reduced the number of areas where 'endemic goitre' still occurs (see below).

Thyroid function tests

Immunoassays for free T_4, free T_3 and TSH are widely available. There are only minor circadian rhythms, and measurements may be made at any time. Particular uses of the tests are summarized in Table 18.10, with typical findings in common disorders.

TSH measurement

TSH levels can discriminate between hyperthyroidism, hypothyroidism and euthyroidism. There are pitfalls, however. These are mainly with hypopituitarism, with the 'sick euthyroid' syndrome and with dysthyroid eye disease, all of which may give 'false' (i.e. misleading, not incorrect) low results implying hyperthyroidism. As a single test of thyroid function it is the most sensitive in most circumstances, but accurate diagnosis requires at least two tests – for example, TSH plus free T_4 or free T_3 where hyperthyroidism is suspected, TSH plus serum free T_4 where hypothyroidism is likely.

TRH test

This has been rendered almost obsolete by modern sensitive TSH assays except for investigation of hypothalamic–pituitary dysfunction. TRH (protirelin) is occasionally used to differentiate between thyroid hormone resistance and TSHoma in the context of raised fT_4 and TSH levels. Typically, there is a rise in TSH in thyroid hormone resistance, whilst in TSHoma there is a flat response due to continued autonomous TSH secretion which does not respond to TRH.

Problems in interpretation of thyroid function tests

There are three major areas of difficulty.

Serious acute or chronic illness

Thyroid function is affected in several ways:

- reduced concentration and affinity of binding proteins,
- decreased peripheral conversion of T_4 to T_3 with more rT_3
- reduced hypothalamic–pituitary TSH production.

Systemically ill patients can therefore have an apparently low total and free T_4 and T_3 with a normal or low basal TSH (the 'sick euthyroid' syndrome). Levels are usually only mildly below normal and are thought to be mediated by interleukins IL-1 and IL-6; the tests should be repeated after resolution of the underlying illness.

Pregnancy and oral contraceptives

These lead to greatly increased TBG levels and thus to high or high-normal total T_4. Free T_4 is usually normal. Normal ranges for free T_4 and TSH alter with the normal physiological changes during pregnancy and TSH is often slightly suppressed in the first trimester, but this rarely causes clinical problems.

Drugs

Amiodarone decreases T_4 to T_3 conversion and free T_4 levels may therefore be above normal in a euthyroid patient; conversely amiodarone may induce both hyper- and hypothyroidism – the TSH level is usually reliable.

Many drugs affect thyroid function tests by interfering with protein binding but this now rarely causes a problem with free T_4 assays.

Antithyroid antibodies

Serum antibodies to the thyroid are common and may be either destructive or stimulating; both occasionally coexist in the same patient.

Destructive antibodies are directed against the microsomes or against thyroglobulin; the antigen for thyroid microsomal antibodies is the thyroid peroxidase (TPO) enzyme. TPO antibodies are found in up to 20% of the normal population, especially older women, but only 10–20% of these develop overt hypothyroidism.

TSH receptor IgG antibodies (TRAb) typically stimulate, but occasionally block, the receptor; they can be measured in two ways:

- by the inhibition of binding of TSH to its receptors (TSH-binding inhibitory immunoglobulin, TBII)
- by demonstrating that they stimulate the release of cyclic AMP (thyroid-stimulating immunoglobulin/antibody TSI, TSAb).

Table 18.10 **Characteristics of thyroid function tests in common thyroid disorders (the clinically most informative tests in each situation are shown in bold)**

	TSH (0.3–3.5 mU/L)	Free T_4 (10–25 pmol/L)	Free T_3 (3.5–7.5 pmol/L)
Thyrotoxicosis	**Suppressed (<0.05 mU/L)**	**Increased**	**Increased**
Primary hypothyroidism	**Increased (>10 mU/L)**	Low/low-normal	Normal or low
TSH deficiency	Low-normal or subnormal	**Low/low-normal**	Normal or low
T_3 toxicosis	Suppressed (<0.05 mU/L)	Normal	**Increased**
Compensated euthyroidism	**Slightly increased (5–10 mU/L)**	**Normal**	Normal

HYPOTHYROIDISM

Pathophysiology

Underactivity of the thyroid is usually primary, from disease of the thyroid, but may be secondary to hypothalamic–pituitary disease (reduced TSH drive) (Table 18.11). Primary hypothyroidism is one of the most common endocrine conditions with an overall UK prevalence of over 1% in women, but under 0.1% in men; lifetime prevalence for an individual is higher – perhaps as high as 9% for women and 1% for men with mean age at diagnosis around 60 years. The worldwide prevalence of subclinical hypothyroidism varies from 1 to 10%.

Causes of primary hypothyroidism (Table 18.11)

Autoimmune

Atrophic (autoimmune) hypothyroidism. This is the most common cause of hypothyroidism and is associated with antithyroid autoantibodies leading to lymphoid infiltration of the gland and eventual atrophy and fibrosis. It is six times more common in females and the incidence increases with age. The condition is associated with other autoimmune disease such as pernicious anaemia, vitiligo and other endocrine deficiencies (p. 965). In some instances intermittent hypothyroidism occurs with recovery from disease; antibodies which block the TSH receptor may sometimes be involved in the aetiology.

Hashimoto's thyroiditis. This form of autoimmune thyroiditis, again more common in women and most common in late middle age, produces atrophic changes with regeneration, leading to goitre formation. The gland is usually firm and rubbery but may range from soft to hard. TPO antibodies are present, often in very high titres (>1000 IU/L). Patients may be hypothyroid or euthyroid, though they may go through an initial toxic phase, 'Hashi-toxicity'. Levothyroxine therapy may shrink the goitre even when the patient is not hypothyroid.

Postpartum thyroiditis. This is usually a transient phenomenon observed following pregnancy. It may cause hyperthyroidism, hypothyroidism or the two sequentially. It is believed to result from the modifications to the immune system necessary in pregnancy, and histologically is a lymphocytic thyroiditis. The process is normally self-limiting, but when conventional antibodies are found there is a high chance of this proceeding to permanent hypothyroidism. Postpartum thyroiditis may be misdiagnosed as postnatal depression, emphasizing the need for thyroid function tests in this situation.

Table 18.11 Causes of hypothyroidism

PRIMARY DISEASE OF THYROID	**Infective**
Congenital	Post-subacute thyroiditis
Agenesis	**Post-surgery**
Ectopic thyroid remnants	**Post-irradiation**
Defects of hormone synthesis	Radioactive iodine therapy
Iodine deficiency	External neck irradiation
Dyshormonogenesis	**Infiltration**
Antithyroid drugs	Tumour
Other drugs (e.g. lithium, amiodarone, interferon)	**SECONDARY (TO HYPOTHALAMIC–PITUITARY DISEASE)**
Autoimmune	**Hypopituitarism**
Atrophic thyroiditis	Isolated TSH deficiency
Hashimoto's thyroiditis	**PERIPHERAL RESISTANCE TO THYROID HORMONE**
Postpartum thyroiditis	

Defects of hormone synthesis

Iodine deficiency. Dietary iodine deficiency still exists (p. 225) in some areas as 'endemic goitre' where goitre, occasionally massive, is common. The patients may be euthyroid or hypothyroid depending on the severity of iodine deficiency. The mechanism is thought to be borderline hypothyroidism leading to TSH stimulation and thyroid enlargement in the face of continuing iodine deficiency. Iodine deficiency is still a problem in the Netherlands, Western Pacific, South East Asia (e.g. the mountainous areas of the Himalayas) and Africa. Some countries affected by iodine deficiency, e.g. China and Kazakhstan, are taking action by providing iodine in salt but others, e.g. Russia, have so far not done so. Of the 500 million with iodine deficiency in India, about 2 million have cretinism.

Dyshormonogenesis. This rare condition is due to genetic defects in the synthesis of thyroid hormones; patients develop hypothyroidism with a goitre. One particular familial form is associated with sensorineural deafness due to a deletion mutation in chromosome 7, causing a defect of the transporter pendrin (Pendred's syndrome) (see Fig. 18.14).

Clinical features (Fig. 18.15)

Hypothyroidism produces many symptoms. The alternative term 'myxoedema' refers to the accumulation of mucopolysaccharide in subcutaneous tissues. The classic picture of the slow, dry-haired, thick-skinned, deep-voiced patient with weight gain, cold intolerance, bradycardia and constipation makes the diagnosis easy. Milder symptoms are, however, more common and hard to distinguish from other causes of non-specific tiredness. Many cases are detected on biochemical screening.

Special difficulties in diagnosis may arise in certain circumstances:

- *Children with hypothyroidism* may not show classic features but often have a slow growth velocity, poor school performance and sometimes arrest of pubertal development.
- *Young women with hypothyroidism* may not show obvious signs. Hypothyroidism should be excluded in all patients with oligomenorrhoea/amenorrhoea, menorrhagia, infertility or hyperprolactinaemia.
- *The elderly* show many clinical features that are difficult to differentiate from normal ageing.

Investigation of primary hypothyroidism

Serum TSH is the investigation of choice; a high TSH level confirms primary hypothyroidism. A low free T_4 level confirms the hypothyroid state (and is also essential to exclude TSH deficiency if clinical hypothyroidism is strongly suspected and TSH is normal or low).

Thyroid and other organ-specific antibodies may be present. Other abnormalities include the following:

Symptoms

Tiredness/malaise
Weight gain
Anorexia
Cold intolerance
Poor memory
Change in appearance
Depression
Poor libido
Goitre
Puffy eyes
Dry, brittle unmanageable hair
Dry, coarse skin
Arthralgia
Myalgia
Muscle weakness/Stiffness
Constipation
Menorrhagia or oligomenorrhoea in women
Psychosis
Coma
Deafness

A history from a relative is often revealing
Symptoms of other autoimmune disease may be present

Signs

Mental slowness
Psychosis/dementia
Ataxia
Poverty of movement
Deafness

'Peaches and cream' complexion
Dry thin hair
Loss of eyebrows

Hypertension
Hypothermia
Heart failure
Bradycardia
Pericardial effusion

Cold peripheries
Carpal tunnel syndrome
Oedema

Periorbital oedema
Deep voice
(Goitre)

Dry skin
Mild obesity

Myotonia
Muscular hypertrophy
Proximal myopathy
Slow-relaxing reflexes

Anaemia

Fig. 18.15 **Hypothyroidism – symptoms and signs.** Bold type indicates signs of greater discriminant value. A history from a relative is often revealing. Symptoms of other autoimmune disease may be present.

- *anaemia*, which is usually normochromic and normocytic in type but may be macrocytic (sometimes this is due to associated pernicious anaemia) or microcytic (in women, due to menorrhagia)
- *increased serum aspartate transferase levels*, from muscle and/or liver
- *increased serum creatine kinase levels*, with associated myopathy
- *hypercholesterolaemia* and *hypertriglyceridaemia*
- *hyponatraemia* due to an increase in ADH and impaired free water clearance.

Treatment

Replacement therapy with levothyroxine (thyroxine, i.e. T_4) is given for life. The starting dose will depend upon the severity of the deficiency and on the age and fitness of the patient, especially their cardiac performance: 100 µg daily for the young and fit, 50 µg (increasing to 100 µg after 2–4 weeks) for the small, old or frail. Patients with ischaemic heart disease require even lower initial doses, especially if the hypothyroidism is severe and long-standing. Most physicians would then begin with 25 µg daily and perform serial ECGs, increasing the dose at 3- to 4-week intervals if angina does not occur or worsen and the ECG does not deteriorate.

Monitoring. The aim is to restore T_4 and TSH to well within the normal range. Adequacy of replacement is assessed clinically and by thyroid function tests after at least 6 weeks on a steady dose. If serum TSH remains high, the dose of T_4 should be increased in increments of 25–50 µg with the tests repeated at 6–8 weeks intervals until TSH becomes normal. Complete suppression of TSH should be avoided because of the risk of atrial fibrillation and osteoporosis. The usual maintenance dose is 100–150 µg given as a single daily dose. An annual thyroid function test is recommended – this is usually performed in the primary care setting, often assisted and prompted by district 'thyroid registers'.

Clinical improvement on T_4 may not begin for 2 weeks or more, and full resolution of symptoms may take 6 months. The necessity of lifelong therapy must be emphasized and the possibility of other autoimmune endocrine disease developing, especially Addison's disease or pernicious anaemia, should be considered. During pregnancy, an increase in T_4 dosage of about 25–50 µg is often needed to maintain normal TSH levels, and the necessity of optimal replacement during pregnancy is emphasized by the finding of reductions in cognitive function in children of mothers with elevated TSH during pregnancy.

A few patients with primary hypothyroidism complain of incomplete symptomatic response to T_4 replacement. Combination T_4 and T_3 replacement has been advocated in this context, but randomized clinical trials show no consistent benefit in quality of life symptoms.

Borderline hypothyroidism or 'compensated euthyroidism'

Patients are frequently seen with low-normal serum T_4 levels and slightly raised TSH levels. Sometimes this follows surgery or radioactive iodine therapy when it can reasonably be seen as 'compensatory'. Treatment with levothyroxine is normally recommended where the TSH is consistently above 10 mU/L, or when possible symptoms, high-titre thyroid antibodies, or lipid abnormalities are present. Where the TSH is only marginally raised, the tests should be repeated 3–6 months later. Conversion to overt hypothyroidism is more common in men or when TPO antibodies are present. In practice, vague symptoms in patients with marginally elevated TSH (below 10 mU/L) rarely respond to treatment, but a 'therapeutic trial' of replacement may be needed to

confirm that symptoms are unrelated to the thyroid. It is also considered best to normalize TSH during (and ideally before) pregnancy to avoid fetal adverse effects.

Myxoedema coma

Severe hypothyroidism, especially in the elderly, may present with confusion or even coma. Myxoedema coma is very rare: hypothermia is often present and the patient may have severe cardiac failure, hypoventilation, hypoglycaemia and hyponatraemia. The mortality was previously at least 50% and patients require full intensive care. Optimal treatment is controversial and data lacking; most physicians would advise T_3 orally or intravenously in doses of 2.5–5 µg every 8 hours, then increasing as above. Large intravenous doses should not be used. Additional measures, though unproven, should include:

- oxygen (by ventilation if necessary)
- monitoring of cardiac output and pressures
- gradual rewarming
- hydrocortisone 100 mg i.v. 8-hourly
- glucose infusion to prevent hypoglycaemia.

'Myxoedema madness' Depression is common in hypothyroidism but rarely with severe hypothyroidism in the elderly the patient may become frankly demented or psychotic, sometimes with striking delusions. This may occur shortly after starting T_4 replacement.

Screening for hypothyroidism

- The incidence of *congenital* hypothyroidism is approximately 1 in 3500 births. Untreated, severe hypothyroidism produces permanent neurological and intellectual damage ('cretinism'). Routine screening of the newborn using a blood spot, as in the Guthrie test, to detect a high TSH level as an indicator of primary hypothyroidism is efficient and cost-effective; cretinism is prevented if T_4 is started within the first few months of life.
- Screening of *elderly* patients for thyroid dysfunction has a low pick-up rate and is controversial and not currently recommended. However, patients who have undergone thyroid surgery or received radioiodine should have regular thyroid function tests, as should those receiving lithium or amiodarone therapy.

HYPERTHYROIDISM

Hyperthyroidism (thyroid overactivity, thyrotoxicosis) is common, affecting perhaps 2–5% of all females at some time and with a sex ratio of 5:1, most often between the ages of 20 and 40 years. Nearly all cases (>99%) are caused by intrinsic thyroid disease; a pituitary cause is extremely rare (Table 18.12).

Table 18.12 Causes of hyperthyroidism

Common
Graves' disease (autoimmune)
Toxic multinodular goitre
Solitary toxic nodule/adenoma
Uncommon
Acute thyroiditis
viral (e.g. de Quervain's)
autoimmune
post-irradiation
postpartum
Gestational thyrotoxicosis (HCG stimulated)
Neonatal thyrotoxicosis (maternal thyroid antibodies)
Exogenous iodine
Drugs – amiodarone
Thyrotoxicosis factitia (secret T_4 consumption)
Rare
TSH-secreting pituitary tumours
Metastatic differentiated thyroid carcinoma
HCG-producing tumours
Hyperfunctioning ovarian teratoma (struma ovarii)

Graves' disease

This is the most common cause of hyperthyroidism and is due to an autoimmune process. Serum IgG antibodies bind to TSH receptors in the thyroid, stimulating thyroid hormone production, i.e. they behave like TSH. These TSH receptor antibodies (TSHR-Ab) are specific for Graves' disease, can be measured in serum, are present in 85–90% of cases and decline with treatment. Persistent high levels predict a relapse when drug treatment is stopped. There is an association with HLA-B8, DR3 and DR2 and a 40% concordance rate amongst monozygotic twins with a 5% concordance rate in dizygotic twins. There is a weak association with cytotoxic T lymphocyte-associated antigen 4 (CTLA-4), HLA-DRB*08 and DRB3*0202 on chromosome 6.

Yersinia enterocolitica, Escherichia coli and other Gram-negative organisms contain TSH binding sites. This raises the possibility that the initiating event in the pathogenesis may be an infection with possible 'molecular mimicry' in a genetically susceptible individual, but the precise initiating mechanisms remain unproven in most cases.

Thyroid eye disease accompanies the hyperthyroidism in many cases (see below) but other components of Graves' disease, e.g. Graves' dermopathy, are very rare. Rarely lymphadenopathy and splenomegaly may occur. Graves' disease is also associated with other autoimmune disorders such as pernicious anaemia, vitiligo and myasthenia gravis.

The natural history is one of fluctuation, many patients showing a pattern of alternating relapse and remission; perhaps only 40% of subjects have a single episode. Many patients eventually become hypothyroid.

Other causes of hyperthyroidism/thyrotoxicosis

Solitary toxic adenoma/nodule

This is the cause of about 5% of cases of hyperthyroidism. It does not usually remit after a course of antithyroid drugs.

Toxic multinodular goitre

This commonly occurs in older women. Again, antithyroid drugs are rarely successful in inducing a remission, although they can control the hyperthyroidism.

de Quervain's thyroiditis

This is transient hyperthyroidism from an acute inflammatory process, probably viral in origin. Apart from the toxicosis, there is usually fever, malaise and pain in the neck with tachycardia and local thyroid tenderness. Thyroid function tests show initial hyperthyroidism, the erythrocyte sedimentation rate (ESR) and plasma viscosity are raised, and thyroid uptake scans show suppression of uptake in the acute phase. Hypothyroidism, usually transient, may then follow after a few weeks. Treatment of the acute phase is with aspirin, using short-term prednisolone in severely symptomatic cases.

Postpartum thyroiditis

This is described on page 985.

Amiodarone-induced thyrotoxicosis

Amiodarone, a class III antiarrhythmic drug (see p. 730), causes two types of hyperthyroidism.

Type I amiodarone-induced thyrotoxicosis (AIT) is associated with pre-existing Graves' disease or multinodular goitre. In this situation hyperthyroidism is probably triggered by the high iodine content of amiodarone.

Type II AIT is not associated with previous thyroid disease and is thought to be due to a direct effect of the drug on thyroid follicular cells leading to a destructive thyroiditis with release of T_4 and T_3. Type II AIT may be associated with a hypothyroid phase several months after presentation. Because amiodarone inhibits the deiodination of T_4 to T_3, biochemical presentation of both types of AIT may be associated with higher T_4:T_3 ratios than usual.

Clinical features of hyperthyroidism

The symptoms and signs of hyperthyroidism affect many systems (Fig. 18.16).

Symptomatology and signs vary with age and with the underlying aetiology.

- *The eye signs, pretibial myxoedema and thyroid acropachy* occur only in Graves' disease. Pretibial myxoedema is an infiltration on the shin, essentially occurring only with eye disease (see below). Thyroid acropachy is very rare and consists of clubbing, swollen fingers and periosteal new bone formation.
- *In the elderly*, a frequent presentation is with atrial fibrillation, other tachycardias and/or heart failure, often with few other signs. Thyroid function tests are mandatory in any patient with atrial fibrillation.
- *Children* frequently present with excessive height or excessive growth rate, or with behavioural problems such as hyperactivity. They may also show weight gain rather than loss.
- *So-called 'apathetic thyrotoxicosis'* in some elderly patients presents with a clinical picture more like hypothyroidism. There may be very few signs and a high degree of clinical suspicion is essential.

Differential diagnosis

Hyperthyroidism is often clinically obvious but treatment should never be instituted without biochemical confirmation.

Differentiation of the mild case from anxiety states may be difficult; useful positive clinical markers are eye signs, a diffuse goitre, proximal myopathy and wasting. Weight loss despite a normal or increased appetite is a very useful clinical symptom of hyperthyroidism. The hyperdynamic circulation with warm peripheries seen with hyperthyroidism can be contrasted with the clammy hands of anxiety.

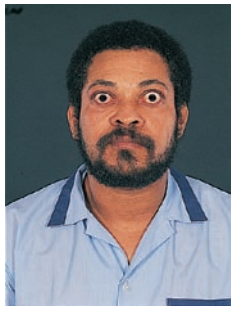
Investigations

- Serum TSH is suppressed in hyperthyroidism (<0.05 mU/L), except for the very rare instances of TSH hypersecretion.
- A raised free T_4 or T_3 confirms the diagnosis; T_4 is almost always raised but T_3 is more sensitive as there are occasional cases of isolated 'T_3 toxicosis'.

Symptoms

Weight loss
Increased appetite
Irritability/behaviour change
Restlessness
Malaise
Stiffness
Muscle weakness
Tremor
Choreoathetosis
Breathlessness
Palpitation
Heat intolerance
Itching
Thirst
Vomiting
Diarrhoea
Eye complaints*
Goitre
Oligomenorrhoea
Loss of libido
Gynaecomastia
Onycholysis
Tall stature (in children)
Sweating

*Only in Graves' disease



Signs

Tremor
Hyperkinesis
Psychosis

Tachycardia or atrial fibrillation
Full pulse
Warm vasodilated peripheries
Systolic hypertension
Cardiac failure

Exophthalmos*
Lid lag and 'stare'
Conjunctival oedema
Ophthalmoplegia*
Periorbital oedema
Goitre, bruit
Weight loss

Proximal myopathy
Proximal muscle wasting
Onycholysis
Palmar erythema

Graves' dermopathy*
Thyroid acropachy
Pretibial myxoedema

*Only in Graves' disease

Fig. 18.16 **Hyperthyroidism – symptoms and signs.** Bold type indicates signs of greater discriminant value.

- TPO and thyroglobulin antibodies are present in most cases of Graves' disease.

TSH receptor antibodies are not measured routinely, but are commonly present: thyroid-stimulating immunoglobulin (TSI) 80% positive, TSH-binding inhibitory immunoglobulin (TBII) 60–90% in Graves' disease (see p. 987).

Treatment

Three possibilities are available: antithyroid drugs, radioiodine and surgery. Practices and beliefs differ widely within and between countries.

Antithyroid drugs

Carbimazole is most often used in the UK, and propylthiouracil is also used. *Thiamazole* (methimazole), the active metabolite of carbimazole, is used in the USA. These drugs inhibit the formation of thyroid hormones and also have other minor actions; carbimazole/thiamazole is also an immunosuppressive agent. Initial doses and side-effects are detailed in Table 18.13.

Although thyroid hormone synthesis is reduced very quickly, the long half-life of T_4 (7 days) means that clinical benefit is not apparent for 10–20 days. As many of the manifestations of hyperthyroidism are mediated via the sympathetic system, beta-blockers are used to provide rapid partial symptomatic control; they also decrease peripheral conversion of T_4 to T_3. Drugs preferred are those without intrinsic sympathomimetic activity, e.g. propranolol (Table 18.13). They should not be used alone for hyperthyroidism except when the condition is self-limiting, as in subacute thyroiditis.

Subsequent management is either by gradual dose titration or a 'block and replace' regimen. Neither regimen has been shown to be unequivocally superior. TSH often remains suppressed for many months after clinical improvement and normalization of T_4 and T_3.

Gradual dose titration

1. Start carbimazole 20–40 mg daily.
2. Review after 4–6 weeks and reduce dose of carbimazole depending on clinical state and T_4/T_3 levels. TSH levels may remain suppressed for several months and are unhelpful at this stage.
3. When clinically and biochemically euthyroid, stop beta-blockers.
4. Review after 2–3 months and, if controlled, reduce carbimazole.
5. Gradually reduce dose to 5 mg daily over 6–24 months if hyperthyroidism remains controlled.
6. When the patient is euthyroid on 5 mg daily carbimazole, discontinue.

Propylthiouracil is used in similar fashion (Table 18.13).

'Block and replace' regimen

With this policy, full doses of antithyroid drugs, usually carbimazole 40 mg daily, are given to suppress the thyroid completely while replacing thyroid activity with 100 µg of levothyroxine daily once euthyroidism has been achieved. This is continued usually for 18 months, the claimed advantages being the avoidance of over- or undertreatment and the better use of the immunosuppressive action of carbimazole. This regimen is contraindicated in pregnancy as T_4 crosses the placenta less well than carbimazole.

Relapse

About 50% of patients will relapse after a course of carbimazole or propylthiouracil, mostly within the following 2 years but occasionally much later. Long-term antithyroid therapy is then used or surgery or radiotherapy is considered (see below). Most patients (90%) with hyperthyroidism have a diffuse goitre but those with large single or multinodular goitres are unlikely to remit after a course of antithyroid drugs and will usually require definitive treatment. Severe biochemical hyperthyroidism is also less likely to remain in remission.

Toxicity

The major side-effect of drug therapy is agranulocytosis that occurs in approximately 1 in 1000 patients, usually within 3 months of treatment. All patients must be warned to seek immediate medical attention for a white blood cell count if they develop unexplained fever or sore throat – written information is essential. Rashes are more frequent and usually require a change of drug. If toxicity occurs on carbimazole, propylthiouracil may be used and vice versa; side-effects are only occasionally repeated on the other drug.

Radioactive iodine

Radioactive iodine (RAI) is given to patients of all ages, although it is contraindicated in pregnancy and while breastfeeding. RAI is the most common treatment modality in the USA whereas antithyroid drugs tend to be favoured in Europe.

Table 18.13 Drugs used in the treatment of hyperthyroidism

Drug	Usual starting dose	Side-effects	Remarks
Antithyroid drugs			
Carbimazole	20–40 mg daily, 8-hourly, or in single dose	Rash, nausea, vomiting, arthralgia, agranulocytosis (0.1%), jaundice	Active metabolite is thiamazole (methimazole) Mild immunosuppressive activity
Propylthiouracil	100–200 mg 8-hourly	Rash, nausea, vomiting, agranulocytosis	Additionally blocks conversion of T_4 to T_3
Beta-blocker for symptomatic control May need higher doses than normal			
Propranolol	40–80 mg every 6–8 hours	Avoid in asthma	Use agents without intrinsic sympathomimetic activity as receptors highly sensitive

^{131}*Iodine* is given in an empirical dose (usually 200–550 MBq) because of variable uptake and radiosensitivity of the gland. It accumulates in the thyroid and destroys the gland by local radiation although it takes several months to be fully effective.

Patients must be rendered euthyroid before treatment. They should stop antithyroid drugs at least 4 days before radioiodine, and not recommence until 3 days after radioiodine. Patients on propylthiouracil should stop antithyroid medication earlier than those on carbimazole before RAI because it has a radioprotective action. Many patients do not need to restart antithyroid medication after treatment.

Early discomfort in the neck and immediate worsening of hyperthyroidism are sometimes seen; if worsening occurs, the patient should receive propranolol (Table 18.13); if necessary carbimazole can be restarted. Euthyroidism normally returns in 2–3 months. Patients with dysthyroid eye disease are more likely to show worsening of eye problems after radioiodine than after antithyroid drugs; this represents a partial contraindication to RAI, although worsening can usually be prevented by steroid administration.

Hypothyroidism affects the majority of subjects over the following 20 years. Seventy five per cent of patients are rendered euthyroid in the short term but a small proportion remain hyperthyroid and may require a second dose of radioiodine. Long-term surveillance of thyroid function is necessary with frequent tests in the first year after therapy, and at least annually thereafter.

Risk of carcinogenesis has been long debated, but the overwhelming evidence suggests that overall cancer incidence and mortality are not increased after radioactive iodine (and indeed are significantly reduced in some studies) but the risk of thyroid cancer is significantly increased, although the risk remains very low in absolute terms.

Surgery: subtotal thyroidectomy

Thyroidectomy should be performed only in patients who have previously been rendered euthyroid. Conventional practice is to stop the antithyroid drug 10–14 days before operation and to give potassium iodide (60 mg three times daily), which reduces the vascularity of the gland.

The operation should be performed only by experienced surgeons to reduce the chance of complications:

- Early postoperative bleeding causing tracheal compression and asphyxia is a rare emergency requiring immediate removal of all clips/sutures to allow escape of the blood/haematoma.
- Laryngeal nerve palsy occurs in 1%. Vocal cord movement should be checked preoperatively. Mild hoarseness is more common and thyroidectomy is best avoided in professional singers!
- Transient hypocalcaemia occurs in up to 10% but with permanent hypoparathyroidism in fewer than 1%.
- Recurrent hyperthyroidism occurs in 1–3% within 1 year, then 1% per year.
- Hypothyroidism occurs in about 10% of patients within 1 year, and this percentage increases with time. It is likeliest if TPO antibodies are positive. Automated computer thyroid registers with annual TSH screening are used in some regions, and have demonstrated that a high proportion of patients become hypothyroid in the long term.

Choice of therapy

Indications for either surgery or radioiodine are:

- patient choice
- persistent drug side-effects
- poor compliance with drug therapy
- recurrent hyperthyroidism after drugs.

Particular indications for surgery include:

- a large goitre, which is unlikely to remit after antithyroid medication.

Special situations in hyperthyroidism

Thyroid crisis or 'thyroid storm'

This rare condition, with a mortality of 10%, is a rapid deterioration of hyperthyroidism with hyperpyrexia, severe tachycardia, extreme restlessness, cardiac failure and liver dysfunction. It is usually precipitated by stress, infection or surgery in an unprepared patient, or radioiodine therapy. With careful management it should no longer occur and most cases referred as 'crisis' are simply severe but uncomplicated thyrotoxicosis.

Treatment is urgent. Propranolol in full doses is started immediately together with potassium iodide, antithyroid drugs, corticosteroids (which suppress many of the manifestations of hyperthyroidism) and full supportive measures. Control of cardiac failure and tachycardia is also necessary.

Hyperthyroidism in pregnancy and neonatal life

The high level of HCG found in pregnancy is a weak stimulator of the TSH receptor, commonly causing suppressed TSH with slightly elevated fT_4/fT_3 in the first trimester which may be associated with hyperemesis gravidarum. True maternal hyperthyroidism during pregnancy is however uncommon and usually mild. Diagnosis can be difficult because of the overlap with symptoms of normal pregnancy and misleading thyroid function tests, although TSH is largely reliable. The pathogenesis is almost always Graves' disease. Thyroid-stimulating immunoglobulin (TSI) crosses the placenta to stimulate the fetal thyroid. Carbimazole also crosses the placenta, but T_4 does so poorly, so a 'block-and-replace' regimen is contraindicated. The smallest dose of carbimazole necessary is used and the fetus must be monitored (see below). The paediatrician should be informed and the infant checked immediately after birth – overtreatment with carbimazole can cause fetal goitre. Breast-feeding while on usual doses of carbimazole or propylthiouracil appears to be safe.

If necessary (high doses needed, poor patient compliance or drug side-effects), surgery can be performed, preferably in the second trimester. Radioactive iodine is absolutely contraindicated.

The fetus and maternal Graves' disease

Any mother with a history of Graves' disease may have circulating TSI. Even if she has been treated (e.g. by surgery), the immunoglobulin may still be present to stimulate the fetal thyroid, and the fetus can thus become hyperthyroid, while the mother remains euthyroid.

Any such patient should therefore be monitored during pregnancy. Fetal heart rate provides a direct biological assay

Ask the authors
www.studentconsult.com

of fetal thyroid status, and monitoring should be performed at least monthly. Rates above 160 per minute are strongly suggestive of fetal hyperthyroidism, and maternal treatment with carbimazole and/or propranolol is used. Direct measurement of TSHR-Ab may be helpful to predict neonatal thyrotoxicosis in this situation. To prevent the mother becoming hypothyroid, T_4 may be given as this does not easily cross the placenta. Sympathomimetics, used to prevent premature labour, are contraindicated as they may provoke fatal tachycardia in the fetus.

Hyperthyroidism may also develop in the neonatal period as TSI has a half-life of approximately 3 weeks. Manifestations in the newborn include irritability, failure to thrive and persisting weight loss, diarrhoea and eye signs. Thyroid function tests are difficult to interpret as neonatal normal ranges vary with age.

Untreated neonatal hyperthyroidism is probably associated with hyperactivity in later childhood.

Thyroid hormone resistance

Thyroid hormone resistance is an inherited condition caused by an abnormality of the thyroid hormone receptor. Mutations to the receptor (TRβ) result in the need for higher levels of thyroid hormones to achieve the same intracellular effect. As a result, the normal feedback control mechanisms (see Fig. 18.2, p. 967) result in high blood levels of T_4 with a normal TSH in order to maintain a euthyroid state. This has two consequences:

- First, thyroid function tests appear abnormal even when the patient is euthyroid and requires no treatment. Specialist review is required to differentiate from hyperthyroidism due to inappropriate TSH secretion.
- Second, different tissues contain different thyroid hormone receptors and, in some families, receptors in certain tissues may have normal activity. In this case the level of thyroid hormones to maintain euthyroidism at pituitary and hypothalamic levels (which controls secretion of TSH) may be higher than that required in other tissues such as heart and bone, so that these tissues may exhibit 'thyrotoxic' effects in spite of a normal serum TSH. This 'partial thyroid hormone resistance' can be very difficult to manage effectively.

Long-term consequences of hyperthyroidism

Long-term follow-up studies of hyperthyroidism show a slight increase in overall mortality, which affects all age groups, is not fully explained and tends to occur in the first year after diagnosis. Thereafter, the only long-term risk of adequately treated hyperthyroidism appears to be an increased risk of osteoporosis. Patients with persistently suppressed TSH levels have an increased likelihood of developing atrial fibrillation which may predispose to thromboembolic disease.

THYROID EYE DISEASE

This is also known as dysthyroid eye disease or ophthalmic Graves' disease.

Pathophysiology

The ophthalmopathy of Graves' disease is due to a specific immune response that causes retro-orbital inflammation. Swelling and oedema of the extraocular muscles lead to limitation of movement and to proptosis which is usually bilateral but can sometimes be unilateral. Ultimately increased pressure on the optic nerve may cause optic atrophy. Histology shows focal oedema and glycosaminoglycan deposition followed by fibrosis. The precise autoantigen which leads to the immune response remains to be identified, but it appears to be an antigen in retro-orbital tissue with similar immunoreactivity to the TSH receptor.

Eye disease is a manifestation of Graves' disease and can occur in patients who may be hyperthyroid, euthyroid or hypothyroid. Thyroid dysfunction and ophthalmopathy usually occur within two years of each other although sometimes a gap of many years is seen. TSH receptor antibodies are almost invariably found in the serum but their role in the pathogenesis is unclear. Ophthalmopathy is more common and more severe in smokers.

Clinical features

The clinical appearances are characteristic (Fig. 18.16) but thyroid eye disease demonstrates a wide range of severity. A high proportion of patients with Graves' disease notice some soreness, painful watering or prominence of the eyes, and the 'stare' of lid retraction is relatively common. More severe proptosis occurs in a minority of cases, and limitation and discomfort of eye movement and visual impairment due to optic nerve compression are relatively uncommon. Proptosis and lid retraction may limit the ability to close the eyes completely so that corneal damage may occur. There is periorbital oedema and conjunctival oedema and inflammation.

Eye manifestations do not parallel the degree of biochemical thyrotoxicosis, or the need for antithyroid therapy, but exacerbation of eye disease is more common after radioiodine treatment (15% vs 3% on antithyroid drugs). Sight is threatened in only 5–10% of cases, but the discomfort and cosmetic problems cause great patient anxiety.

Investigations

Few investigations are necessary if the appearances are characteristic and bilateral. TSH, T_3 and free T_4 are measured.

There are a variety of grading systems but none is universally accepted. It is essential to clearly document eye movements and the degree of oedema and inflammation. The exophthalmos should be measured to allow progress to be monitored. MRI or CT of the orbits will exclude retro-orbital space-occupying lesions and show enlarged muscles and oedema, and may show a taut optic nerve due to raised intra-orbital pressure.

Treatment

If the patient is thyrotoxic this should be treated, but this will not directly result in an improvement of the ophthalmopathy, and hypothyroidism must be avoided as it may exacerbate the eye problem. Smoking should be stopped. Treatment of the eyes may be either local or systemic, and always requires close liaison between specialist endocrinologist and ophthalmologist:

- *Methylcellulose or hypromellose eyedrops* are given to aid lubrication and improve comfort.
- *Some patients gain relief by sleeping upright.*
- *The eyelids can be taped to ensure closure at night.* In the rare case, a tarsorrhaphy (eyelids partly sewn together) is required.

- *Systemic steroids* (prednisolone 30–120 mg daily) usually reduce inflammation if more severe symptoms are present. Pulse intravenous methylprednisolone may be more rapidly effective in severe cases.
- *Irradiation of the orbits* (20 Gy in divided doses) is used in severe instances. This improves inflammation and ocular motility but has little effect on proptosis and its precise role is debated.
- *Lid surgery* will protect the cornea if lids cannot be closed.
- *Surgical decompression of the orbit(s)* may be required, particularly if pressure of orbital contents on the optic nerve threatens vision, and at a later, stable stage for cosmetic reasons.
- *Corrective eye muscle surgery* may improve diplopia due to muscle changes, but should be deferred until the situation has been stable for 6 months and should follow any orbital decompression. Plastic surgery around the eyes may also be of value.

GOITRE (THYROID ENLARGEMENT)

Goitre is more common in women than in men and may be either physiological or pathological.

Clinical features

Goitres are present on examination in up to 9% of the population. Most commonly a goitre is noticed as a cosmetic defect by the patient or by friends or relatives. The majority are painless, but pain or discomfort can occur in acute varieties. Large goitres can produce dysphagia and difficulty in breathing, implying oesophageal or tracheal compression.

A small goitre may be more easily visible (on swallowing) than palpable. Clinical examination should record the size, shape, consistency and mobility of the gland as well as whether its lower margin can be demarcated (thus implying the absence of retrosternal extension). A bruit may be present. Associated lymph nodes should be sought and the tracheal position determined if possible. Examination should never omit an assessment of the patient's clinical thyroid status.

Specific enquiry should be made about any medication, especially iodine-containing preparations, and possible exposure to radiation.

Particular points of note are:

- Puberty and pregnancy may produce a diffuse increase in size of the thyroid.
- Pain in a goitre may be caused by thyroiditis, bleeding into a cyst or (rarely) a thyroid tumour.
- Excessive doses of carbimazole or propylthiouracil will induce goitre.
- Iodine deficiency and dyshormonogenesis (see above) can also cause goitre.

Assessment

There are two major aspects of any goitre: its pathological nature and the patient's thyroid status.

The nature can often be judged clinically. Goitres (Table 18.14) are usually separable into diffuse and nodular types, the causes of which differ.

Table 18.14 Goitre – causes and types

Diffuse
Simple
Physiological (puberty, pregnancy)
Autoimmune
Graves' disease
Hashimoto's disease
Thyroiditis
Acute (de Quervain's thyroiditis)
Iodine deficiency (endemic goitre)
Dyshormonogenesis
Goitrogens (e.g. sulfonylureas)
Nodular
Multinodular goitre
Solitary nodular
Fibrotic (Reidel's thyroiditis)
Cysts
Tumours
Adenomas
Carcinoma
Lymphomas
Miscellaneous
Sarcoidosis
Tuberculosis

Diffuse goitre

Simple goitre

In this instance no clear cause is found for enlargement of the thyroid, which is usually smooth and soft. It may be associated with thyroid growth-stimulating antibodies.

Autoimmune thyroid disease

Hashimoto's thyroiditis and thyrotoxicosis are both associated with firm diffuse goitre of variable size. A bruit is often present in thyrotoxicosis.

Thyroiditis

Acute tenderness in a diffuse swelling, sometimes with severe pain, is suggestive of an acute viral thyroiditis (de Quervain's). It may produce transient clinical hyperthyroidism with an increase in serum T_4 (see p. 968).

Nodular goitres

Multinodular goitre

Most common is the multinodular goitre, especially in older patients. The patient is usually euthyroid but may be hyperthyroid or borderline with suppressed TSH levels but normal T_4 and T_3. Multinodular goitre is the most common cause of tracheal and/or oesophageal compression and can cause laryngeal nerve palsy. It may also extend retrosternally. The classical 'multinodular goitre' is usually readily apparent clinically, but it should be noted that modern, high-resolution ultrasound frequently reports multiple small nodules in glands which are clinically diffusely enlarged and associated with autoimmune thyroid disease. These nodules are also found in up to 40% of the normal population.

Solitary nodular goitre

Such a goitre presents a difficult problem of diagnosis. Malignancy should be considered in any solitary nodule –

however, the majority of such nodules are cystic or benign and, indeed, may simply be the largest nodule of a multinodular goitre. The diagnostic challenge is to identify the small minority of malignant nodules, which require surgery, from the majority of benign nodules, which do not. A history of rapid enlargement, associated lymph nodes or occasionally pain in such a situation suggests the possibility of thyroid carcinoma, but investigations are paramount. Risk factors for malignancy include previous irradiation, long-standing iodine deficiency and occasional familial cases.

Solitary toxic nodules are quite uncommon and may be associated with T_3 toxicosis.

Fibrotic goitre

Fibrotic goitre (Riedel's thyroiditis) is a rare condition, usually producing a 'woody' gland. It is associated with other midline fibrosis and is often difficult to distinguish from carcinoma, being irregular and hard. Clinical clues include systemic symptoms of inflammation and elevation in inflammatory markers.

Malignancy

In addition to thyroid carcinomas (see below), the thyroid is rarely the site of a metastatic deposit or the site of origin of a lymphoma.

Investigations

Clinical findings will dictate appropriate initial tests:

- **Thyroid function tests** – TSH plus free T_4 or T_3 (see Table 18.10).
- **Thyroid antibodies** – to exclude autoimmune aetiology.
- **Ultrasound**. Ultrasound with high resolution is a sensitive method for delineating nodules and can demonstrate whether they are cystic or solid. In addition, a multinodular goitre may be demonstrated when only a single nodule is palpable. Unfortunately, even cystic lesions can be malignant and thyroid tumours may arise within a multinodular goitre; therefore fine-needle aspiration (see below) is often required and performed under ultrasound control at the same time as the scan.
- **Chest and thoracic inlet X-rays** to detect tracheal compression and large retrosternal extensions in patients with very large goitre or clinical symptoms.
- **Fine-needle aspiration (FNA)**. In patients with a solitary nodule or a dominant nodule in a multinodular goitre, there is a 5% chance of malignancy; in view of this, FNA should be performed. This can be done in the outpatient clinic. Cytology in expert hands can usually differentiate the suspicious or definitely malignant nodule.

  FNA reduces the necessity for surgery, but there is a 5% false-negative rate which must be borne in mind (and the patient appropriately counselled). Continued observation is required when an isolated thyroid nodule is assumed to be benign without excision.
- **Thyroid scan** (^{125}I or ^{131}I) can be useful to distinguish between functioning (hot) or non-functioning (cold) nodules. A hot nodule is only rarely malignant; however, a cold nodule is malignant in only 10% of cases and FNA has largely replaced isotope scans in the diagnosis of thyroid nodules.

Treatment

Euthyroid goitre

Many goitres are small, cause no symptoms and can be observed (including self-monitoring by the patient in the long term). In particular, during puberty and pregnancy a goitre associated with euthyroidism rarely requires intervention and the patient can be reassured that spontaneous resolution is likely. When thyroid function is abnormal the patient should be rendered euthyroid. Indications for surgical intervention are:

- *The possibility of malignancy*. A history of rapid growth, pain, cervical lymphadenopathy, change in voice or previous irradiation to the neck are worrying features. A positive or suspicious FNA makes surgery mandatory and surgery may be necessary if doubt persists even in the presence of a negative FNA (especially if the patient is concerned by the false negative rate).
- *Pressure symptoms on the trachea or, more rarely, oesophagus*. The possibility of retrosternal extension should be excluded.
- *Cosmetic reasons*. A large goitre is often a considerable anxiety to the patient even though functionally and anatomically benign.

Recent guidelines suggest possible use of RAI in euthyroid goiter, particularly when surgery is an unattractive option.

Toxic nodule

This is initially with antithyroid drugs but surgery or radioiodine is often required.

THYROID CARCINOMA

Types of thyroid carcinoma, their characteristics and treatment are listed in Table 18.15. While not common, these tumours are responsible for 400 deaths annually in the UK and an annual incidence of 30 000 cases in the USA. Over 75% occur in women. In 90% of cases they present as thyroid nodules (see above), but occasionally with cervical lymphadenopathy (about 5%), or with lung, cerebral, hepatic or bone metastases.

Table 18.15 Types of thyroid malignancy

Cell type	Frequency	Behaviour	Spread	Prognosis
Papillary	70%	Occurs in young people	Local, sometimes lung/bone secondaries	Good, especially in young
Follicular	20%	More common in females	Metastases to lung/bone	Good if resectable
Anaplastic	<5%	Aggressive	Locally invasive	Very poor
Lymphoma	2%	Variable		Sometimes responsive to radiotherapy
Medullary cell	5%	Often familial	Local and metastases	Poor, but indolent course

Carcinomas derived from thyroid epithelium may be papillary or follicular (differentiated) or anaplastic (undifferentiated), while medullary carcinomas (about 5% of all thyroid cancers) arise from the calcitonin-producing C cells. The pathogenesis of thyroid epithelial carcinomas is not understood except for occasional familial papillary carcinoma and those cases related to previous head and neck irradiation or ingestion of radioactive iodine (e.g. post Chernobyl). These tumours are minimally active hormonally and are extremely rarely associated with hyperthyroidism; over 90%, however, secrete thyroglobulin, which can therefore act as a tumour marker.

Papillary and follicular carcinomas

The primary treatment is surgical, normally total or near-total thyroidectomy for local disease. Regional or more extensive neck dissection is needed where there is local nodal spread or involvement of local structures.

Most tumours will take up iodine, and UK guidelines currently recommend radioactive iodine (RAI) ablation of residual thyroid tissue post-operatively for most patients with differentiated thyroid cancer. After ablation of normal thyroid in this way, RAI may be used to localize residual disease (scanning using low doses) or to treat it (using high doses – 5.5–7.5 GBq). When recurrence does occur, local invasion and lymph node involvement is most common, and lungs and bone are the most common sites of distant metastases.

Patients are treated with suppressive doses of levothyroxine (sufficient to suppress TSH levels below the normal range) in order to minimize risk of recurrence. Patient progress is monitored both clinically and biochemically using serum thyroglobulin levels as a tumour marker. The measurement of thyroglobulin is most sensitive when TSH is high but this requires the withdrawal of levothyroxine therapy. Recombinant TSH (thyrotropin alfa, rhTSH) 900 μg (2 doses over 48 hours) is used to stimulate thyroglobulin without stopping thyroxine therapy. Detectable thyroglobulin suggests recurrence, in which case whole body ^{131}I scanning is required.

The prognosis is extremely good when these types of tumour are excised while confined to the thyroid gland, and the specific therapies available lead to a relatively good prognosis even in the presence of metastases at diagnosis. Accepted markers of high risk include greater age (>40 years), larger primary tumour size (>4 cm) and macroscopic invasion of capsule and surrounding tissues.

Anaplastic carcinomas and lymphoma

These do not respond to radioactive iodine, and external radiotherapy produces only a brief respite.

Medullary carcinoma

Medullary carcinoma (MTC) is a neuroendocrine tumour of the calcitonin-producing C cells of the thyroid. This condition is often associated with multiple endocrine neoplasia type 2 (MEN 2, see p. 1026), although the other manifestations of MEN 2 may be absent. Approximately 25% of patients diagnosed with MTC have a mutation of the *RET* proto-oncogene, hence the importance of genetic counselling and family screening. Total thyroidectomy and wide lymph node clearance is usually indicated in MTC. Local invasion or metastasis is frequent, and the tumour responds poorly to treatment, although progression is often slow.

FURTHER READING

Brent GA. Graves' disease. *New England Journal of Medicine* 2008; **358**: 2594–2605.

Hegedus L. *The thyroid nodule. New England Journal of Medicine* 2004; **351**: 1764–1771.

Report of a Working Party. Radioiodine in the management of benign thyroid disease: Clinical guidelines. Royal College of Physicians (London), 2007. Available at http://www.rcplondon.ac.uk/pubs/

Roberts GGP. Hypothyroidism. *Lancet* 2004; **36**: 793–798.

Sherman SI. Thyroid carcinoma. *Lancet* 2003; **361**: 501–511.

Zimmermann MB, Jooste PL, Pandar CS. Iodine deficiency disorders. *Lancet* 2008; **372**: 1251–1262.

REPRODUCTION AND SEX

The normal physiology of the female and male reproductive systems will be considered first, followed by their common disorders. Some relevant terminology is set out in Box 18.4.

Embryology

Up to 8 weeks of gestation the sexes share a common development, with a primitive genital tract including the Wolffian and Müllerian ducts. There are additionally a primitive perineum and primitive gonads.

- In the presence of a Y chromosome, the potential testis develops while the ovary regresses.
- In the absence of a Y chromosome, the potential ovary develops and related ducts form a uterus and the upper vagina.

Production of Müllerian inhibitory factor from the early 'testis' produces atrophy of the Müllerian duct, while, under the influence of testosterone and dihydrotestosterone, the Wolffian duct differentiates into an epididymis, vas deferens, seminal vesicles and prostate. Androgens induce transformation of the perineum to include a penis, penile urethra and scrotum containing the testes, which descend in response to androgenic stimulation. At birth, testicular volume is 0.5–1 mL.

Physiology

The male

An outline of the hypothalamic–pituitary–testicular axis is shown in Figure 18.17.

1. Pulses of gonadotrophin releasing hormone (GnRH) are released from the hypothalamus and stimulate LH and FSH release from the pituitary. LH and FSH are composed of two glycoprotein chains (α and β subunits). The α subunits are identical and are shared with TSH, whilst the β subunit confers specific biological activity.
2. LH stimulates testosterone production from Leydig cells of the testis.
3. Testosterone acts via nuclear androgen receptors which interact with co-regulatory proteins to produce the appropriate tissue responses – male secondary sexual characteristics, anabolism and the maintenance of

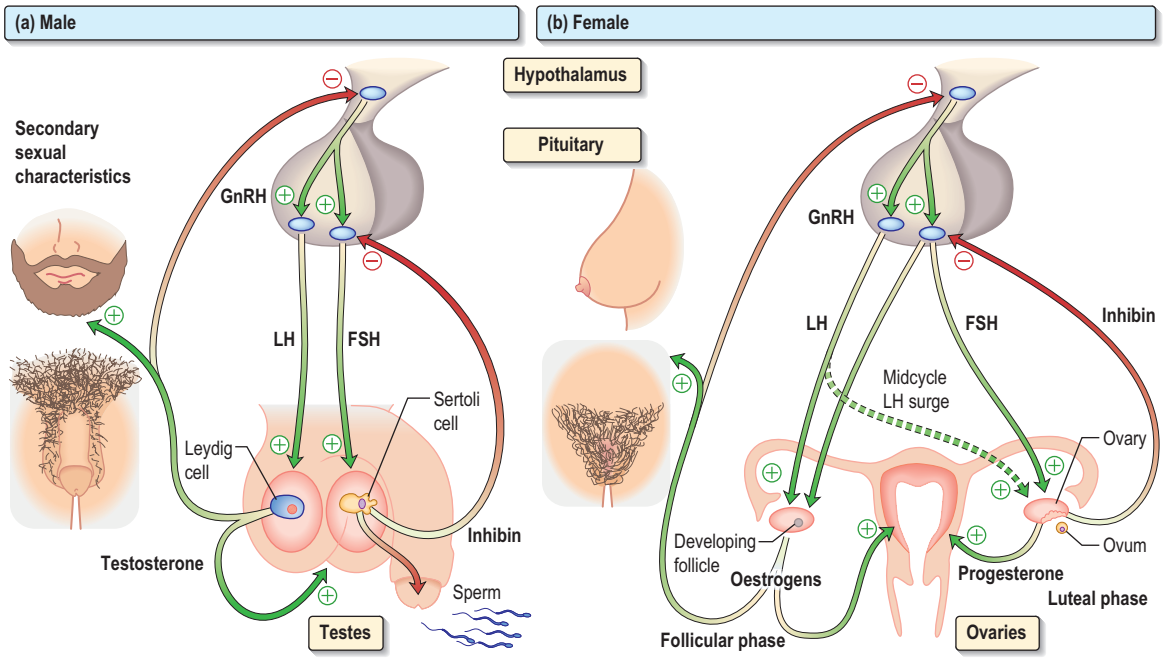


Fig. 18.17 **Male and female hypothalamic–pituitary–gonadal axes.** LH and FSH are secreted in pulses in response to hypothalamic GnRH and have differential effects on the gonads (see text). Testosterone and oestrogens have multiple peripheral effects and exert negative feedback on LH secretion. Inhibin exerts negative feedback on FSH secretion.

Box 18.4 Definitions in reproductive medicine

Menarche	Age at first period
Primary amenorrhoea	Failure to begin spontaneous menstruation by age 16
Secondary amenorrhoea	Absence of menstruation for 3 months in a woman who has previously had cycles
Oligomenorrhoea	Irregular long cycles; often used for any length of cycle above 32 days
Dyspareunia	Pain or discomfort in the female during intercourse
Libido	Sexual interest or desire; often difficult to assess and is greatly affected by stress, tiredness and psychological factors
Menstruation	Onset of spontaneous (usually regular) uterine bleeding in the female
Erectile dysfunction	Inability of the male to achieve or sustain an erection adequate for satisfactory intercourse
Azoospermia	Absence of sperm in the ejaculate
Oligospermia	Reduced numbers of sperm in the ejaculate; normal values are disputed
Virilization	Occurrence of male secondary sexual characteristics in the female

libido. It also acts locally within the testis to aid spermatogenesis. Testosterone circulates largely bound to sex hormone-binding globulin (SHBG) (see p. 966). Testosterone feeds back on the hypothalamus/pituitary to inhibit GnRH secretion.

4. FSH stimulates the Sertoli cells in the seminiferous tubules to produce mature sperm and the inhibins A and B.
5. Inhibin feeds back to the pituitary to decrease FSH secretion. Activin, a related peptide, counteracts inhibin.

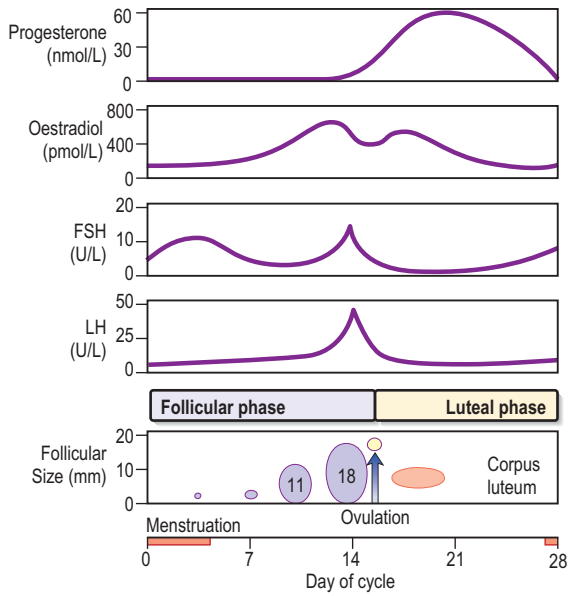


Fig. 18.18 **Hormonal and follicular changes during the normal menstrual cycle.**

The secondary sexual characteristics of the male for which testosterone is necessary are the growth of pubic, axillary and facial hair, enlargement of the external genitalia, deepening of the voice, sebum secretion, muscle growth and frontal balding.

The female

The female situation is more complex (Figs 18.17 and 18.18).

1. In the adult female, higher brain centres impose a menstrual cycle of 28 days upon the activity of hypothalamic GnRH.
2. Pulses of GnRH, at about 2-hour intervals, stimulate release of pituitary LH and FSH.

3. LH stimulates ovarian androgen production by the ovarian theca cells.
4. FSH stimulates follicular development and aromatase activity (an enzyme required to convert ovarian androgens to oestrogens) in the ovarian granulosa cells. FSH also stimulates release of inhibin from ovarian stromal cells, which inhibits FSH release. Activin counteracts inhibin (Fig. 18.17).
5. Although many follicles are 'recruited' for development in early folliculogenesis, by day 8–10 a 'leading' (or 'dominant') follicle is selected for development into a mature Graafian follicle.
6. Oestrogens show a double feedback action on the pituitary (Fig. 18.17), initially inhibiting gonadotrophin secretion (negative feedback), but later high-level exposure results in increased GnRH secretion and increased LH sensitivity to GnRH (positive feedback), which leads to the mid-cycle LH surge inducing ovulation from the leading follicle (Fig. 18.18).
7. The follicle then differentiates into a corpus luteum, which secretes both progesterone and oestradiol during the second half of the cycle (luteal phase).
8. Oestrogen initially and then progesterone cause uterine endometrial proliferation in preparation for possible implantation; if implantation does not occur, the corpus luteum regresses and progesterone secretion and inhibin levels fall so that the endometrium is shed (menstruation) allowing increased GnRH and FSH secretion.
9. If implantation and pregnancy follow, human chorionic gonadotrophin (HCG) production from the trophoblast maintains corpus luteum function until 10–12 weeks of gestation, by which time the placenta will be making sufficient oestrogen and progesterone to support itself.

Oestrogens also induce secondary sexual characteristics, especially development of the breast and nipples, vaginal and vulval growth and pubic hair development. They also induce growth and maturation of the uterus and Fallopian tubes. They circulate largely bound to SHBG.

Puberty

The mechanisms initiating puberty are poorly understood but are thought to result from withdrawal of central inhibition of GnRH release. Environmental and physical factors are involved in the timing of puberty (including body fat changes, physical exercise) as well as genetic factors (e.g. a G protein-coupled receptor gene, *GPR54*) required for pubertal maturation. Kisspeptin is the endogenous ligand for *GPR54* and this peptide is believed to play a crucial role in the regulation of GnRH production and the timing of puberty.

LH and FSH are both low in the prepubertal child. In early puberty, FSH begins to rise first, initially in nocturnal pulses; this is followed by a rise in LH with a subsequent increase in testosterone/oestrogen levels. The milestones of puberty in the two sexes are shown in Figure 18.19.

In boys, pubertal changes begin at between 10 and 14 years and are complete at between 15 and 17 years. The genitalia develop, testes enlarge and the area of pubic hair increases. Peak height velocity is reached between ages 12 and 17 years during stage 4 of testicular development. Full spermatogenesis occurs comparatively late.

In girls, events start a year earlier. Breast bud enlargement begins at age 9–13 years and continues to 12–18 years.

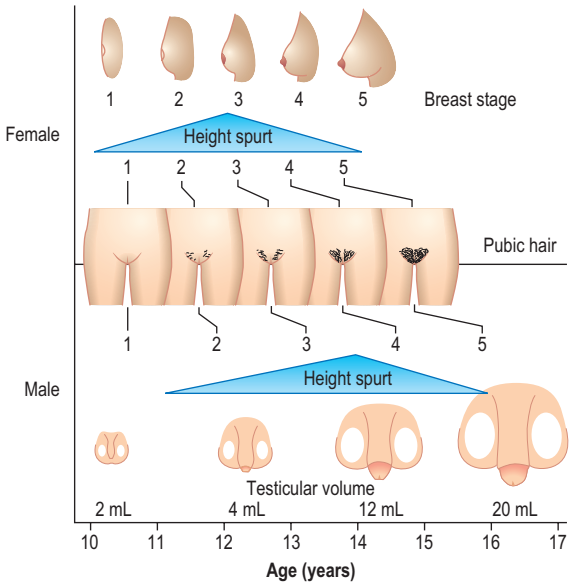


Fig. 18.19 **The age of development of features of puberty.** Stages and testicular size show mean ages, and all vary considerably between individuals. The same is true of height spurt, shown here in relation to other data. Numbers 2 to 5 indicate stages of development (see the text).

Pubic hair growth commences at ages 9–14 years and is completed at 12–16 years. Menarche occurs relatively late (age 11–15 years) but peak height velocity is reached earlier (at age 10–13 years), and growth is completed much earlier than in boys.

Precocious puberty

Development of secondary sexual characteristics, or menarche in girls, at or before the age of 9 years is premature. All cases require assessment by a paediatric endocrinologist.

Idiopathic (true) precocity is most common in girls and very rare in boys. This is a diagnosis of exclusion. With no apparent cause for premature breast or pubic hair development, and an early growth spurt, it may be normal and may run in families. Treatment with long-acting GnRH analogues (given by nasal spray, by subcutaneous injection or by implant) causes suppression of gonadotrophin release via downregulation of the receptor – and therefore reduced sex hormone production – and is moderately effective; cyproterone acetate, an antiandrogen with progestational activity, is also used.

Other forms of precocity include:

- *Cerebral precocity*. Many causes of hypothalamic disease, especially tumours, present in this way. In boys this must be rigorously excluded. MRI scan is almost always indicated to exclude this diagnosis.
- *McCune–Albright syndrome*. This usually occurs in girls, with precocity, polyostotic fibrous dysplasia and skin pigmentation (café-au-lait). See also page 1026.
- *Premature thelarche*. This is early breast development alone, usually transient, at age 2–4 years. It may regress or persist until puberty. There is no evidence of follicular development.
- *Premature adrenarche*. This is early development of pubic hair without significant other changes, usually after the age of 5 years and more commonly in girls.

FURTHER READING

Carel J-C, Léger J. Precocious puberty. *New England Journal of Medicine* 2008; **358**: 2366–2377.

It is also more common in obese children due to reduced SHBG levels leading to higher free circulating androgens. In boys with precocious adrenarche, the rare possibility of an androgen-secreting testicular tumour should be looked for if serum androgens are high and LH is suppressed.

Delayed puberty

Over 95% of children show signs of pubertal development by the age of 14 years. In its absence, investigation should begin by the age of 15 years. Causes of hypogonadism (see below) are clearly relevant but most cases represent constitutional delay.

In constitutional delay, pubertal development, bone age and stature are in parallel. A family history may confirm that other family members experienced the same delayed development, which is common in boys but very rare in girls.

In boys, a testicular volume >5 mL indicates the onset of puberty. A rising serum testosterone is an earlier clue.

In girls, the breast bud is the first sign. Ultrasound allows accurate assessment of ovarian and uterine development.

Basal LH/FSH levels may identify the site of a defect, and GnRH (LHRH) tests can indicate the stage of early puberty.

If any progression into puberty is evident clinically, investigations are not required. When delay is great and problems are serious (e.g. severe teasing at school), low-dose short-term sex hormone therapy is used.

The menopause

The menopause, or cessation of periods, naturally occurs at about the age of 45–55 years. During the late forties, FSH initially, and then LH concentrations begin to rise, probably as follicle supply diminishes. Oestrogen levels fall and the cycle becomes disrupted. Most women notice irregular scanty periods coming on over a variable period, though in some sudden amenorrhoea or menorrhagia occurs. Eventually the menopausal pattern of low oestradiol levels with grossly elevated LH and FSH levels (usually >50 and >25 U/L, respectively) is established. Premature menopause may also occur surgically, with radiotherapy to the ovaries and with ovarian disease.

Clinical features and treatment

Features of oestrogen deficiency are hot flushes (which occur in most women and can be disabling), vaginal dryness and atrophy of the breasts. There may also be vague symptoms of loss of libido, loss of self-esteem, non-specific aches and pains, irritability, depression, loss of concentration and weight gain.

Women show a rapid loss of bone density in the 10 years following the menopause (osteoporosis, see p. 561) and the premenopausal protection from ischaemic heart disease disappears.

Hormone replacement therapy (HRT). Symptomatic patients should usually be treated but the previous widespread use of HRT has been thrown into doubt by a number of large prospective studies which have reported in recent years. Although there is no universal agreement, the overall benefits and risks are summarized as follows (percentages are from the Women's Health Initiative (WHI) study of 16600 women):

- *Symptomatic improvement in most menopausal symptoms* for the majority of women. Oestrogen-deficient symptoms respond well to oestrogen replacement, the vaguer symptoms generally, but not always, less well. Vaginal symptoms also respond to local oestrogen preparations.
- *Protection against fractures of wrist, spine and hip, secondary to osteoporosis* (~24–33%) (see Ch. 10), owing to protection of predominantly trabecular bone (p. 557). However, this is not a recommended indication for therapy.
- *A significant reduction in the risk of large bowel cancer* (~33%).
- *A significant increase in the risk of breast cancer* (+26%), but no change in breast cancer mortality, and some studies suggest that breast cancers diagnosed on HRT are easier to treat effectively. This increased risk has been disputed.
- *A significant increase in the risk of endometrial cancer* when unopposed oestrogens are given to women with a uterus.
- *A significant increase in the risk of ischaemic heart disease (+29%) and stroke (+41%).*
- The inconvenience of withdrawal bleeds, unless a hysterectomy has been performed or regimens which include continuous oestrogen and progesterone are used.
- *Other disputed effects* include possible increase in general well-being but there is no reduction in the incidence of Alzheimer's disease.

Absolute risks and benefits for individual women clearly depend on their background risk of that disease, and there is as yet no evidence on the relative risks of different hormone preparations or routes of administration (oral, transdermal or implant). Overall, the WHI study estimated that, over 5 years of treatment, an extra 1 woman in every 100 would develop an illness that would not have occurred had she not been taking HRT. However, the decision about whether or not a woman takes HRT is now very much an individual decision based on the severity of that woman's menopausal symptoms, her personal risk of conditions which may be prevented or made more likely by HRT, and ultimately individual patient choice. HRT is not recommended purely for prevention of postmenopausal osteoporosis in the absence of menopausal symptoms.

Symptomatic treatment is the main indication with the lowest effective dose given for short-term rather than long-term treatment.

Selective oestrogen receptor modulators, SERMs (e.g. raloxifene), offer a potentially attractive combination of positive oestrogen effects on bone and cardiovascular system with no effects on oestrogen receptors of uterus and breast and possible reduction in breast cancer incidence; long-term outcome studies, however, are still awaited.

Premature menopause

The most common cause of premature menopause in women (before age 40) is ovarian failure, which may be autoimmune. It is rarely caused by identifiable genetic causes such as the fragile X pre-mutation but is most commonly of unknown aetiology (although often familial). Repeat LH/FSH levels are necessary before giving a diagnosis of premature menopause because of the psychological impact of this diagnosis

and the possibility that a single elevation of LH/FSH might simply be the mid-cycle ovulatory surge. Bilateral oophorectomy causes the same oestrogen deficiency state. HRT should almost always be given, as the risk of osteoporosis and other conditions related to oestrogen deficiency almost always outweighs the risks of HRT at this younger age. HRT may still also be actively recommended when normal menopause occurs relatively early (e.g. before the age of 50).

The ageing male

In the male there is no sudden 'change of life'. However, there is a progressive loss in sexual function with reduction in morning erections and frequency of intercourse.

The age of onset varies widely, but overall testicular volume diminishes and sex hormone-binding globulin (SHBG) and gonadotrophin levels gradually rise. If premature hypogonadism is present for any reason, replacement testosterone therapy should be given to prevent osteoporosis (see p. 565). More widespread androgen replacement of the ageing male is under evaluation, primarily in the USA, with variable results in earlier, small studies, but no long-term follow-up data are available. However, the increasing choice of preparations for testosterone replacement largely reflects the interest of the pharmaceutical industry in this situation.

Conversely, lowering of androgens forms part of the therapy of prostate hypertrophy and prostate cancer. Finasteride, an inhibitor of 5α-reductase, is used in benign prostatic hypertrophy. It prevents the conversion of testosterone to dihydrotestosterone, which causes prostatic hypertrophy, and is effective, though somewhat delayed in action (see p. 645). GnRH analogues are used to lower testosterone levels and induce disease remission in patients with prostate cancer (p. 492).

CLINICAL FEATURES OF DISORDERS OF SEX AND REPRODUCTION

A detailed history and examination of all systems is required (Box 18.5). A man having regular satisfactory intercourse or a woman with regular ovulatory periods is most unlikely to have significant endocrine disease, assuming the history is accurate (check with the partner!).

Box 18.5 Sexual and menstrual disorders

History

Menstruation – timing of bleeding and cycle
Relationship of symptoms to cycle
Breasts (tenderness/galactorrhoea)
Hirsutism and acne
Libido and potency
Problems with intercourse
Past fertility and future plans

Physical signs

Evidence of systemic disease
Secondary sexual characteristics
Extent/distribution of hair
Genital size (testes, ovaries, uterus)
Clitoromegaly
Breast development, gynaecomastia
Galactorrhoea

Tests of gonadal function

Basal measurements of the gonadotrophins, oestrogens/testosterone and prolactin:

- Low testosterone or oestradiol with high gonadotrophins indicates primary gonadal disease.
- Low levels of testosterone/oestradiol with low or normal LH/FSH imply hypothalamic–pituitary disease.
- Demonstration of ovulation (by measurement of luteal phase serum progesterone and/or by serial ovarian ultrasound in the follicular phase) or a healthy sperm count (20–200 million/mL, >60% grade I motility and <20% abnormal forms), provide absolute confirmation of normal female or male reproductive endocrinology, but these tests are not always essential.
- Pregnancy provides complete demonstration of normal male and female function.
- Hyperprolactinaemia can be confirmed or excluded by direct measurement. Levels may increase with stress; if this is suspected, a cannula should be inserted and samples taken through it 30 minutes later.

More detailed tests are indicated in Table 18.16.

DISORDERS IN THE MALE

Hypogonadism

Clinical features

Male hypogonadism may be a presenting complaint or an incidental finding, such as during investigation for subfertility. The testes may be small and soft and there may be gynae-

Table 18.16 Tests of gonadal function

Test	Uses/comments
Male	
Basal testosterone	Normal levels exclude hypogonadism
Sperm count	Normal count excludes deficiency Motility and abnormal sperm forms should be noted
Female	
Basal oestradiol	Normal levels exclude hypogonadism
Luteal phase progesterone (days 18–24 of cycle)	If >30 nmol/L, suggests ovulation
Ultrasound of ovaries	To confirm ovulation
Both sexes	
Basal LH/FSH	Demonstrates state of feedback system for hormone production (LH) and germ cell production (FSH)
HCG test (testosterone or oestradiol measured)	Response shows potential of ovary or testis; failure demonstrates primary gonadal problem
Clomifene test (LH and FSH measured)	Tests hypothalamic negative feedback system; clomifene is oestrogen antagonist and causes LH/FSH to rise
LHRH test (rarely used)	Shows adequacy (or otherwise) of LH and FSH stores in pituitary

comastia. Except with subfertility, the symptoms are usually of androgen deficiency, primarily poor libido, erectile dysfunction and loss of secondary sexual hair (Table 18.17) rather than deficiency of semen production. Sperm makes up only a very small proportion of seminal fluid volume; most is prostatic fluid.

Causes of male hypogonadism are shown in Table 18.18.

Investigations

Testicular disease may be immediately apparent but basal levels of testosterone, LH and FSH should be measured. These will allow the distinction between primary gonadal (testicular) failure and hypothalamic–pituitary disease to be made. Depending on the causes, semen analysis, chromosomal analysis (e.g. to exclude Klinefelter's syndrome) and bone age estimation are required.

In clear-cut gonadotrophin deficiency, pituitary MRI scan, prolactin levels and other pituitary function tests are needed.

However, equivocal lowering of serum testosterone (7–10 nmol/L) without elevation of gonadotrophins is a relatively common biochemical finding, and is a frequent cause of referral in men with poor libido or erectile dysfunction. Such tests are compatible with mild gonadotrophin deficiency, but may also be seen in acute illness of any cause and often simply represent the lower end of the normal range or the normal circadian rhythm of testosterone when bloods are checked in afternoon or evening surgeries. A therapeutic trial of testosterone replacement is often justified and forms part of the investigation in many patients; full pituitary evaluation may be required in such cases to exclude other pituitary disease. 'Anabolic' steroid (i.e. androgen) abuse causes similar biochemical findings, and is likely if the patient appears well virilized.

Treatment

The cause can rarely be reversed. Replacement therapy should be commenced (Table 18.19). Primary gonadal failure should be treated with testosterone. Patients with hypothalamic–pituitary disease are routinely given testosterone replacement, but should be given LH and FSH (purified or synthetic) or pulsatile GnRH when fertility is required.

Table 18.17 Effects of androgens and consequences of androgen deficiency in the male

Physiological effect	Consequences of deficiency
General	
Maintenance of libido	Loss of libido
Deepening of voice	High-pitched voice (if prepubertal)
Frontotemporal balding	No temporal recession
Facial, axillary and limb hair	Decreased hair
Maintenance of erectile and ejaculatory function	Loss of erections/ejaculation
Pubic hair	
Maintenance of male pattern	Thinning and loss of pubic hair
Testes and scrotum	
Maintenance of testicular size/consistency (needs gonadotrophins as well)	Small soft testes
Rugosity of scrotum	Poorly developed penis/scrotum
Stimulation of spermatogenesis	Subfertility
Musculoskeletal	
Epiphyseal fusion	Eunuchoidism (if prepubertal)
Maintenance of muscle bulk and power	Decreased muscle bulk
Maintenance of bone mass	Osteoporosis

Table 18.18 Causes of male hypogonadism

Reduced gonadotrophins (hypothalamic–pituitary disease)
Hypopituitarism
Selective gonadotrophic deficiency (Kallmann's syndrome), normosmic idiopathic hypogonadotropic hypogonadism
Severe systemic illness
Severe underweight
Hyperprolactinaemia
Primary gonadal disease (congenital)
Anorchia/Leydig cell agenesis
Cryptorchidism (testicular maldescent)
Chromosome abnormality (e.g. Klinefelter's syndrome)
Enzyme defects: 5α-reductase deficiency
Primary gonadal disease (acquired)
Testicular torsion
Orchidectomy
Local testicular disease
Chemotherapy/radiation toxicity
Orchitis (e.g. mumps)
Chronic kidney disease
Cirrhosis/alcohol
Sickle cell disease
Androgen receptor deficiency/abnormality

Table 18.19 Androgen replacement therapy

Preparation	Dose	Route	Remarks
Testosterone mixed esters	250 mg × every 3 weeks	i.m	Usual first-line maintenance therapy
Testosterone enanthate			Injection can be painful
Testosterone propionate	50–100 mg 2–3 × weekly	i.m.	Frequent injections needed as half-life is short Good initial therapy
Testosterone undecanoate	80–240 mg daily, in divided doses	Oral	Variable dose, irregular absorption
	1000 mg every 3 months	i.m.	4 mL injection
Testosterone implant	600 mg every 4–5 months	Implant	Requires implant procedure
Testosterone transdermal	2.5–7.5 mg/24 h	Dermal	Patch or gel preparations

Special instances of hypogonadism

Cryptorchidism

By the age of 5 years both testes should be in the scrotum. After that age, the germinal epithelium is increasingly at risk, and lack of descent by puberty is associated with subfertility. Surgical exploration and orchidopexy are usually undertaken but a short trial of HCG occasionally induces descent: an HCG test with a testosterone response 72 hours later excludes anorchia. Intra-abdominal testes have an increased risk of developing malignancy; if presentation is after puberty, orchidectomy is advised.

Klinefelter's syndrome

Klinefelter's syndrome is a common congenital abnormality, affecting 1 in 1000 males. It is a chromosomal disorder (47XXY and variants, e.g. 46XY/47XXY mosaicism), i.e. a male with an extra X chromosome. There is both a loss of Leydig cells and seminiferous tubular dysgenesis. Patients usually present in adolescence with poor sexual development, small or undescended testes, gynaecomastia or infertility. In 47XXY there is long leg length with tall stature. The androgen deficiency leads to lack of epiphyseal closure in puberty. Patients occasionally have behavioural problems and learning difficulties. There is also a predisposition to diabetes mellitus, breast cancer, emphysema and bronchiectasis; these are all unrelated to the testosterone deficiency.

Clinical examination shows a wide spectrum of features with small pea-sized but firm testes, usually gynaecomastia and other signs of *androgen deficiency*. Some patients have a normal puberty and may present later with infertility. Confirmation is by chromosomal analysis. Treatment is androgen replacement therapy unless testosterone levels are normal. No treatment is possible for the abnormal seminiferous tubules and infertility.

Kallmann's syndrome

This is isolated GnRH deficiency. It is associated with decreased or absent sense of smell (anosmia), and sometimes with other bony (cleft palate), renal and cerebral abnormalities (e.g. colour blindness). It is often familial and is usually X-linked, resulting from a mutation in the *KAL1* gene which encodes anosmin-1 (producing loss of smell); one sex-linked form is due to an abnormality of a cell adhesion molecule. Management is that of secondary hypogonadism (see p. 999). Fertility is possible.

Normosmic idiopathic hypogonadotropic hypogonadism

This refers to isolated GnRH deficiency in the absence of anosmia. Known mutations account for less than 15% of normosmic idiopathic hypogonadotropic hypogonadism (nIHH). Mutations include the *KISS1* gene which codes for kisspeptin, the protein which acts on GPR54 receptor, and the *FGFR1* gene.

Oligospermia or azoospermia

These may be secondary to androgen deficiency and can be corrected by androgen replacement. More often they result from primary testicular diseases, in which case they are rarely treatable.

Azoospermia with normal testicular size and low FSH levels suggests a vas deferens block, which is sometimes reversible by surgical intervention.

Lack of libido and erectile dysfunction

Lack of libido is a loss of sexual desire which may in turn lead to erectile dysfunction (ED). ED may be psychological, neurogenic, vascular, endocrine or related to drugs, and often includes contributions from several causes. ED is a common symptom in hypogonadism, but most patients with ED have normal hormones and many have no definable organic cause. Vascular disease is a common aetiology, especially in smokers, and is often associated with vascular problems elsewhere. Autonomic neuropathy, most commonly from diabetes mellitus, is a common partial, if not total, identifiable cause (see p. 1056). Many drugs produce ED (see Table 22.28). The endocrine causes are those of hypogonadism (see above) and can be excluded by normal testosterone, gonadotrophin and prolactin levels. The presence of nocturnal emissions and frequent satisfactory morning erections makes endocrine disease unlikely.

Psychogenic erectile dysfunction is frequently a diagnosis of exclusion, though complex tests of penile vasculature and function are available in some centres.

Treatment. Offending drugs should be stopped. Phosphodiesterase type-5 inhibitors (sildenafil, tadalafil, vardenafil) which increase penile blood flow (see p. 1056) are first choice for therapy. Other treatments include apomorphine, intracavernosal injections of alprostadil, papaverine or phentolamine, vacuum expanders and penile implants.

If no organic disease is found, or if there is clear evidence of psychological problems, the couple should receive psychosexual counselling.

Gynaecomastia

Gynaecomastia is development of breast tissue in the male. Causes are shown in Table 18.20. It is due to an imbalance

Table 18.20 Causes of gynaecomastia

- Physiological
  - Neonatal
  - Pubertal
  - Old age
- Hyperthyroidism
- Hyperprolactinaemia
- Renal disease
- Liver disease
- Hypogonadism (see Table 18.18)
- Oestrogen-producing tumours (testis, adrenal)
- HCG-producing tumours (testis, lung)
- Starvation/refeeding
- Carcinoma of breast

Drugs

- Oestrogenic:
  - oestrogens
  - digoxin
  - cannabis
  - diamorphine
- Anti-androgens:
  - spironolactone
  - cimetidine
  - cyproterone
- Others:
  - gonadotrophins
  - cytotoxics

between free oestrogen and free androgen effects on breast tissue.

Pubertal gynaecomastia occurs in perhaps 50% of normal boys, often asymmetrically. It usually resolves spontaneously within 6–18 months, but after this duration may require surgical removal, as fibrous tissue will have been laid down. The cause is thought to be relative oestrogen excess, and the oestrogen antagonist tamoxifen is occasionally helpful.

In *the older male*, gynaecomastia requires a full assessment to exclude potentially serious underlying disease, such as bronchial carcinoma and testicular tumours (e.g. Leydig cell tumour). However, aromatase activity (p. 996) increases with age and may be the cause of gynaecomastia in this group. Aromatase is an enzyme of the cytochrome P450 family and converts androgens to produce oestrogens. Drug effects are common (especially digoxin and spironolactone), and once these and significant liver disease are excluded most cases have no definable cause. Surgery is occasionally necessary.

DISORDERS IN THE FEMALE

Hypogonadism

Impaired ovarian function, whether primary or secondary, will lead both to oestrogen deficiency and abnormalities of the menstrual cycle. The latter is very sensitive to disruption, cycles becoming anovulatory and irregular before disappearing altogether. Symptoms will depend on the age at which the failure develops. Thus, before puberty, primary amenorrhoea will occur, possibly with delayed puberty; if after puberty, secondary amenorrhoea and hypogonadism will result.

Oestrogen deficiency

The physiological effects of oestrogens and symptoms/signs of deficiency are shown in Table 18.21.

Table 18.21 Effects of oestrogens and consequences of oestrogen deficiency

Physiological effect	Consequence of deficiency
Breast	
Development of connective and duct tissue	Small, atrophic breast
Nipple enlargement and areolar pigmentation	
Pubic hair	
Maintenance of female pattern	Thinning and loss of pubic hair
Vulva and vagina	
Vulval growth	Atrophic vulva
Vaginal glandular and epithelial proliferation	Atrophic vagina
Vaginal lubrication	Dry vagina and dyspareunia
Uterus and tubes	
Myometrial and tubal hypertrophy	Small, atrophic uterus and tubes
Endometrial proliferation	Amenorrhoea
Skeletal	
Epiphyseal fusion	Eunuchoidism (if prepubertal)
Maintenance of bone mass	Osteoporosis

Amenorrhoea

Absence of periods or markedly irregular infrequent periods (oligomenorrhoea) is the commonest presentation of female gonadal disease. The clinical assessment of such patients is shown in Box 18.6, and common causes are listed in Table 18.22.

Polycystic ovary syndrome

Polycystic ovary syndrome is the most common cause of oligomenorrhoea and amenorrhoea (see below).

Weight-related amenorrhoea

A minimum body weight is necessary for regular menstruation. While anorexia nervosa is the extreme form of weight loss (see p. 1219), amenorrhoea is common and may be seen at weights within the 'normal' range. The biochemistry is indistinguishable from gonadotrophin deficiency and some patients have additional mild endocrine disease (e.g. polycystic ovarian disease). Restoration of bodyweight to above the 50th centile for height is usually effective in restoring menstruation, but in the many cases where this cannot be achieved then oestrogen replacement is necessary. Similar problems occur with intensive physical training in athletes and dancers.

Hypothalamic amenorrhoea

Amenorrhoea with low oestrogen and gonadotrophins in the absence of organic pituitary disease, weight loss or excessive exercise is described as hypothalamic amenorrhoea. This may be related to 'stress', to previous weight loss or stopping the contraceptive pill, but some patients appear to have defective cycling mechanisms without apparent explanation.

Hypothyroidism

Oligomenorrhoea and amenorrhoea are frequent findings in severe hypothyroidism in young women.

Box 18.6 Clinical assessment of amenorrhoea

History

? Pregnant
Date of onset
Age of menarche, if any
Sudden or gradual onset
General health
Weight, absolute and changes in recent past
Stress (job, lifestyle, exams, relationships)
Excessive exercise
Drugs
Hirsutism, acne, virilization
Headaches/visual symptoms
Sense of smell
Past history of pregnancies
Past history of gynaecological surgery

Examination

General health
Body shape and skeletal abnormalities
Weight and height
Hirsutism and acne
Evidence of virilization
Maturity of secondary sexual characteristics
Galactorrhoea
Normality of vagina, cervix and uterus

Table 18.22 Amenorrhoea – differential diagnosis and investigation

Hormone results					Diagnoses to consider	Secondary tests
LH	FSH	E2	PRL	T		
↑	↑	↓	N	N	**Ovarian failure**	
					Ovarian dysgenesis*	Karyotype
					Premature ovarian failure*	Ultrasound of ovary/uterus
					Steroid biosynthetic defect*	Laparoscopy/biopsy of ovary
					(Oophorectomy)	HCG stimulation
					(Chemotherapy)	
					Resistant ovary syndrome	
↓	↓	↓	N	N	**Gonadotrophin failure**	
					Hypothalamic–pituitary disease*	Pituitary MRI if diagnosis unclear
					Kallmann's syndrome/nIHH*	Clomifene test
					Possible hypothalamic causes:	Possibly LHRH test
					Hypothalamic amenorrhoea*	Serum free T_4
					Weight-related amenorrhoea*	Consider full assessment of pituitary function
					Exercise-induced amenorrhoea and anorexia*	
					Post-pill amenorrhoea	
					General illness*	
↓	↓	↓	↑	N	**Hyperprolactinaemia**	
			or ↑↑		Prolactinoma*	See page 980 (hyperprolactinaemia)
					Idiopathic hyperprolactinaemia*	Serum free T_4/TSH
					Hypothyroidism*	Pituitary MRI
					Polycystic ovarian disease*	Other tests for PCOS
					Physiological in lactation	
					Dopamine antagonist drugs	
↑/N	N	N	N/↑	N/↑	**Polycystic ovary syndrome***	
					Rarely Cushing's syndrome	Androstenedione, DHEAS
						SHBG
						Ultrasound of ovary
						Progesterone challenge
						See page 1012 (Cushing's syndrome)
N/↓	N/↓	N/↓	N	↑↑	**Androgen excess**	
					Gonadal or adrenal tumour	Imaging ovary/adrenal
					Congenital adrenal hyperplasia*	17α-OH-progesterone
N	N	↑	↑	N	**Pregnancy**	Pregnancy test
N	N	N	N	N	**Uterine/vaginal abnormality**	
					Imperforate hymen*	Examination findings (?EUA)
					Absent uterus*	Ultrasound of pelvis
					Lack of endometrium	Progesterone challenge
						Hysteroscopy

*These conditions may present as primary amenorrhoea.
LH, Luteinizing hormone; FSH, follicle-stimulating hormone; E2, oestradiol; PRL, prolactin; T, testosterone; DHEAS, dehydroepiandrosterone sulphate; SHBG, sex hormone-binding globulin; HCG, human chorionic gonadotrophin; LHRH, luteinizing hormone-releasing hormone; nIHH, normosmic idiopathic hypogonadotropic hypogonadism.

Other

Other causes include pregnancy. Genital tract abnormalities, such as an imperforate hymen, cause primary amenorrhoea.

Severe illness, even in the absence of weight loss, can lead to amenorrhoea.

Turner's syndrome (p. 1007) is a cause of primary amenorrhoea. The phenotype is female with female external genitalia. There is gonadal dysgenesis with streak ovaries. Features include short stature, webbing of the neck (up to 40%), a wide carrying angle of the elbows, high arched palate and low-set ears. These patients also have an increased incidence of autoimmune disease (2%), bicuspid aortic valves aortic coarctation and dissection and coronary artery disease, hypertension, type 2 diabetes, abnormal biochemistry, horseshoe kidneys, lymphoedema, reduced bone density and inflammatory bowel disease (0.3%).

Investigations

Basal levels of FSH, LH, oestrogen and prolactin allow initial distinction between primary gonadal and hypothalamic–pituitary causes (Table 18.22). Ovarian biopsy may occasionally be necessary to confirm the diagnosis of primary ovarian failure, although elevation of LH and FSH to menopausal levels is usually adequate. Subsequent investigations are shown in Table 18.22.

Treatment

Treatment is that of the cause wherever possible (e.g. hypothyroidism, low weight, stress, excessive exercise).

Primary ovarian disease is rarely treatable except in the rare condition of 'resistant' ovary, where high-dose gonadotrophin therapy can occasionally lead to folliculogenesis. Hyperprolactinaemia should be corrected (see below). Polycystic ovary syndrome is discussed in detail below. In all other cases oestrogen replacement is usually indicated to prevent the long-term consequences of deficiency.

Hirsutism and polycystic ovary syndrome

Normal hair vs hirsutism

The extent of normal hair growth varies between individuals, families and races, being more extensive in the Mediterranean and some Asian subcontinent populations. These normal variations in body hair, and the more extensive hair growth seen in patients complaining of hirsutism, represent a continuum from no visible hair to extensive cover with thick dark hair. It is therefore impossible to draw an absolute dividing line between 'normal' and 'abnormal' degrees of facial and body hair in the female. Soft vellus hair is normally present all over the body, and this type of hair on the face and elsewhere is 'normal' and is not sex hormone dependent. Hair in the beard, moustache, breast, chest, axilla, abdominal midline, pubic and thigh areas is sex-hormone dependent. Any excess in the latter regions is thus a marker of increased ovarian or adrenal androgen production, most commonly polycystic ovary syndrome (PCOS) but occasionally other rarer causes.

Causes of hirsutism

Polycystic ovary syndrome. PCOS is the most common cause of hirsutism in clinical practice affecting about 1 in 5 women world-wide. It is characterized by multiple small cysts within the ovary (which represent arrested follicular development) and by excess androgen production from the ovaries (and to a lesser extent from the adrenals). It was originally described in its severe form as the Stein–Leventhal syndrome.

Measured levels of androgens in blood vary widely from patient to patient and may remain within the normal range but SHBG levels are often low (due to high insulin levels), and therefore free androgen levels are high. In PCOS there is thought to be increased frequency of the GnRH pulse generator, leading to an increase in LH pulses and androgen secretion. In addition, androgens are normally converted to oestrogens in adipose tissue, but aromatase levels are low (partly due to decreased FSH levels) so that, instead of androgen (mainly androstenedione) being converted to oestrogen, androstenedione is secreted and converted to testosterone in peripheral tissue. The response of the hair follicle to circulating androgens also seems to vary between individuals with otherwise identical clinical and biochemical features, and the reason for this variation in end-organ response remains poorly understood.

PCOS is frequently associated with hyperinsulinaemia and insulin resistance. The prevalence of type 2 diabetes is ten times higher than in normal women. It is also associated with hypertension, hyperlipidaemia and increased cardiovascular disease (the metabolic syndrome, p. 231), which is 2–3 times higher in PCOS. Obesity with PCOS is an additional risk factor for insulin resistance. The precise mechanisms which link the aetiology of polycystic ovaries, hyperandrogenism, anovulation and insulin resistance remain to be elucidated and whether the basic defect is in the ovary, adrenal, pituitary or a more generalized metabolic defect remains unknown.

In routine clinical practice, the majority of patients with objective signs of androgen-dependent hirsutism will have PCOS, and investigation is mainly required to exclude rarer and more serious causes of virilization.

Idiopathic hirsutism. Patients with hirsutism, no elevation of serum androgen levels and no other clinical features are sometimes labelled as having 'idiopathic hirsutism'. However, studies suggest that most patients with 'idiopathic hirsutism' have some radiological or biochemical evidence of PCOS on more detailed investigation, and several studies have demonstrated evidence of mild PCOS in up to 20% of the normal female population.

Familial or idiopathic hirsutism does occur, but usually involves a distribution of hair growth which is not typically androgenic.

Other causes. Rarer and more serious endocrine causes of hirsutism and virilization include congenital adrenal hyperplasia (CAH, see p. 1014), Cushing's syndrome (p. 1012) and virilizing tumours of the ovary and adrenal.

Ovarian hyperthecosis is a non-malignant ovarian disorder characterized by luteinized thecal cells in the ovarian stroma which secrete testosterone. The clinical features are similar to PCOS but tend to present in perimenopausal women, and serum testosterone levels are higher than typically seen in PCOS.

Iatrogenic hirsutism also occurs after treatment with androgens, or more weakly androgenic drugs such as progestogens or danazol.

Non-androgen-dependent hair growth (hypertrichosis) occurs with drugs such as phenytoin, diazoxide, minoxidil and ciclosporin.

Clinical features of PCOS

PCOS presents with amenorrhoea/oligomenorrhoea, hirsutism and acne (alone or in combination), usually beginning shortly after menarche. Clinical, biochemical and radiological features of PCOS merge imperceptibly into those of the normal populations. The development of hirsutism commonly provokes severe distress in young women and may lead to avoidance of normal social activities.

- *Hirsutism* should be recorded objectively, ideally using a scoring system, to document the problem and to monitor treatment. The method and frequency of physical removal (e.g. shaving, plucking) should also be recorded. Most patients who complain of hirsutism will have an objective excess of hair on examination, but occasionally very little will be found (and appropriate counselling is then indicated).
- *Age and speed of onset*. Hirsutism related to PCOS usually begins around the time of the menarche and increases slowly and steadily in the teens and twenties. Rapid progression and prepubertal or late onset suggest a more serious cause.
- *Accompanying virilization*. Hirsutism due to PCOS may be severe and affect all androgen-dependent areas on the face and body. However, more severe virilization (clitoromegaly, recent-onset frontal balding, male phenotype) implies substantial androgen excess, and usually indicates a rarer cause rather than PCOS. Thinning of head hair in a male pattern – androgenic alopecia – occurs in a proportion of women with uncomplicated PCOS, typically with a familial tendency for premature androgen-related hair loss in both sexes.
- *Menstruation*. Most patients with hirsutism will have some disturbance of menstruation, typically oligo-/amenorrhoea, although more frequent erratic bleeding

can also occur. However, PCOS can present as hirsutism with regular periods, or as irregular periods with no evidence of hirsutism or acne.

- *Weight*. Many patients with hirsutism are also overweight or obese. This worsens the underlying androgen excess and insulin resistance and inhibits the response to treatment, and is an indication for appropriate advice on diet and exercise. In severe cases the insulin resistance may have a visible manifestation as acanthosis nigricans on the neck and in the axillae (see Fig. 23.24).

Investigations and differential diagnosis

A variety of investigations aid the diagnosis of patients with hirsutism:

- **Serum total testosterone** is often elevated in PCOS and is invariably substantially raised in virilizing tumours (usually >5 nmol/L). Patients with hirsutism and normal testosterone levels frequently have low levels of sex hormone-binding globulin (SHBG), leading to high free androgen levels. *The free androgen index* (total testosterone/SHBG concentration) is often used and is high; free testosterone is difficult to measure.
- **Other androgens**. Androstenedione and dehydroepiandrosterone sulphate are frequently elevated in PCOS, and even more elevated in congenital adrenal hyperplasia (CAH) and virilizing tumours.
- **17α-Hydroxyprogesterone** is elevated in classical CAH, but may be apparent in late-onset CAH only after stimulation tests.
- **Gonadotrophin levels**. LH hypersecretion is a frequent feature of PCOS, but the pulsatile nature of secretion of this hormone means that a 'classic' increased LH/FSH ratio is not always observed on a random sample.
- **Oestrogen levels**. Oestradiol is usually normal in PCOS, but oestrone levels (which are rarely measured) are elevated because of peripheral conversion. Levels are variable in other causes.
- **Ovarian ultrasound** is a useful investigation (Fig. 18.20). Typical features are those of a thickened capsule, multiple 3–5 mm cysts and a hyperechogenic stroma. Prolonged hyperandrogenization from any cause may lead to polycystic changes in the ovary. Ultrasound may also reveal virilizing ovarian tumours, although these are often small.
- **Serum prolactin**. Mild hyperprolactinaemia is common in PCOS but rarely exceeds 1500 mU/L.

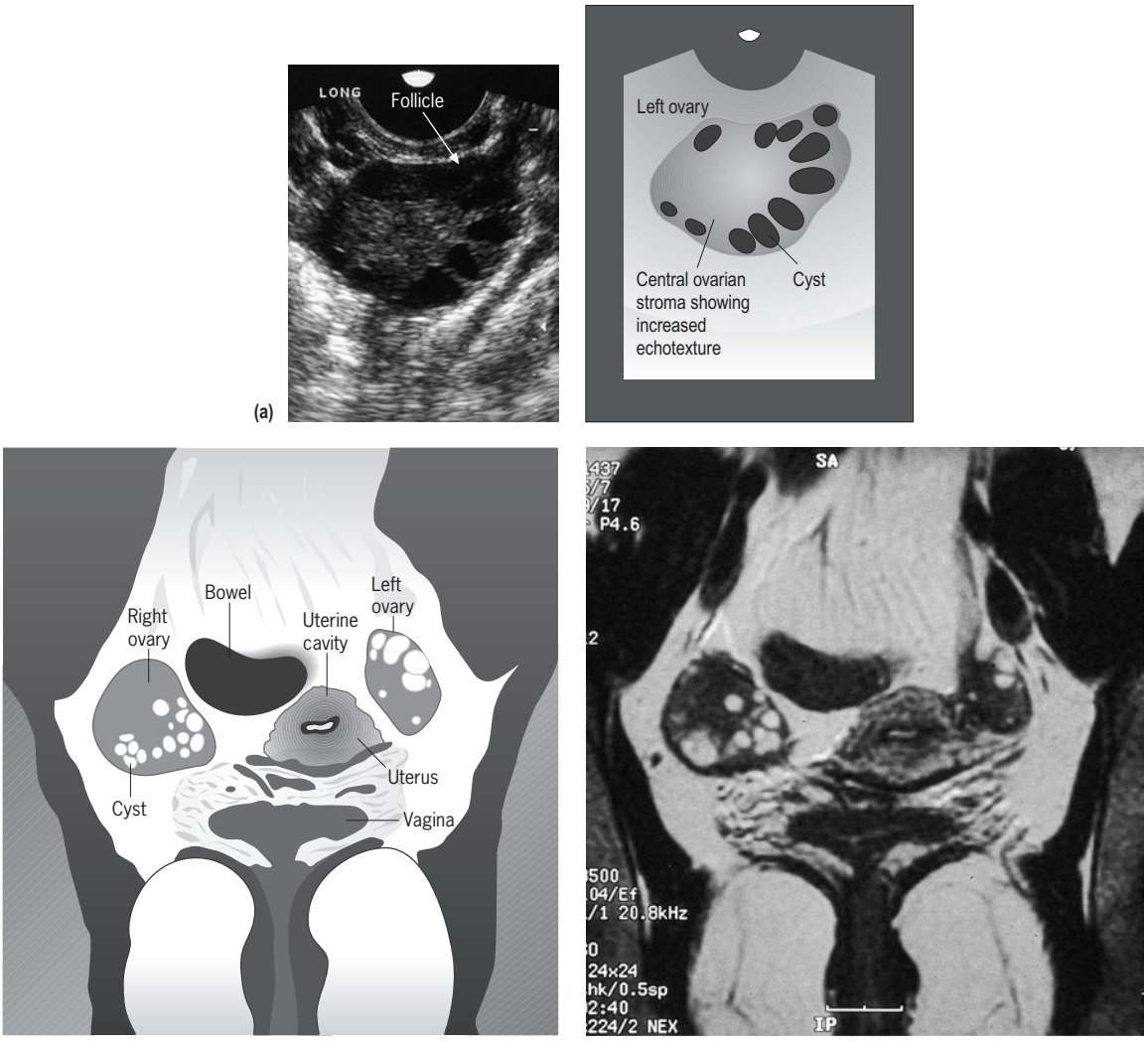


Fig. 18.20 **Polycystic ovary syndrome. (a) Longitudinal transvaginal ultrasound of ovary,** revealing multiple cysts with central ovarian stroma showing increased echo texture. **(b) MR image (coronal) of polycystic ovaries,** also showing pelvic anatomy. Reproduced by kind permission of Barbara Hochstein and Geoffrey Cox, Auckland Radiology Group.

If a virilizing tumour is suspected clinically or after investigation, then more complex tests include dexamethasone suppression tests, CT or MRI of adrenals, and selective venous sampling.

Diagnosis

Most patients presenting with a combination of hirsutism and menstrual disturbance will be shown to have polycystic ovary syndrome, but the rarer alternative diagnoses should be excluded, e.g. late-onset congenital adrenal hyperplasia (early-onset, raised serum 17α-OH-progesterone), Cushing's syndrome (look for other clinical features) and virilizing tumours of the ovary or adrenals (severe virilization, markedly elevated serum testosterone).

The consensus (Rotterdam) criteria 2003 for diagnosis of PCOS are at least two of:

- clinical or biochemical evidence of hyperandrogenism
- evidence of oligo- or anovulation
- presence of polycystic ovaries on ultrasound.

Treatments

The underlying cause should be removed in the rare instances where this is possible (e.g. drugs, adrenal or ovarian tumours). Treatment of CAH and Cushing's is discussed on page 1014 and page 1012, respectively. Other therapy depends upon whether the aim is to reduce hirsutism, regularize periods or produce fertility.

Local therapy for hirsutism

Regular plucking, bleaching, depilatory cream, waxing or shaving is used. Such removal neither worsens nor improves the underlying severity of hirsutism. More 'permanent' solutions include electrolysis and a variety of 'laser' hair removal systems – all appear effective but have not been evaluated in long-term studies, are expensive, and still often require repeated long-term treatment. *Eflornithine cream* (an antiprotozoal) inhibits hair growth by inhibiting ornithine decarboxylase but is effective in only a minority of cases and should be discontinued if there is no improvement after four months.

Systemic therapy for hirsutism

This always requires a year or more of treatment for maximal benefit, and long-term treatment is frequently required as the problem tends to recur when treatment is stopped. The patient must therefore always be an active participant in the decision to use systemic therapy and must understand the rare risks as well as the benefits.

- *Oestrogens* (e.g. oral contraceptives) suppress ovarian androgen production and reduce free androgens by increasing SHBG levels. Combined hormone pills, which contain ethinylestradiol and a non-androgenic progestogen, e.g. desogestrel drospirenone, or cyproterone acetate (co-cyprindiol), will result in a slow improvement in hirsutism in a majority of cases and should normally be used first unless there is a contraindication, e.g. history of thrombosis. The risk of venous thrombosis appears to be 2–4-fold higher than on other low-dose oral contraceptive pills. After the menopause, HRT preparations which contain medroxyprogesterone (rather than more androgenic progestogens) may be helpful.
- *Cyproterone acetate* (50–100 mg daily) is an antiandrogen but is also a progestogen, teratogenic and a weak glucocorticoid. Given continuously it produces amenorrhoea, and so is normally given for days 1–14 of each cycle. In women of childbearing age, contraception is essential.
- *Spironolactone* (200 mg daily) also has antiandrogen activity and can cause useful improvements in hirsutism.
- *Finasteride* (5 mg daily), a 5α-reductase inhibitor which prevents the formation of dihydrotestosterone in the skin, has also been shown to be effective but long-term experience is limited.
- *Flutamide*, another antiandrogen, is less commonly used owing to the high incidence of hepatic side-effects.

Treatment of menstrual disturbance

- Cyclical oestrogen/progestogen administration will regulate the menstrual cycle and remove the symptom of oligo- or amenorrhoea. This is most frequently an additional benefit of the treatment of hirsutism, but may also be used when menstrual disturbance is the only symptom.
- Drugs to improve the hyperinsulinaemia associated with PCOS and obesity are increasingly used (and requested by patients). Metformin (500 mg three times daily) improves menstrual cyclicity and ovulation in short-term studies, and some patients also report improvement in hirsutism and ease of weight loss, but gastrointestinal upset may limit use.
- Thiazolidinediones are also effective, but there are relatively few studies which are mostly short-term, and there are concerns about the long-term cardiovascular effects of these agents.

Treatment for fertility

- Metformin alone may improve ovulation and achieve conception.
- Clomifene 50–100 mg can be given daily on days 2–6 of the cycle and is more effective than metformin alone in achieving ovulation. This can occasionally cause the *ovarian hyperstimulation syndrome*, an iatrogenic complication of ovulation induction therapy, consisting of ovarian enlargement, oedema, hypovolaemia, renal failure, and possibly shock; specialist supervision is essential. It is recommended that clomifene should not normally be used for longer than six cycles (owing to a possible increased risk of ovarian cancer in patients treated for longer than recommended).
- *Reverse circadian rhythm*. Prednisolone (2.5 mg in the morning, 5 mg on retiring) suppresses pituitary production of ACTH, upon which adrenal androgens partly depend. Regular ovulatory cycles often ensue. A steroid instruction leaflet and a card must be supplied.

More intensive techniques to stimulate ovulation may also be indicated in specialist hands, including low-dose gonadotrophin therapy, and ovarian hyperstimulation techniques associated with in vitro fertilization.

Wedge resection of the ovary was a traditional therapy but is now rarely required, although laparoscopic ovarian electrodiathermy may be helpful.

Oral contraception

The combined oestrogen–progestogen pill is widely used for contraception and has a low failure rate (<1 per 100 woman-years). 'Pills' contain 20–40 μg of oestrogen, usually ethinyl-

estradiol, together with a variable amount of one of several progestogens. The mechanism of action is twofold:

- suppression by oestrogen of gonadotrophins, thus preventing follicular development, ovulation and luteinization
- progestogen effects on cervical mucus, making it hostile to sperm, and on tubal motility and the endometrium.

Side-effects of these preparations are shown in Box 18.7. Most of the serious ones are rare and are less common with typical modern 20–30 μg oestrogen pills, although evidence suggests that thromboembolism may be slightly more common with 'third-generation pills' containing desogestrel and gestodene (approximately 30/100 000 woman-years compared with 15/100 000 on older pills and 5/100 000 on no treatment). While some problems require immediate cessation of the pill, other milder side-effects must be judged against the hazards of pregnancy occurring with inadequate contraception, especially if other effective methods are not practicable or acceptable.

Hazards of the combined pill are increased in smokers, in obesity and in those with other risk factors for cardiovascular disease (e.g. hypertension, hyperlipidaemia, diabetes) especially in women aged over 35 years (avoid if over 50 years). The 'mini-pill' (progestogen only, usually norethisterone) is less effective but is often suitable where oestrogens are contraindicated (Box 18.7). A progesterone antagonist, mifepristone, in combination with a prostaglandin analogue (vaginal gemeprost), induces abortion of pregnancy at up to 9 weeks' gestation. It prevents progesterone-induced inhibition of uterine contraction.

Box 18.7 Adverse effects and drug interactions of oral contraceptives (mixed oestrogen–progesterone combinations)

General

Weight gain
Loss of libido
Pigmentation (chloasma)
Breast tenderness
Increased growth rate of some malignancies

Cardiovascular

Increased blood pressure*
Deep vein thrombosis*
Myocardial infarction
Stroke

Gastrointestinal

Nausea and vomiting
Abnormal liver biochemistry*
Gallstones increased
Hepatic tumours

Nervous system

Headache
Migraine*
Depression*

Malignancy

Increase in cancer of the breast (but reduced risk of ovarian and endometrial cancer)

Gynaecological

Amenorrhoea
'Spotting'
Cervical erosion

Haematological

Increased clotting tendency

Endocrine/metabolic

Mild impairment of glucose tolerance
Worsened lipid profile, though variable

Drug interactions (reduced contraceptive effect owing to enzyme induction)

Antibiotics
Barbiturates
Phenytoin
Carbamazepine
Rifampicin
St John's wort

*Common reasons for stopping oral contraceptives.

SUBFERTILITY

Subfertility, or 'infertility', is defined as the inability of a couple to conceive after 1 year of unprotected intercourse. Investigation requires the combined skills of gynaecologist, endocrinologist and, ideally, andrologist. Both partners must be involved and every aspect of the physiology critically examined.

Causes (Fig. 18.21)

A significant proportion of couples have both male and female contributing factors.

Inadequate intercourse, hostile cervical mucus and vaginal factors are uncommon (5%). Fifteen per cent of cases appear to be idiopathic, and natural fertility decreases with increasing age.

Male factors

About 30–40% of couples have a major identifiable male factor. There is some evidence that male sperm counts are declining in many populations. Untreated male hypogonadism of any cause (see Table 18.18) is likely to be associated with subfertility.

Female factors

Female tubal problems account for perhaps 20%; a similar proportion have ovulatory disorders. Any cause of oligomenorrhoea or amenorrhoea (see Table 18.22) is likely to be associated with suboptimal ovulation or anovulation.

Clinical assessment

Both partners should be seen and the following factors checked:

- *The man*. Look for previous testicular damage (orchitis, trauma), undescended testes, urethral symptoms and evidence of sexually transmitted infection, local surgery, and use of alcohol and drugs. A semen analysis early in the investigations is essential.
- *The woman*. Look for previous pelvic infection, regularity of periods, previous surgery, alcohol intake and smoking, and bodyweight (see p. 1001).
- *Together*. Check the frequency and adequacy of intercourse, and the use of lubricants.

Investigations

See Figure 18.21.

www.studentconsult.com
Ask the authors

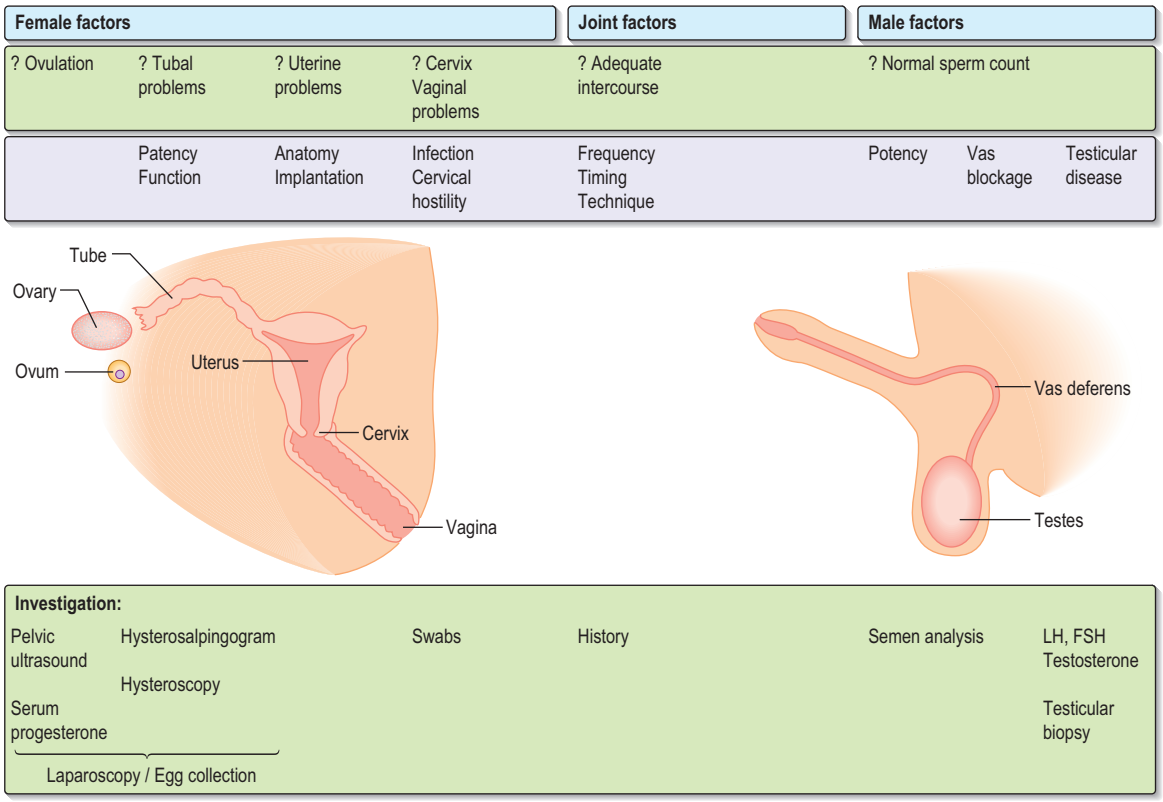


Fig. 18.21 **Major factors involved in subfertility and their investigation.** LH, luteinizing hormone; FSH, follicle-stimulating hormone.

Table 18.23 Disorders of sexual differentiation

Condition	Chromosomes	Gonads	Phenotype	Remarks
Turner's syndrome	45X (50%) 46,X,i (Xq) (5–10%) 45,X, mosaicism (remainder)	Streak	Female	Often morphological features (e.g. short stature, web neck, coarctation of aorta)
Gonadal dysgenesis	46XY	Streak or minimal testes*	Immature female	
Congenital adrenal hyperplasia	46XX	Ovary	Female with variable virilization	Obvious androgen excess
Virilizing tumour	46XX	Ovary	Female with variable virilization	Obvious androgen excess
True hermaphroditism	46XX/XY or mosaic	Testis and ovary	Male or ambiguous	
Klinefelter's syndrome	47XXY	Small testes	Male, often with gynaecomastia	Many are hypogonadal
Testicular feminization	46XY	Testes*	Ambiguous or infantile female	Androgen receptor defective
Testicular synthetic defects	46XY	Testes*	Cryptorchid, ambiguous	
5α-Reductase deficiency	46XY	Testes	Cryptorchid, ambiguous	Impaired conversion of testosterone to dihydrotestosterone
Anorchia	46XY	Absent	Immature female	

*Gonadectomy advised because of high risk of malignancy. i, isochromosome.

Treatment

Counselling of both partners is essential. Any defect(s) found should be treated if possible. Ovulation can usually be induced by exogenous hormones if simpler measures fail, while in vitro fertilization (IVF) and similar techniques are widely used, especially where there is tubal blockage, oligospermia or 'idiopathic subfertility'. Intracytoplasmic sperm injection (ICSI) appears particularly effective for severe oligospermia and poor sperm function.

DISORDERS OF SEXUAL DIFFERENTIATION

Disorders of sexual differentiation are rare but may affect chromosomal, gonadal, endocrine and phenotypic development (Table 18.23). Such cases always require extensive,

FURTHER READING

Grady D. Management of menopausal symptoms. *New England Journal of Medicine* 2006; **355**: 2338–2347.

MacLaughlin DT, Donahoe PK. Sex determination and differentiation. *New England Journal of Medicine* 2004; **350**: 367–378.

Million Women Study Collaborators. Breast cancer and HRT in the Million Women Study. *Lancet* 2003; **362**: 419–427.

Norman RJ, Dewailly D, Legro RS et al. Polycystic ovary syndrome. *Lancet* 2007; **370**: 685–697.

Rosenfield RL. Clinical practice. Hirsutism. *New England Journal of Medicine* 2005; **353**: 2578–2588.

multidisciplinary clinical management. An individual's sex can be defined in several ways:

- *Chromosomal sex*. The normal female is 46XX, the normal male 46XY. The Y chromosome confers male sex; if it is not present, development follows female lines.
- *Gonadal sex*. This is obviously determined predominantly by chromosomal sex, but requires normal embryological development.
- *Phenotypic sex*. This describes the normal physical appearance and characteristics of male and female body shape. This in turn is a manifestation of gonadal sex and subsequent sex hormone production.
- *Social sex (gender)*. This is heavily dependent on phenotypic sex and normally assigned on appearance of the external genitalia at birth.
- *Sexual orientation* – heterosexual, homosexual or bisexual.

THE GLUCOCORTICOID AXIS

Adrenal gland – anatomy and function

(see Table 18.3)

The human adrenals weigh 8–10 g together and comprise an outer cortex with three zones (reticularis, fasciculata and glomerulosa) producing steroids, and an inner medulla that synthesizes, stores and secretes catecholamines (see adrenal medulla, p. 1024).

The adrenal steroids are grouped into three classes based on their predominant physiological effects.

Glucocorticoids

These are so named after their effects on carbohydrate metabolism. Major actions are listed in Table 18.24. They act on intracellular corticosteroid receptors and combine with coactivating proteins to bind the 'glucocorticoid response element' (GRE) in specific regions of DNA to cause gene transcription. Glucocorticoid action is modified locally by the action of 11β-hydroxysteroid dehydrogenase (11βHSD). 11βHSD Type 1 converts inactive cortisone in cortisol, hence amplifying the hormone signal, whilst 11βHSD type 2 does the opposite.

The relative potency of common steroids is shown in Table 18.25.

Table 18.24 The major actions of glucocorticoids

Increased or stimulated	Decreased or inhibited
Gluconeogenesis	Protein synthesis
Glycogen deposition	Host response to infection
Protein catabolism	Lymphocyte transformation
Fat deposition	Delayed hypersensitivity
Sodium retention	Circulating lymphocytes
Potassium loss	Circulating eosinophils
Free water clearance	
Uric acid production	
Circulating neutrophils	

Mineralocorticoids

The predominant effect of mineralocorticoids is on the extracellular balance of sodium and potassium in the distal tubule of the kidney. Aldosterone, produced solely in the zona glomerulosa, is the predominant mineralocorticoid in humans (about 50%); corticosterone makes a small contribution. Mineralocorticoids act on type 1 corticosteroid receptors, whilst glucocorticoids act on type 2 receptors, both having a very similar structure. The mineralocorticoid activity of cortisol is weak but it is present in considerable excess. The mineralocorticoid receptor in the kidney is largely protected from this excess by the intrarenal conversion ('shuttle') of cortisol to cortisone by 11β-hydroxysteroid dehydrogenase type 2.

Androgens

Although secreted in considerable quantities, most androgens have only relatively weak intrinsic androgenic activity until metabolized peripherally to testosterone or dihydrotestosterone. Dihydrotestosterone is metabolized from testosterone by 5α-reductase and is a potent androgen receptor agonist. The androgen receptor has been well characterized (p. 966) and mutations within this gene may cause androgen insensitivity syndromes.

Biochemistry

All steroids have the same basic skeleton (Fig. 18.22b) and the chemical differences between them are slight. The major biosynthetic pathways are shown in Figure 18.22a.

Physiology

Glucocorticoid production by the adrenal is under hypothalamic–pituitary control (Fig. 18.23).

Corticotrophin-releasing hormone (CRH) is secreted in the hypothalamus in response to circadian rhythm, stress and other stimuli. CRH travels down the portal system to stimulate *adrenocorticotrophin (ACTH)* release from the anterior pituitary. ACTH is derived from the prohormone pro-opiomelanocortin (POMC), which undergoes complex processing within the pituitary to produce ACTH and a number of other peptides including beta-lipotrophin and beta-endorphin. Many of these peptides, including ACTH, contain melanocyte-stimulating hormone (MSH)-like sequences which cause pigmentation when levels of ACTH are markedly raised.

Circulating ACTH stimulates cortisol production in the adrenal. The cortisol secreted (or any other synthetic corti-

Table 18.25 The relative glucocorticoid and mineralocorticoid potency of equal amounts of common natural and synthetic steroids

Steroid	Glucocorticoid effect	Mineralocorticoid effect
Cortisol (hydrocortisone)*	1	1
Prednisolone	4	0.7
Dexamethasone	40	2
Aldosterone	0.1	400
Fludrocortisone	10	400

*Cortisol is arbitrarily defined as 1.

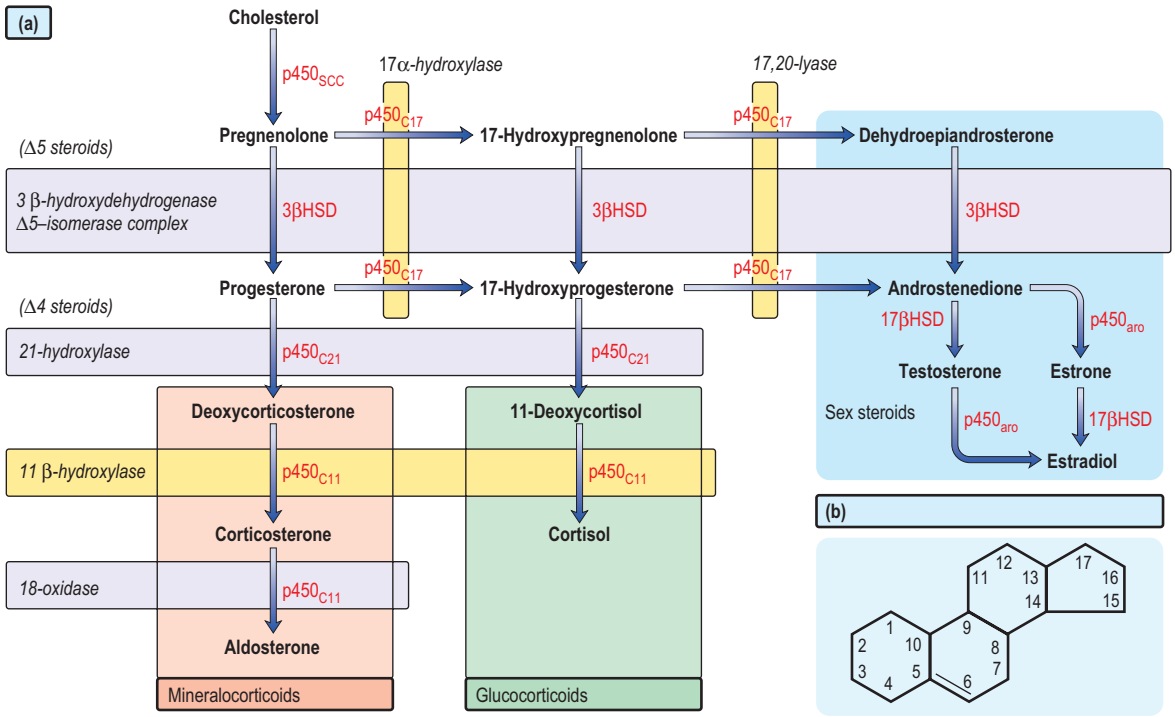


Fig. 18.22 **(a) The major steroid biosynthetic pathways.** Steroid hormones are synthesized in adrenal and gonads from the initial substrate cholesterol via a series of interconnected enzymatic steps. The reactions catalysed are shown by shaded boxes with italic labels. The molecular identity of the enzymes is shown in red – some catalyse more than one reaction. p450 enzymes are in mitochondria, 3βHSD (hydroxysteroid dehydrogenase) in cytoplasm. 17βHSD and $p450_{aro}$ (aromatase) are found mainly in gonads. **(b) The steroid molecule.**

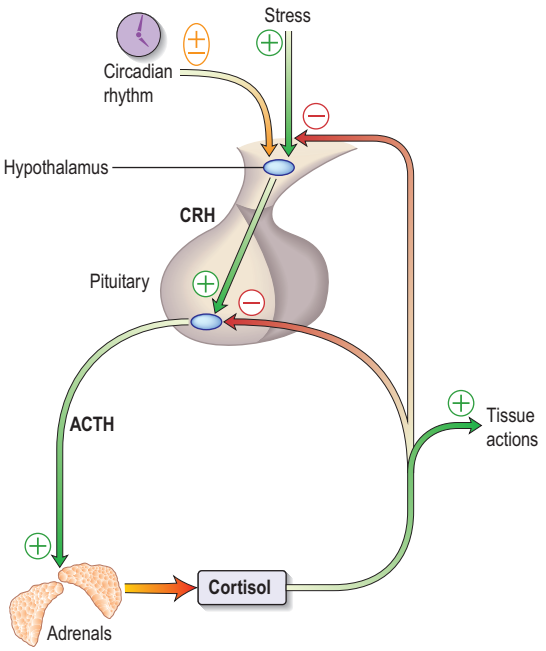


Fig. 18.23 **Control of the hypothalamic–pituitary–adrenal axis.** Pituitary ACTH is secreted in response to hypothalamic CRH (corticotrophin-releasing hormone) triggered by circadian rhythm, stress and other factors, and stimulates secretion of cortisol from the adrenal. Cortisol has multiple actions in peripheral tissues and exerts negative feedback on pituitary and hypothalamus.

costeroid administered to the patient) causes negative feedback on the hypothalamus and pituitary to inhibit further CRH/ACTH release. The set-point of this system clearly varies through the day according to the circadian rhythm, and is usually overridden by severe stress. Unlike cortisol, mineralocorticoids and sex steroids do not cause negative feedback on the CRH/ACTH axis.

Following adrenalectomy or other adrenal damage (e.g. Addison's disease), cortisol secretion will be absent or reduced; ACTH levels will therefore rise.

Mineralocorticoid secretion is mainly controlled by the renin–angiotensin system (see p. 1023).

Investigation of glucocorticoid abnormalities

Basal levels

ACTH and cortisol are released episodically and in response to stress. When taking a blood sample, remember:

- Sampling time should be recorded. Basal levels should be taken between 0800 h and 0900 h near the peak of the circadian variation.
- Stress should be minimized.
- Appropriate reference ranges (for time and assay method) should be used.

Suppression and stimulation tests are used in suspected excess and deficient cortisol production, respectively.

Dexamethasone suppression tests

Administration of a synthetic glucocorticoid to a normal subject produces prompt feedback suppression of CRH and ACTH levels and thus of endogenous cortisol secretion (dexamethasone is not measured by most cortisol assays).

Table 18.26 Details of dexamethasone suppression and ACTH (tetracosactide) tests in the diagnosis of Cushing's syndrome and Addison's disease

Test and protocol	Measure	Normal test result or positive suppression	Use and explanation
Dexamethasone (for Cushing's)			
Overnight			
Take 1 mg on going to bed at 2300 h	Plasma cortisol at 0900 h next morning	Plasma cortisol <100 nmol/L	Outpatient screening test Some 'false positives'
'Low-dose'			
0.5 mg 6-hourly Eight doses from 0900 h on day 0	Plasma cortisol at 0900 h on days 0 and +2	Plasma cortisol <50 nmol/L on second sample	For diagnosis of Cushing's syndrome
'High-dose' used in differential diagnosis			
2 mg 6-hourly Eight doses from 0900 h on day 0	Plasma cortisol at 0900h on days 0 and +2	Plasma cortisol on day +2 less than 50% of that on day 0 suggests pituitary-dependent disease	Differential diagnosis of Cushing's syndrome Pituitary-dependent disease suppresses in about 90% of cases
ACTH (tetracosactide) (for Addison's)			
Short			
Tetracosactide 250 µg i.v. or i.m. at time 0	Plasma cortisol at times 0, +30 min	Cortisol at +30 min >600 nmol/L	To exclude primary adrenal failure
Long			
Depot tetracosactide 1 mg i.m. at time 0	Plasma cortisol at times 0, +1, +2, +3, +4, +5, +8 and +24 h	Maximum >1000 nmol/L Rise >550 nmol/L	To demonstrate or exclude adrenal suppression (rather than primary adrenal failure)

Plasma cortisol values are very dependent upon the assay used – local reference ranges must be consulted.

Three forms of the test, used in the diagnosis and differential diagnosis of Cushing's syndrome, are available (Table 18.26).

ACTH stimulation tests

Synthetic ACTH (tetracosactide, which consists of the first 24 amino acids of human ACTH) is given to stimulate adrenal cortisol production. Details are given in Table 18.26 and Box 18.1 on page 968.

Addison's disease: primary hypoadrenalism

Pathophysiology and causes

In this condition there is destruction of the entire adrenal cortex. Glucocorticoid, mineralocorticoid and sex steroid production are therefore all reduced. (This differs from hypothalamic–pituitary disease, in which mineralocorticoid secretion remains largely intact, being predominantly stimulated by angiotensin II. Adrenal sex steroid production is also largely independent of pituitary action.) In Addison's disease reduced cortisol levels lead, through feedback, to increased CRH and ACTH production, the latter being directly responsible for the hyperpigmentation.

Incidence. Addison's disease is rare, with an incidence of 3–4/million/year and prevalence of 40–60/million. Primary hypoadrenalism shows a marked female preponderance and is most often caused by autoimmune disease (>90% in UK) but in countries with a high prevalence of HIV/AIDS, tuberculosis is an increasing cause. Autoimmune adrenalitis results from the destruction of the adrenal cortex by organ-specific autoantibodies, with 21-hydroxylase as the common antigen. There are associations with other autoimmune conditions in the polyglandular autoimmune syndromes types I and II (e.g. type 1 diabetes mellitus, pernicious anaemia, thyroiditis, hypoparathyroidism, premature ovarian failure) (p. 1026).

All other causes are rare (Table 18.27).

Table 18.27 Causes of primary hypoadrenalism

- Autoimmune disease
- Tuberculosis (<10% in UK)
- Surgical removal
- Haemorrhage/infarction
  - Meningococcal septicaemia
  - Venography
- Infiltration
  - Malignant destruction
  - Amyloid
- Schilder's disease (adrenal leucodystrophy)

Clinical features

These are shown in Figure 18.24. The symptomatology of Addison's disease is often vague and non-specific. These symptoms may be the prelude to an Addisonian crisis with severe hypotension and dehydration precipitated by intercurrent illness, accident or operation.

Pigmentation (dull, slaty, grey-brown) is the predominant sign in over 90% of cases.

Postural systolic hypotension, due to hypovolaemia and sodium loss, is present in 80–90% of cases, even if supine blood pressure is normal. Mineralocorticoid deficiency is the cause of the hypotension.

Investigations

Once Addison's disease is suspected, investigation is urgent. If the patient is seriously ill or hypotensive, hydrocortisone

Symptoms

Weight loss
Anorexia
Malaise
Weakness
Fever
Depression
Impotence/amenorrhoea
Nausea/vomiting
Diarrhoea
Confusion
Syncope from postural hypotension
Abdominal pain
Constipation
Myalgia
Joint or back pain

Signs

Pigmentation, especially of new scars and palmar creases
Buccal pigmentation
Postural hypotension
Loss of weight
General wasting
Dehydration
Loss of body hair
(Vitiligo)

Fig. 18.24 **Primary hypoadrenalism (Addison's disease) – symptoms and signs.** Bold type indicates signs of greater discriminant value.

100 mg should be given intramuscularly together with intravenous saline. Ideally this should be done immediately after a blood sample is taken for later measurement of plasma cortisol. Alternatively, an ACTH stimulation test can be performed immediately. Full investigation should be delayed until emergency treatment (see below) has improved the patient's condition. Otherwise, tests are as follows:

- **Single cortisol measurements** are of little value, although a random cortisol below 100 nmol/L during the day is highly suggestive, and a random cortisol >550 nmol/L makes the diagnosis unlikely (but not impossible).
- **The short ACTH stimulation test** should be performed (see Table 18.26). Note that an absent or impaired cortisol response confirms the presence of hypoadrenalism but does not differentiate Addison's disease from ACTH deficiency or iatrogenic suppression by steroid medication.
- **A 0900 h plasma ACTH level** – a high level (>80 ng/L) with low or low-normal cortisol confirms primary hypoadrenalism.
- **A long ACTH stimulation test** can also exclude adrenal suppression by steroids or ACTH deficiency.
- **Electrolytes and urea** classically show hyponatraemia, hyperkalaemia and a high urea, but they can be normal.
- **Blood glucose** may be low, with symptomatic hypoglycaemia.
- **Adrenal antibodies** are present in many cases of autoimmune adrenalitis.
- **Chest and abdominal X-rays** may show evidence of tuberculosis and/or calcified adrenals.
- **Serum aldosterone** is reduced with high plasma renin activity.
- **Hypercalcaemia and anaemia** (after rehydration) are sometimes seen. They resolve on treatment, but are occasionally the first clue to the diagnosis.

Treatment

Acute hypoadrenalism needs urgent treatment (Emergency box 18.1).

Long-term treatment is with replacement glucocorticoid and mineralocorticoid; tuberculosis must be treated if present or suspected. Replacement dosage details are shown in Table 18.28. Dehydroepiandrosterone (DHEA) replacement has also been advocated, and studies suggest that this may cause symptomatic improvements, although long-term results are awaited.

Emergency Box 18.1

Management of acute hypoadrenalism

- Clinical context: hypotension, hyponatraemia, hyperkalaemia, hypoglycaemia, dehydration, pigmentation often with precipitating infection, infarction, trauma or operation. The major deficiencies are of salt, steroid and glucose.
- Assuming normal cardiovascular function, the following are required:
- One litre of 0.9% saline should be given over 30–60 minutes with 100 mg of intravenous bolus hydrocortisone.
- Subsequent requirements are several litres of saline within 24 hours (assessing with central venous pressure line if necessary) plus hydrocortisone, 100 mg i.m., 6-hourly, until the patient is clinically stable.
- Glucose should be infused if there is hypoglycaemia.
- Oral replacement medication is then started, unless unable to take oral medication, initially hydrocortisone 20 mg, 8-hourly, reducing to 20–30 mg in divided doses over a few days (Table 18.28).
- Fludrocortisone is unnecessary acutely as the high cortisol doses provide sufficient mineralocorticoid activity – it should be introduced later.

Adequacy of glucocorticoid dose is judged by:

- clinical well-being and restoration of normal, but not excessive, weight
- normal cortisol levels during the day while on replacement hydrocortisone (cortisol levels cannot be used for synthetic steroids).

Fludrocortisone replacement is assessed by:

- restoration of serum electrolytes to normal
- blood pressure response to posture (it should not fall >10 mmHg systolic after 2 minutes' standing)
- suppression of plasma renin activity to normal.

Patient advice

All patients requiring replacement steroids should:

Table 18.28 Average replacement steroid dosages for adults with primary hypoadrenalism

Drug	Dose
Glucocorticoid	
Hydrocortisone	20–30 mg daily e.g. 10 mg on waking, 5 mg at 1200 h, 5 mg at 1800 h
or	
Prednisolone	7.5 mg daily 5 mg on waking, 2.5 mg at 1800 h
rarely	
Dexamethasone	0.75 mg daily 0.5 mg on waking, 0.25 mg at 1800 h
Mineralocorticoid	
Fludrocortisone	50–300 μg daily

- know how to increase steroid replacement dose for intercurrent illness
- carry a 'Steroid Card'
- wear a Medic-Alert bracelet (or similar), which gives details of their condition so that emergency replacement therapy can be given if found unconscious
- keep an (up-to-date) ampoule of hydrocortisone at home in case oral therapy is impossible, for administration by self, family or GP.

Secondary hypoadrenalism

This may arise from hypothalamic–pituitary disease (inadequate ACTH production) or from long-term steroid therapy leading to hypothalamic–pituitary–adrenal suppression.

Most patients with the former have panhypopituitarism (see p. 974) and need T_4 replacement as well as cortisol; in this case hydrocortisone must be started before T_4.

The most common cause of secondary hypoadrenalism is long-term corticosteroid medication for non-endocrine disease. The hypothalamic–pituitary axis and the adrenal may both be suppressed and the patient may have vague symptoms of feeling unwell. The long ACTH stimulation test should demonstrate a delayed cortisol response. Weaning off steroids is often a long and difficult process.

Cushing's syndrome

Cushing's syndrome is the term used to describe the clinical state of increased free circulating glucocorticoid. It occurs most often following the therapeutic administration of synthetic steroids or ACTH (see below). All the spontaneous forms of the syndrome are rare.

Pathophysiology and causes

Spontaneous Cushing's syndrome is rare, with an incidence of <5/million/year.

Causes of Cushing's syndrome are usually subdivided into two groups (Table 18.29):

- increased circulating ACTH from the pituitary (65% of cases), known as Cushing's disease, or from an 'ectopic', non-pituitary, ACTH-producing tumour elsewhere in the body (10%) with consequent glucocorticoid excess ('ACTH dependent' Cushing's)

Table 18.29 Causes of Cushing's syndrome

ACTH-dependent disease
Pituitary-dependent (Cushing's disease)
Ectopic ACTH-producing tumours
ACTH administration
Non-ACTH-dependent causes
Adrenal adenomas
Adrenal carcinomas
Glucocorticoid administration
Others
Alcohol-induced pseudo-Cushing's syndrome

- a primary excess of endogenous cortisol secretion (25% of spontaneous cases) by an adrenal tumour or nodular hyperplasia, with subsequent (physiological) suppression of ACTH. Rare cases are due to aberrant expression of receptors for other hormones (e.g. glucose-dependent insulinotrophic peptide (GIP), LH or catecholamines) in adrenal cortical cells ('ACTH independent' Cushing's).

Clinical features

The clinical features of Cushing's syndrome are those of glucocorticoid excess and are illustrated in Figure 18.25.

- *Pigmentation* occurs only with ACTH-dependent causes.
- *A Cushingoid appearance* can be caused by excess alcohol consumption (pseudo-Cushing's syndrome) – the pathophysiology is poorly understood.
- *Impaired glucose tolerance* or frank diabetes is common, especially in the ectopic ACTH syndrome.
- *Hypokalaemia* due to the mineralocorticoid activity of cortisol is common with ectopic ACTH secretion.

Diagnosis

There are two phases to the investigation:

1. Confirmation

Most obese, hirsute, hypertensive patients do not have Cushing's syndrome, and some cases of genuine Cushing's have relatively subtle clinical signs. Confirmation rests on demonstrating inappropriate cortisol secretion, not suppressed by exogenous glucocorticoids: difficulties occur with obesity and depression where cortisol dynamics are often abnormal. Random cortisol measurements are of no value. Occasional patients are seen with so-called 'cyclical Cushing's' where the abnormalities come and go.

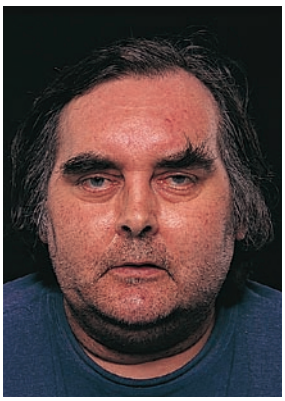
Investigations to confirm the diagnosis include:

- **48-hour low-dose dexamethasone test** (see Table 18.26). Normal individuals suppress plasma cortisol to <50 nmol/L. Patients with Cushing's syndrome fail to show complete suppression of plasma cortisol levels (although levels may fall substantially in a few cases). This test is highly sensitive (>97%). The overnight dexamethasone test is slightly simpler, but has a higher false-positive rate.
- **24-hour urinary free cortisol measurements**. This is simple, but less reliable – repeatedly normal values render the diagnosis most unlikely, but some patients with Cushing's have normal values on some collections (approximately 10%).
- **Circadian rhythm**. After 48 hours in hospital, cortisol samples are taken at 0900 h and 2400 h (without

Symptoms

Weight gain (central)
Change of appearance
Depression
Insomnia
Amenorrhoea/ oligomenorrhoea
Poor libido
Thin skin/easy bruising
Hair growth/acne
Muscular weakness
Growth arrest in children
Back pain
Polyuria/polydipsia
Psychosis

Old photographs may be useful



Signs

Moon face
Plethora
Depression/psychosis
Acne
Hirsutism
Frontal balding (female)
Thin skin
Bruising
Poor wound healing
Pigmentation
Skin infections
Hypertension
Osteoporosis
Pathological fractures (especially vertebrae and ribs)
Kyphosis
'Buffalo hump' (dorsal fat pad)
Central obesity
Striae (purple or red)
Rib fractures
Oedema
Proximal myopathy
Proximal muscle wasting
Glycosuria

Fig. 18.25 **Cushing's syndrome – symptoms and signs.** Bold type indicates signs of most value in discriminating Cushing's syndrome from simple obesity and hirsutism.

warning the patient). Normal subjects show a pronounced circadian variation (see Fig. 18.3, p. 968); those with Cushing's syndrome have high midnight cortisol levels (>100 nmol/L), though the 0900 h value may be normal. Midnight salivary cortisol collected at home gives the same information more simply where the assay is available.

- **Other tests**. There are frequent exceptions to the classic responses to diagnostic tests in Cushing's syndrome. If any clinical suspicion of Cushing's remains after preliminary tests then specialist investigations are still indicated. These may include insulin stress test, desmopressin stimulation test (p. 1017) and CRH tests.

2. Differential diagnosis of the cause

This can be extremely difficult since all causes can result in clinically identical Cushing's syndrome. The classical ectopic ACTH syndrome is distinguished by a short history, pigmentation and weight loss, unprovoked hypokalaemia, clinical or chemical diabetes and plasma ACTH levels above 200 ng/L, but many ectopic tumours are benign and mimic pituitary disease closely, both clinically and biochemically. Severe hirsutism/virilization suggests an adrenal tumour.

Biochemical and radiological procedures for diagnosis include:

- **Adrenal CT or MRI scan**. Adrenal adenomas and carcinomas causing Cushing's syndrome are relatively large and always detectable by CT scan. Carcinomas are distinguished by large size, irregular outline and signs of infiltration or metastases. Bilateral adrenal hyperplasia may be seen in ACTH-dependent causes or in ACTH-independent nodular hyperplasia.
- **Pituitary MRI**. A pituitary adenoma may be seen but the adenoma is often small and not visible in a significant proportion of cases.
- **Plasma potassium levels**. Hypokalaemia is common with ectopic ACTH secretion. (All diuretics must be stopped.)
- **High-dose dexamethasone** test (Table 18.26). Failure of significant plasma cortisol suppression suggests an ectopic source of ACTH or an adrenal tumour.
- **Plasma ACTH levels**. Low or undetectable ACTH levels (<10 ng/L) on two or more occasions are a reliable indicator of non-ACTH-dependent disease.
- **CRH test.** An exaggerated ACTH and cortisol response to exogenous CRH suggests pituitary-dependent Cushing's disease, as ectopic sources rarely respond.
- **Chest X-ray** to look for a carcinoma of the bronchus or a bronchial carcinoid. Carcinoid lesions may be very small; if ectopic ACTH is suspected, whole-lung, mediastinal and abdominal CT scanning should be performed.

Further investigations may involve selective catheterization of the inferior petrosal sinus to measure ACTH for pituitary lesions, or blood samples taken throughout the body in a search for ectopic sources. Radiolabelled octreotide (^{111}In octreotide) is occasionally helpful in locating ectopic ACTH sites.

Treatment

Untreated Cushing's syndrome has a very bad prognosis, with death from hypertension, myocardial infarction, infection and heart failure. Whatever the underlying cause, cortisol hypersecretion should be controlled prior to surgery or radiotherapy. Considerable morbidity and mortality is otherwise associated with operating on unprepared patients, especially when abdominal surgery is required. The usual drug is metyrapone, an 11β-hydroxylase blocker, which is given in doses of 750 mg to 4 g daily in three to four divided doses. Ketoconazole (200 mg three times daily) is also used and is synergistic with metyrapone. Plasma cortisol should be monitored, aiming to reduce the mean level during the day

to 150–300 nmol/L, equivalent to normal production rates. Aminoglutethimide and trilostane (which reversibly inhibits 3-hydroxysteroid dehydrogenase/δ-5,4 isomers) are occasionally used.

Choice of further treatment depends upon the cause.

Cushing's disease (pituitary-dependent hyperadrenalism)

- *Trans-sphenoidal removal of the tumour* is the treatment of choice. Selective adenomectomy nearly always leaves the patient ACTH deficient immediately postoperatively, and this is a good prognostic sign. Overall, pituitary surgery results in remission in 75–80% of cases, but results vary considerably and an experienced surgeon is essential.
- *External pituitary irradiation* alone is slow acting, only effective in 50–60% even after prolonged follow-up and mainly used after failed pituitary surgery. Children, however, respond much better to radiotherapy, 80% being cured.
- *Medical therapy* to reduce ACTH (e.g. bromocriptine, cyproheptadine) is rarely effective.
- *Bilateral adrenalectomy* is an effective last resort if other measures fail to control the disease (see Nelson's syndrome). This can be performed laparoscopically.

Other causes

Adrenal adenomas should be resected after achievement of clinical remission with metyrapone or ketoconazole. Contralateral adrenal suppression may last for a year or more.

Adrenal carcinomas are highly aggressive and the prognosis is poor. In general, if there are no widespread metastases, tumour bulk should be reduced surgically. The adrenolytic drug mitotane may inhibit growth of the tumour and prolong survival, though it can cause nausea and ataxia. Some would also give radiotherapy to the tumour bed after surgery.

Tumours secreting ACTH ectopically should be removed if possible. Otherwise chemotherapy/radiotherapy may be used, depending on the tumour. Control of the Cushing's syndrome with metyrapone or ketoconazole is beneficial for symptoms, and bilateral adrenalectomy may be appropriate to give complete control of the Cushing's syndrome if prognosis from the tumour itself is reasonable.

If the source of ACTH is not clear, cortisol hypersecretion should be controlled with medical therapy until a diagnosis can be made.

Nelson's syndrome

Nelson's syndrome is increased pigmentation (because of high levels of ACTH) associated with an enlarging pituitary tumour, which occurs in about 20% of cases after bilateral adrenalectomy for Cushing's disease. The syndrome is rare now that adrenalectomy is an uncommon primary treatment, and its incidence may be reduced by pituitary radiotherapy soon after adrenalectomy. The Nelson's adenoma may be treated by pituitary surgery and/or radiotherapy (unless given previously).

Incidental adrenal tumours ('incidentalomas')

With the advent of abdominal CT, MRI and high-resolution ultrasound scanning, unsuspected adrenal masses have been discovered in 3–10% of scans (increasing with age). These obviously include the adrenal tumours described above, but cysts, myelolipomas and metastases are also seen. Functional tests to exclude secretory activity should be performed (adenomas often secrete cortisol at a low level); if none is found then most authorities recommend removal of large (>4–5 cm) and functional tumours but observation of smaller hormonally inactive lesions. Phaeochromocytoma must be excluded before surgery due to the risk of perioperative hypertensive or hypotensive crises (see p. 1025).

Congenital adrenal hyperplasia (CAH)

Pathophysiology

This condition results from an autosomal recessive deficiency of an enzyme in the cortisol synthetic pathways. There are six major types, but most common is 21-hydroxylase deficiency (CYP21A2) which occurs in about 1 in 15000 births and which has been shown to be due to defects on chromosome 6 near the HLA region affecting one of the cytochrome p450 enzymes (p450C21).

As a result, cortisol secretion is reduced and feedback leads to increased ACTH secretion to maintain adequate cortisol – leading to adrenal hyperplasia. Diversion of the steroid precursors into the androgenic steroid pathways occurs (see Fig. 18.22a). Thus, 17-hydroxyprogesterone, androstenedione and testosterone levels are increased, leading to virilization. Aldosterone synthesis may be impaired with resultant salt wasting.

The other forms affect 11β-hydroxylase, 17α-hydroxylase, 3β-hydroxysteroid dehydrogenase and a cholesterol side-chain cleavage enzyme ($p450_{scc}$) (see Fig. 18.22a).

Clinical features

If severe, CAH presents at birth with sexual ambiguity or adrenal failure (collapse, hypotension, hypoglycaemia), sometimes with a salt-losing state (hypotension, hyponatraemia). In the female, clitoral hypertrophy, urogenital abnormalities and labioscrotal fusion are common, but the syndrome may be unrecognized in the male.

Precocious puberty with hirsutism is a later presentation, whereas rare, milder cases only present in adult life, usually accompanied by primary amenorrhoea. Hirsutism developing before puberty is suggestive of CAH.

Investigations

Expert advice is essential in the confirmation and differential diagnosis of 21-hydroxylase deficiency, and with ambiguous genitalia such advice must be sought urgently before any assignment of gender is made.

A profile of adrenocortical hormones is measured before and one hour after ACTH administration.

- 17-Hydroxyprogesterone levels are increased.
- Urinary pregnanetriol excretion is increased.
- Androstenedione levels are raised
- Basal ACTH levels are raised.

Treatment

Glucocorticoid activity must be replaced, as must mineralocorticoid activity if deficient. In CAH the larger dose of glucocorticoid is given at night to suppress the morning ACTH peak with a smaller dose in the morning (cf. Addison's disease, p. 1012; Table 18.28). Correct dosage is often difficult to establish in the child but should ensure normal 17-

hydroxyprogesterone levels while allowing normal growth; excessive replacement leads to stunting of growth. In adults, clinical features and biochemistry (plasma renin and 17-OH-progesterone) are used to modify treatment. Genetic counselling (p. 45) and antenatal diagnosis is essential, particularly in 21-hydroxylase deficiency. The mother of an affected fetus can take dexamethasone daily to prevent virilization.

Uses and problems of therapeutic steroid therapy

Apart from their use as therapeutic replacement for endocrine deficiency states, synthetic glucocorticoids are widely used for many non-endocrine conditions (Box 18.8). Short-term use (e.g. for acute asthma) carries only small risks of significant side-effects except for the simultaneous suppression of immune responses. The danger lies in their continuance, often through medical oversight or patient default. In general, therapy for 3 weeks or less, or a dose of prednisolone less than 5 mg per day, will not result in significant long-term suppression of the normal adrenal axis.

Long-term therapy with synthetic or natural steroids will, in most respects, mimic endogenous Cushing's syndrome. Exceptions are the relative absence of hirsutism, acne, hypertension and severe sodium retention, as the common synthetic steroids have low androgenic and mineralocorticoid activity.

Excessive doses of steroids may also be absorbed from skin when strong dermatological preparations are used, but inhaled steroids rarely cause Cushing's syndrome, although they commonly cause adrenal suppression.

The major hazards are detailed in Box 18.9. In the long term, many are of such severity that the clinical need for high-dose steroids should be continually and critically assessed. Steroid-sparing agents (e.g. azathioprine) should always be considered and screening and prophylactic therapy for osteoporosis introduced (see p. 565). New targeted biological therapies for inflammatory conditions may reduce the incidence of steroid-induced adrenal suppression.

Supervision of steroid therapy

All patients receiving steroids should carry a 'Steroid Card'. They should be made aware of the following points:

- Long-term steroid therapy must never be stopped suddenly.
- Doses should be reduced very gradually, with most being given in the morning at the time of withdrawal –

Box 18.8 Common therapeutic uses of glucocorticoids

Respiratory disease

Asthma
Chronic obstructive pulmonary disease
Sarcoidosis
Hayfever (usually topical)
Prevention/treatment of ARDS (see p. 917)

Cardiac disease

Post-myocardial infarction syndrome

Renal disease

Some nephrotic syndromes
Some glomerulonephritides

Gastrointestinal disease

Ulcerative colitis
Crohn's disease
Autoimmune hepatitis

Rheumatological disease

Systemic lupus erythematosus
Polymyalgia rheumatica
Cranial arteritis
Juvenile idiopathic arthritis
Vasculitides
Rheumatoid arthritis

Neurological disease

Cerebral oedema

Skin disease

Pemphigus, eczema

Tumours

Hodgkin's lymphoma
Other lymphomas

Transplantation

Immunosuppression

Box 18.9 Major adverse effects of corticosteroid therapy

Physiological

Adrenal and/or pituitary suppression

Pathological

Cardiovascular
Increased blood pressure

Gastrointestinal
Pancreatitis

Renal
Polyuria
Nocturia

Central nervous
Depression
Euphoria
Psychosis
Insomnia

Endocrine
Weight gain
Glycosuria/hyperglycaemia/diabetes
Impaired growth
Amenorrhoea

Bone and muscle
Osteoporosis
Proximal myopathy and wasting
Aseptic necrosis of the hip
Pathological fractures

Skin
Thinning
Easy bruising

Eyes
Cataracts (including inhaled drug)

Increased susceptibility to infection (signs and fever are frequently masked)
Septicaemia
Fungal infections
Reactivation of TB
Skin (e.g. fungi)

Table 18.30 Steroid cover for operative procedures

Procedure	Premedication	Intra- and postoperative	Resumption of normal maintenance
Simple procedures (e.g. gastroscopy, simple dental extractions)	Hydrocortisone 100 mg i.m	–	Immediately if no complications and eating normally
Minor surgery (e.g. laparoscopic surgery, veins, hernias)	Hydrocortisone 100 mg i.m.	Hydrocortisone 20 mg orally 6-hourly or 50 mg i.m. every 6 hours for 24 h if not eating	After 24 h if no complications
Major surgery (e.g. hip replacement, vascular surgery)	Hydrocortisone 100 mg i.m.	Hydrocortisone 50–100 mg i.m every 6 hours for 72 h	After 72 h if normal progress and no complications Perhaps double normal dose for next 2–3 days
GI tract surgery or major thoracic surgery (not eating or ventilated)	Hydrocortisone 100 mg i.m.	Hydrocortisone 100 mg i.m. every 6 hours for 72 h or longer if still unwell	When patient eating normally again Until then, higher doses (to 50 mg 6-hourly) may be needed

A useful summary of surgical steroid guidelines can be found at: http://www.addisons.org.uk/comms/publications/surgicalguidelines-colour.pdf

this minimizes adrenal suppression. Many authorities believe that 'alternate-day therapy' produces less suppression.

- Doses need to be increased in times of serious intercurrent illness (defined as presence of a fever), accident and stress. Double doses should be taken during these times.
- Other physicians, anaesthetists and dentists must be told about steroid therapy.
- Patients should also be informed of potential side-effects and all this information should be documented in the clinical record.
- Rationale for prophylactic use of bisphosphonate therapy to prevent the development of osteoporosis (NICE guidance).

Steroids and surgery

Any patient receiving steroids or who has recently received them (within the last 12 months) and may still have adrenal suppression requires careful control of steroid medication around the time of surgery. Details are shown in Table 18.30.

FURTHER READING

Arlt W, Allolio B. Adrenal insufficiency. *Lancet* 2003; **361**: 1881–1893.

Newell-Price J, Bertagna X, Grossman AB et al. Cushing's syndrome. *Lancet* 2006; **367**: 1605–1617.

Speiser PW, White PC. Congenital adrenal hyperplasia. *New England Journal of Medicine* 2003; **349**: 776–788.

Young WF. The incidentally discovered adrenal mass. *New England Journal of Medicine* 2007; **356**: 601–610.

THE THIRST AXIS

Thirst and water regulation are largely controlled by vasopressin, also known as antidiuretic hormone (ADH), which is synthesized in the hypothalamus and then migrates in neurosecretory granules along axonal pathways to the posterior pituitary. Pituitary disease alone without hypothalamic involvement therefore does not lead to ADH deficiency as the hormone can still 'leak' from the damaged end of the intact axon.

At normal concentrations the kidney is the predominant site of action of vasopressin. Vasopressin stimulation of the V_2 receptors allows the collecting ducts to become permeable to water via the migration of aquaporin-2 water channels, thus permitting reabsorption of hypotonic luminal fluid (p. 654). Vasopressin therefore reduces diuresis and results in overall retention of water. At high concentrations vasopressin also causes vasoconstriction via the V_1 receptors in vascular tissue.

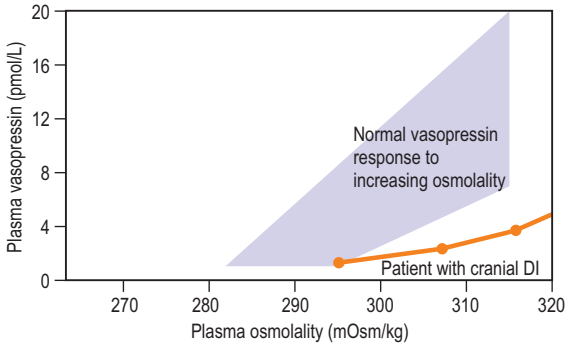


Fig. 18.26 **Plasma vasopressin response to increasing osmolality in normal subjects and in a patient with diabetes insipidus.**

Changes in plasma osmolality are sensed by osmoreceptors in the anterior hypothalamus. Vasopressin secretion is suppressed at levels below 280 mOsm/kg, thus allowing maximal water diuresis. Above this level, plasma vasopressin increases in direct proportion to plasma osmolality. At the upper limit of normal (295 mOsm/kg) maximum antidiuresis is achieved and thirst is experienced at about 298 mOsm/kg (Fig. 18.26).

Other factors affecting vasopressin release are shown in Table 18.31.

Disorders of vasopressin secretion or activity include:

- deficiency as a result of hypothalamic disease ('cranial' diabetes insipidus)
- inappropriate excess of the hormone
- 'nephrogenic' diabetes insipidus – a rare condition in which the renal tubules are insensitive to vasopressin, an example of a receptor abnormality.

While all these are uncommon, they need to be distinguished from the occasional patient with 'primary polydipsia' and those whose renal tubular function has been impaired by

Table 18.31 Factors affecting vasopressin release

Increased by:	Decreased by:
Increased osmolality	Decreased osmolality
Hypovolaemia	Hypervolaemia
Hypotension	Hypertension
Nausea	Ethanol
Hypothyroidism	α-Adrenergic stimulation
Angiotensin II	
Epinephrine (adrenaline)	
Cortisol	
Nicotine	
Antidepressants	

electrolyte abnormalities, such as hypokalaemia or hypercalcaemia.

Diabetes insipidus (DI)

Clinical features

Deficiency of vasopressin or insensitivity to its action leads to polyuria, nocturia and compensatory polydipsia. Daily urine output may reach as much as 10–15 L, leading to dehydration that may be very severe if the thirst mechanisms or consciousness are impaired or the patient is denied fluid.

Causes of DI are listed in Table 18.32. The most common is hypothalamic–pituitary surgery, following which transient DI is common, frequently remitting after a few days or weeks.

DI may be masked by simultaneous cortisol deficiency – cortisol replacement allows a water diuresis and DI then becomes apparent.

Familial isolated vasopressin deficiency causes DI from early childhood and is dominantly inherited. *DIDMOAD (Wolfram) syndrome* is a rare autosomal recessive disorder comprising Diabetes Insipidus, Diabetes Mellitus, Optic Atrophy and Deafness due to mutations in the *WFS1* gene on chromosome 4. MR scanning may show an absent or poorly developed posterior pituitary.

Biochemistry

- High or high-normal plasma osmolality with low urine osmolality (in primary polydipsia plasma osmolality tends to be low).
- Resultant high or high-normal plasma sodium.
- High 24-h urine volumes (less than 2 L excludes the need for further investigation).
- Failure of urinary concentration with fluid deprivation.
- Restoration of urinary concentration with vasopressin or an analogue.

The latter two points are studied with a formal water-deprivation test (Box 18.10). In normal subjects, plasma osmolality remains normal while urine osmolality rises above 600 mOsm/kg. In DI, plasma osmolality rises while the urine remains dilute, only concentrating after exogenous vasopressin is given (in 'cranial' DI) or not concentrating after vasopressin if nephrogenic DI is present. An alternative is measurement of plasma vasopressin during hypertonic saline infusion, but these measurements are not widely available.

Table 18.32 Causes of diabetes insipidus

Cranial diabetes insipidus

Familial (e.g. DIDMOAD)

Idiopathic (often autoimmune)

Tumours
- Craniopharyngioma
- Hypothalamic tumour, e.g. glioma, germinoma
- Metastases, especially breast
- Lymphoma/leukaemia
- Pituitary with suprasellar extension (rare)

Infections
- Tuberculosis
- Meningitis
- Cerebral abscess

Infiltrations
- Sarcoidosis
- Langerhans' cell histiocytosis

Inflammatory
- Hypophysitis

Post-surgical
- Transfrontal
- Trans-sphenoidal

Post-radiotherapy (cranial)

Vascular
- Haemorrhage/thrombosis
- Sheehan's syndrome
- Aneurysm

Trauma (e.g. head injury)

Nephrogenic diabetes insipidus

Familial (e.g. vasopressin receptor gene, aquaporin-2 gene defect)

Idiopathic

Renal disease (e.g. renal tubular acidosis)

Hypokalaemia

Hypercalcaemia

Drugs (e.g. lithium, demeclocycline, glibenclamide)

Sickle cell disease

Mild temporary nephrogenic DI can occur after prolonged polyuria due to any cause, including cranial DI and primary polydipsia.

Box 18.10 Water-deprivation test

Indication

Diagnosis or exclusion of diabetes insipidus

Procedure

Fasting and no fluids from 0730 h (or overnight if only mild DI is expected and polyuria is only modest)
Monitor serum and urine osmolality, urine volume and weight hourly for up to 8 hours
Abandon fluid deprivation if weight loss >3% occurs
If serum osmolality >300 mOsm/kg and/or urine osmolality <600 mOsm/kg give desmopressin 2 μg i.m. at end of test. Allow free fluid but measure urine osmolality for 2–4 hours

Interpretation

Normal response. Serum osmolality remains within normal range (275–295 mOsm/kg). Urine osmolality rises to >600 mOsm/kg.
Diabetes insipidus (DI). Serum osmolality rises above normal without adequate concentration of urine osmolality (i.e. serum osmolality >300 mOsm/kg; urine osmolality <600 mOsm/kg).
Nephrogenic DI – if DDAVP does not concentrate urine.
Cranial DI – if urine osmolality rises by >50% after desmopressin.

Treatment

The synthetic vasopressin analogue desmopressin is the treatment of choice. It is most reliably given intranasally as a spray 10–40 μg once or twice daily, but can also be given orally as 100–200 μg three times daily, or intramuscularly 2–4 μg daily. Response is variable and must be monitored carefully with enquiry about fluid input/output and plasma osmolality measurements. Where there is a reversible underlying cause (e.g. a hypothalamic tumour) this should be investigated and treated.

Alternative agents in mild DI, probably working by sensitizing the renal tubules to endogenous vasopressin, include thiazide diuretics, carbamazepine (200–400 mg daily) or chlorpropamide (200–350 mg daily) but these are rarely used.

Nephrogenic diabetes insipidus

In this condition, renal tubules are resistant to normal or high levels of plasma vasopressin. It may be inherited as a rare sex-linked recessive, with an abnormality in the vasopressin-2 receptor, or as an autosomal post-receptor defect in an ADH-sensitive water channel, aquaporin-2. More commonly it can be acquired as a result of renal disease, sickle cell disease, drug ingestion (e.g. lithium), hypercalcaemia or hypokalaemia. Wherever possible the cause should be reversed. Polyuria is helped by thiazide diuretics.

Other causes of polyuria and polydipsia

Diabetes mellitus, hypokalaemia and hypercalcaemia should be excluded. In the case of diabetes mellitus the cause is an osmotic diuresis secondary to glycosuria which leads to dehydration and an increased perception of thirst owing to hypertonicity of the extracellular fluid.

Primary polydipsia

This is a relatively common cause of thirst and polyuria. It is a psychiatric disturbance characterized by the excessive intake of water. Plasma sodium and osmolality fall as a result and the urine produced is appropriately dilute. Vasopressin levels become virtually undetectable. Prolonged primary polydipsia may lead to the phenomenon of 'renal medullary washout', with a fall in the concentrating ability of the kidney.

Characteristically the diagnosis is made by a water-deprivation test. A low plasma osmolality is usual at the start of the test, and since vasopressin secretion and action can be stimulated, the patient's urine becomes concentrated (albeit 'maximum' concentrating ability may be impaired); the initially low urine osmolality gradually increases with the duration of the water deprivation.

Syndrome of inappropriate antidiuretic hormone (SIADH)

Clinical features

Inappropriate secretion of ADH leads to retention of water and hyponatraemia. The presentation is usually vague, with confusion, nausea, irritability and, later, fits and coma. There is no oedema. Mild symptoms usually occur with plasma sodium levels below 125 mmol/L and serious manifestations are likely below 115 mmol/L. The elderly may show symptoms with milder abnormalities.

The syndrome must be distinguished from dilutional hyponatraemia due to excess infusion of dextrose/water solutions or diuretic administration (thiazides or amiloride, see p. 657).

Table 18.33 Common causes of the syndrome of inappropriate ADH secretion (SIADH)

Tumours
Small-cell carcinoma of lung
Prostate
Thymus
Pancreas
Lymphomas
Pulmonary lesions
Pneumonia
Tuberculosis
Lung abscess
CNS causes
Meningitis
Tumours
Head injury
Subdural haematoma
Cerebral abscess
SLE vasculitis
Metabolic causes
Alcohol withdrawal
Porphyria
Drugs
Chlorpropamide
Carbamazepine
Cyclophosphamide
Vincristine
Phenothiazines

Diagnosis

The usual features are:

- dilutional hyponatraemia due to excessive water retention
- euvolaemia (in contrast to hypovolaemia of sodium and water depletion states)
- low plasma osmolality with 'inappropriate' urine osmolality >100 mOsm/kg (and typically higher than plasma osmolality)
- continued urinary sodium excretion >30 mmol/L (lower levels suggest sodium depletion and should respond to 0.9% saline infusion)
- absence of hypokalaemia (or hypotension)
- normal renal and adrenal and thyroid function.

The causes are listed in Table 18.33.

Hyponatraemia is very common during illness in frail elderly patients and it may sometimes be clinically difficult to distinguish SIADH from salt and water depletion, particularly when mixed clinical features are present. Under these circumstances a trial infusion of 1–2 L 0.9% saline is given. SIADH will not respond (but will excrete the sodium and water load effectively) – sodium depletion will respond.

Treatment

The underlying cause should be corrected where possible. Symptomatic relief can be obtained by the following measures:

- Fluid intake should be restricted to 500–1000 mL daily. If tolerated, and complied with, this will correct the biochemical abnormalities in almost every case.
- Plasma osmolality, serum sodium and bodyweight should be measured frequently.
- If water restriction is poorly tolerated or ineffective, demeclocycline (600–1200 mg daily) is given; this inhibits the action of vasopressin on the kidney, causing a reversible form of nephrogenic diabetes insipidus. It often, however, causes photosensitive rashes.
- When the syndrome is very severe (i.e. acute and symptomatic), hypertonic saline may be indicated but this is potentially dangerous and should only be used with extreme caution (p. 663).
- Vasopressin V_2 antagonists, e.g. tolvaptan, are being used with good initial results.

DISORDERS OF CALCIUM METABOLISM

Serum calcium levels are mainly controlled by parathyroid hormone (PTH) and vitamin D. Hypercalcaemia is much more common than hypocalcaemia and is frequently detected incidentally with multichannel biochemical analysers. Mild asymptomatic hypercalcaemia occurs in about 1 in 1000 of the population, with an incidence of 25–30 per 100 000 population. It occurs mainly in elderly females, and is usually due to primary hyperparathyroidism (primary HPT).

Parathyroid hormone

There are normally four parathyroid glands which are situated posterior to the thyroid, but occasionally additional glands exist or they may be found elsewhere in the neck or mediastinum. PTH, an 84-amino-acid hormone derived from a 115-residue preprohormone, is secreted from the chief cells of the parathyroid glands. PTH levels rise as serum ionized calcium falls. The latter is detected by specific G protein-coupled calcium-sensing receptors on the plasma membrane of the parathyroid cells. PTH has several major actions, all serving to increase plasma calcium by:

- increasing osteoclastic resorption of bone (occurring rapidly)
- increasing intestinal absorption of calcium (a slow response)
- increasing synthesis of $1,25\text{-}(OH)_2D_3$
- increasing renal tubular reabsorption of calcium
- increasing excretion of phosphate.

PTH effects are mediated at specific membrane receptors on the target cells, resulting in an increase of adenyl cyclase messenger activity.

Vitamin D metabolism is discussed on page 559.

PTH measurements

PTH measurements use two-site immunometric assays that measure only the intact PTH molecule; interpretation requires a simultaneous calcium measurement in order to differentiate most causes of hyper- and hypocalcaemia.

Table 18.34 Causes of hypercalcaemia

Excessive parathormone (PTH) secretion
Primary hyperparathyroidism (commonest by far), adenoma (common), hyperplasia or carcinoma (rare)
Tertiary hyperparathyroidism
Ectopic PTH secretion (very rare indeed)
Malignant disease (second commonest cause)
Myeloma
Secondary deposits in bone
Production of osteoclastic factors by tumours
PTH-related protein secretion
Excess action of vitamin D
Iatrogenic or self-administered excess
Granulomatous diseases, e.g. sarcoidosis, TB
Lymphoma
Excessive calcium intake
'Milk–alkali' syndrome
Other endocrine disease (mild hypercalcaemia only)
Thyrotoxicosis
Addison's disease
Drugs
Thiazide diuretics
Vitamin D analogues
Lithium administration (chronic)
Vitamin A
Miscellaneous
Long-term immobility
Familial hypocalciuric hypercalcaemia

Hypercalcaemia

Pathophysiology and causes

The major causes of hypercalcaemia are listed in Table 18.34; primary hyperparathyroidism and malignancies are by far the most common (>90% of cases). Hyperparathyroidism itself may be primary, secondary or tertiary. Primary hyperparathyroidism is caused by single (>80%) parathyroid adenomas or by diffuse hyperplasia of all the glands (15–20%); multiple parathyroid adenomas are rare. Involvement of multiple parathyroid glands may be part of a familial syndrome (e.g. multiple endocrine neoplasia (MEN) syndrome type 1 or 2a). Parathyroid carcinoma is rare (<1%), though it usually produces severe and intractable hypercalcaemia. Hyperparathyroidism–jaw tumour syndrome is a rare familial cause of hyperparathyroidism which may be associated with parathyroid carcinoma and maxillary or mandibular tumours.

Primary hyperparathyroidism is of unknown cause, though it appears that adenomas are monoclonal. Hyperplasia may also be monoclonal. Chromosomal rearrangements in the 5′ regulatory region of the parathyroid hormone gene have been identified, and inactivation of some tumour suppressor genes at a variety of sites may also be involved.

Secondary hyperparathyroidism (see p. 628) is physiological compensatory hypertrophy of all parathyroids because of hypocalcaemia, such as occurs in renal failure or vitamin D deficiency. PTH levels are raised but calcium levels are low or normal, and PTH falls to normal after correction of the cause of hypocalcaemia where this is possible.

Tertiary hyperparathyroidism is the development of apparently autonomous parathyroid hyperplasia after long-standing secondary hyperparathyroidism, most often in renal failure. Plasma calcium and phosphate are both raised, the latter often grossly so. Parathyroidectomy is necessary at this stage.

Symptoms and signs

Mild hypercalcaemia (e.g. adjusted calcium <3 mmol/L) is frequently asymptomatic, but more severe hypercalcaemia can produce a number of symptoms:

- *General features*. There may be tiredness, malaise, dehydration and depression.
- *Renal features*. Renal colic from stones, polyuria or nocturia, haematuria and hypertension occurs. The polyuria results from the effect of hypercalcaemia on renal tubules, reducing their concentrating ability – a form of mild nephrogenic diabetes insipidus. Primary hyperparathyroidism is present in about 5% of patients who present with renal calculi.
- *Bones*. There may be bone pain. Hyperparathyroidism mainly affects cortical bone, and bone cysts and locally destructive 'brown tumours' occur but only in advanced disease. Only 5–10% of all cases have definite bony lesions even when sought. Bone disease may be more apparent when there is coexisting vitamin D deficiency.
- *Abdominal*. There may be abdominal pain.
- *Chondrocalcinosis* and ectopic calcification. These are occasional features.
- *Corneal calcification*. This is a marker of long-standing hypercalcaemia but causes no symptoms.

There may also be symptoms from the underlying cause. Malignant disease is usually advanced by the time hypercalcaemia occurs, typically with bony metastases. The common primary tumours are bronchus, breast, myeloma, oesophagus, thyroid, prostate, lymphoma and renal cell carcinoma. True 'ectopic PTH secretion' by the tumour is very rare, and most cases are associated with raised levels of PTH-related protein. This is a 144-amino-acid polypeptide, the initial sequence of which shows an approximate homology with the biologically active part of PTH, which is necessary in fetal development but does not have a clearly defined role in the adult. Local bone-resorbing cytokines and prostaglandins may be involved locally where there are metastatic skeletal lesions, leading to local mobilization of calcium by osteolysis with subsequent hypercalcaemia.

Severe hypercalcaemia (>3 mmol/L) is usually associated with malignant disease, hyperparathyroidism, renal failure or vitamin D therapy.

Investigations and differential diagnosis

Biochemistry

- Several fasting serum calcium and phosphate samples should be performed.
- *Serum PTH*. The hallmark of primary hyperparathyroidism is hypercalcaemia and hypophosphataemia with detectable or elevated intact PTH levels during hypercalcaemia. When this combination is present in an asymptomatic patient then further investigation is usually unnecessary.
- There is often a mild *hyperchloraemic acidosis*.
- *Renal function* is usually normal but should be measured as a baseline.
- *24-hour urinary calcium* or single calcium creatinine ratio should be measured in a young patient with modest elevation in calcium and PTH to exclude familial hypocalciuric hypercalcaemia (see p. 1021).
- *Elevated serum alkaline phosphatase* is found in severe parathyroid bone disease, but otherwise it suggests an alternative cause for hypercalcaemia.

Where PTH is undetectable or equivocal, a number of other tests may lead to the diagnosis:

- protein electrophoresis: to exclude myeloma
- serum TSH: to exclude hyperthyroidism
- 0900 h cortisol and/or ACTH test: to exclude Addison's disease
- serum ACE: helpful in the diagnosis of sarcoidosis
- hydrocortisone suppression test: hydrocortisone 40 mg three times daily for 10 days leads to suppression of plasma calcium in sarcoidosis, vitamin D-mediated hypercalcaemia and some malignancies.

Imaging

Abdominal X-rays may show renal calculi or nephrocalcinosis. High-definition hand X-rays can show subperiosteal erosions in the middle or terminal phalanges. DXA bone density scan is useful to detect bone effects in asymptomatic patients with hyperparathyroidism (HPT) in whom conservative management is planned.

The success of parathyroid imaging is highly operator dependent and choice therefore depends on local skills and experience. Imaging is frequently far less accurate than parathyroid exploration by an expert surgeon where the success rate is at least 90%. However, with the advent of minimally invasive surgery, preoperative imaging is increasingly used. Methods include:

- ultrasound which, though insensitive for small tumours, is simple and safe
- high-resolution CT scan or MRI (more sensitive)
- radioisotope scanning using ^{99m}Tc-sestamibi, which is approximately 90% sensitive in detecting adenomas.

Treatment of hypercalcaemia

Details of emergency treatment for severe hypercalcaemia are given in Emergency box 18.2. This should be followed by oral therapy unless the underlying disease can be treated.

Treatment of primary hyperparathyroidism

Medical management

There are no effective medical therapies at present for primary hyperparathyroidism, but a high fluid intake should be maintained, a high calcium or vitamin D intake avoided, and exercise encouraged. New therapeutic agents that target the calcium-sensing receptors (e.g. cinacalcet) are of proven value in parathyroid carcinoma and in dialysis patients (p. 633).

Surgery

Indications for surgery in primary hyperparathyroidism remain controversial. There is agreement that surgery is indicated for:

- patients with renal stones or impaired renal function
- bone involvement or marked reduction in cortical bone density

Emergency Box 18.2

Treatment of acute severe hypercalcaemia

Acute hypercalcaemia often presents with dehydration, nausea and vomiting, nocturia and polyuria, drowsiness and altered consciousness. The serum Ca^{2+} is over 3 mmol/L and sometimes as high as 5 mmol/L. While investigation of the cause is under way, immediate treatment is mandatory if the patient is seriously ill or if the Ca^{2+} is above 3.5 mmol/L.

- Rehydrate at least 4–6 L of 0.9% saline on day 1, and 3–4 L for several days thereafter. Central venous pressure (CVP) may need to be monitored to control the hydration rate.
- Intravenous bisphosphonates are the treatment of choice for hypercalcaemia of malignancy or of undiagnosed cause. Pamidronate is preferred (60–90 mg as an intravenous infusion in 0.9% saline or dextrose over 2–4 hours or, if less urgent, over 2–4 days). Levels fall after 24–72 hours, lasting for approximately two weeks. Zoledronate is an alternative.
- Prednisolone (30–60 mg daily) is effective in some instances (e.g. in myeloma, sarcoidosis and vitamin D excess) but in most cases is ineffective.
- Calcitonin (200 units i.v. 6-hourly) has a short-lived action and is little used.
- *Oral* phosphate (sodium cellulose phosphate 5 g three times daily) produces diarrhoea.

- unequivocal marked hypercalcaemia (in UK typically >3.0 mmol/L; USA guidelines state >1 mg/dL above reference range)
- the uncommon younger patient, below age 50 years
- a previous episode of severe acute hypercalcaemia.

The situation where plasma calcium is mildly raised (2.65–3.00 mmol/L) is more controversial. Most authorities feel that young patients should be operated on, as should those who have reduced cortical bone density or significant hypercalciuria, as this is associated with stone formation.

In older patients without these problems, or in those unfit for or unwilling to have surgery, conservative management is indicated. Regular measurement of serum calcium and of renal function is necessary. Bone density of cortical bone should be monitored if conservative management is used.

Surgical technique and complications

Parathyroid surgery should be performed only by experienced surgeons, as the minute glands may be very difficult to define, and it is difficult to distinguish between an adenoma and normal parathyroid. In expert centres over 90% of operations are successful, involving removal of the adenoma, or removal of all four hyperplastic parathyroids. Minimal access surgery is increasingly used, and some centres measure PTH levels intra-operatively to ensure the adenoma has been removed.

Other than postoperative hypocalcaemia (see below), the other rare complications are those of thyroid surgery – bleeding and recurrent laryngeal nerve palsies (<1%). Vocal cord function should be checked preoperatively.

If initial exploration is unsuccessful, a full work-up including venous catheterization and scanning is essential, remembering that parathyroid tissue can be ectopic.

Postoperative care

The major danger after operation is hypocalcaemia, which is more common in patients with significant bone disease – the 'hungry bone' syndrome. Some authorities pretreat such patients with alfacalcidol 2 µg daily from 2 days preoperatively for 10–14 days. Chvostek's and Trousseau's signs (see p. 1022) are monitored as well as biochemistry. Plasma calcium measurements are performed at least daily until stable – with or without replacement – a mild transient hypoparathyroidism often continues for 1–2 weeks. Depending on its severity, oral or intravenous calcium should be given temporarily, as only a few patients (<1%) will develop long-standing surgical hypoparathyroidism.

Table 18.35 Causes of hypocalcaemia

Increased phosphate levels
Chronic kidney disease (common)
Phosphate therapy
Hypoparathyroidism
Surgical – after neck exploration (thyroidectomy, parathyroidectomy – common)
Congenital deficiency (DiGeorge syndrome)
Idiopathic hypoparathyroidism (rare)
Severe hypomagnesaemia
Vitamin D deficiency
Osteomalacia/rickets
Vitamin D resistance
Resistance to PTH
Pseudohypoparathyroidism
Drugs
Calcitonin
Bisphosphonates
Other
Acute pancreatitis (quite common)
Citrated blood in massive transfusion (not uncommon)
Low plasma albumin, e.g. malnutrition, chronic liver disease
Malabsorption, e.g. coeliac disease

Familial hypocalciuric hypercalcaemia

This uncommon autosomal dominant, and usually asymptomatic, condition demonstrates increased renal reabsorption of calcium despite hypercalcaemia. PTH levels are normal or slightly raised and urinary calcium is low. It is caused by loss of function mutations in the gene on the long arm of chromosome 3 encoding for the calcium-ion-sensing G protein-coupled receptor in the kidney and parathyroid gland. Family members are often affected, detected by genetic analysis. Parathyroid surgery is not indicated as the course appears benign. This diagnosis can be differentiated from hyperparathyroidism in an isolated case by the calcium creatinine ratio in blood and urine.

Hypocalcaemia and hypoparathyroidism

Pathophysiology

Hypocalcaemia may be due to deficiencies of calcium homeostatic mechanisms, secondary to high phosphate levels or other causes of hypocalcaemia (Table 18.35). All forms of hypoparathyroidism, except transient surgical effects, are uncommon.

FURTHER READING

Bilezikian JP, Silverberg SJ. Asymptomatic primary hyperparathyroidism. *New England Journal of Medicine* 2004; **350**: 1746–1751.

Ask the authors
www.studentconsult.com

FURTHER READING

Shoback D. Hypoparathyroidism. *New England Journal of Medicine* 2008; **359**: 391–403.

Causes

- *Chronic Kidney disease* is the most common cause of hypocalcaemia.
- *Hypocalcaemia after thyroid or parathyroid surgery* is common but usually transient – fewer than 1% of thyroidectomies leave permanent damage (see above).
- *Idiopathic hypoparathyroidism* is one of the rarer autoimmune disorders, often accompanied by vitiligo, cutaneous candidiasis and other autoimmune disease.

The DiGeorge syndrome (p. 70) is a familial condition in which the hypoparathyroidism is associated with intellectual impairment, cataracts and calcified basal ganglia, and occasionally with specific autoimmune disease.

Pseudohypoparathyroidism is a syndrome of end-organ resistance to PTH owing to a mutation in the $G_S\alpha$-protein (GNAS1) which is coupled to the PTH receptor. It is associated with short stature, short metacarpals, subcutaneous calcification and sometimes intellectual impairment. Variable degrees of resistance involving other G protein-linked hormone receptors may also be seen (TSH, LH, FSH).

Pseudo-pseudohypoparathyroidism describes the phenotypic defects but without any abnormalities of calcium metabolism. These individuals may share the same gene defect as patients with pseudohypoparathyroidism and be members of the same families.

Clinical features

Hypoparathyroidism presents as neuromuscular irritability and neuropsychiatric manifestations. Paraesthesiae, circumoral numbness, cramps, anxiety and tetany (Box 18.11) are followed by convulsions, laryngeal stridor, dystonia and psychosis. Two signs of hypocalcaemia are Chvostek's sign (gentle tapping over the facial nerve causes twitching of the ipsilateral facial muscles) and Trousseau's sign, where inflation of the sphygmomanometer cuff above systolic pressure for 3 minutes induces tetanic spasm of the fingers and wrist. Severe hypocalcaemia may cause papilloedema and frequently a prolonged QT interval on the ECG.

Investigations

The clinical history and picture is usually diagnostic and is confirmed by a low serum calcium (after correction for any albumin abnormality). Additional tests include:

- **serum and urine creatinine** for renal disease
- **PTH levels** in the serum: absent or inappropriately low in hypoparathyroidism, high in other causes of hypocalcaemia
- **parathyroid antibodies** (present in idiopathic hypoparathyroidism)
- **25-hydroxy vitamin D serum level** (low in vitamin D deficiency)
- **magnesium level** – severe hypomagnesaemia results in functional hypoparathyroidism which is reversed by magnesium replacement.
- **X-rays** of metacarpals, showing short fourth metacarpals which occur in pseudohypoparathyroidism.

Box 18.11 Causes of tetany

In the presence of alkalosis

- Hyperventilation
- Excess antacid therapy
- Persistent vomiting
- Hypochloraemic alkalosis, e.g. primary hyperaldosteronism

In the presence of hypocalcaemia (see Table 18.35)

Treatment

Alpha-hydroxylated derivatives of vitamin D are preferred for their shorter half-life, and especially in renal disease as the others require renal hydroxylation. Usual daily maintenance doses are 0.25–2 μg for alfacalcidol (1α-OH-D_3). During treatment, plasma calcium must be monitored frequently to detect hypercalcaemia.

ENDOCRINOLOGY OF BLOOD PRESSURE CONTROL

The control of blood pressure (BP) is complex, involving neural, cardiac, hormonal and many other mechanisms.

BP is dependent upon cardiac output and peripheral resistance. Although cardiac output can be increased in endocrine disease (e.g. hyperthyroidism), the main role of hormonal mechanisms is control of peripheral resistance and of circulating blood volume. The oral contraceptive pill is a common endocrine cause of mild hypertension.

When to investigate for secondary hypertension

Endocrine causes account for 5–10% of all hypertension (Table 18.36). It is impracticable and unnecessary to screen all hypertensive patients for secondary causes. The highest chances of detecting such causes are in:

- subjects under 35 years, especially those without a family history of hypertension
- those with accelerated (malignant) hypertension
- those with indications of renal disease (e.g. proteinuria, unequal renal sizes)
- those with hypokalaemia before diuretic therapy

Table 18.36 Endocrine causes of hypertension

Excessive renin, and thus angiotensin II, production
Renal artery stenosis
Other local renal disease
Renin-secreting tumours
Excessive production of catecholamines
Phaeochromocytoma
Excessive GH production
Acromegaly
Excessive aldosterone production
Adrenal adenoma (Conn's syndrome)
Idiopathic adrenal hyperplasia
Dexamethasone-suppressible hyperaldosteronism
Excessive production of other mineralocorticoids
Cushing's syndrome (massive excess of cortisol, a weak mineralocorticoid)
Congenital adrenal hyperplasia (in rare cases)
Tumours producing other mineralocorticoids, e.g. corticosterone
Exogenous 'mineralocorticoids' or enzyme inhibitors
Liquorice ingestion (inhibits 11β-hydroxylase)
Abuse of mineralocorticoid preparations

- those resistant to conventional antihypertensive therapy (e.g. more than three drugs)
- those with unusual symptoms (e.g. sweating attacks or weakness).

The renin–angiotensin–aldosterone axis: biochemistry and actions

The renin–angiotensin–aldosterone system is illustrated in Figure 18.27.

Angiotensinogen, an α_2-globulin of hepatic origin, circulates in plasma. The enzyme, renin, is secreted by the kidney in response to decreased renal perfusion pressure or flow; it cleaves the decapeptide *angiotensin I* from angiotensinogen. Angiotensin I is inactive but is further cleaved by angiotensin-converting enzyme (ACE; present in lung and vascular endothelium) into the active peptide, *angiotensin II*, which has two major actions (mediated by two types of receptor, AT_1 and AT_2). The AT_1 subtype which is found in the heart, blood vessels, kidney, adrenal cortex, lung and brain mediates the vasoconstrictor effect. AT_2 is probably involved in vascular growth. Angiotension II:

- causes rapid, powerful vasoconstriction
- stimulates the adrenal zona glomerulosa to increase aldosterone production (over hours or days).

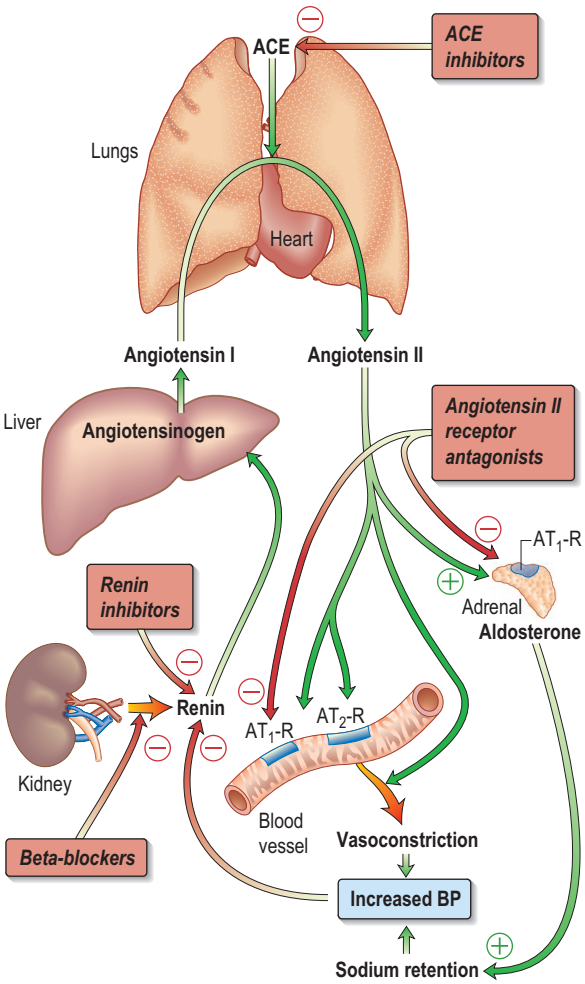


Fig. 18.27 **The renin–angiotensin–aldosterone system.** ACE inhibitors and angiotensin II antagonists can inhibit this system. ACE, angiotensin-converting enzyme.

As BP increases and sodium is retained, the stimuli to renin secretion are reduced. Dietary sodium excess also suppresses renin secretion, whereas sodium deprivation or urinary sodium loss will increase it.

The renin–angiotensin system can be blocked at several points with renin inhibitors, angiotensin-converting enzyme inhibitors (ACEI) and angiotensin II receptor antagonists (A-IIRA). These are useful agents in treatment of hypertension and heart failure (see pp. 738 and 739) but have differences in action: ACEIs also block kinin production while A-IIRAs are specific for the AT-II receptors.

Atrial and brain natriuretic factors/peptides (ANP and BNP)

Atrial natriuretic peptides, a family of varying length forms, are secreted from atrial granules in response to atrial stretch. They produce marked effects on the kidney, increasing sodium and water excretion and glomerular filtration rate and lowering BP, plasma renin activity and plasma aldosterone (p. 652).

Brain natriuretic peptide is found in the ventricle as well as the brain and has moderate sequence homology with ANP; normally its circulating level is much less than for ANP but may exceed it in congestive cardiac failure.

ANP and BNP appear to play a significant role in cardiovascular and fluid homeostasis, but there is no evidence of primary defects in their secretion causing disease. Serum BNP levels accurately reflect the presence and severity of heart failure and are useful in defining prognosis and the need for adjustment of treatment (p. 735).

Agents that both inhibit the endopeptidases that break down ANP and inhibit ACE (the receptors concerned are similar) are becoming available.

RENIN (AND ANGIOTENSIN) DEPENDENT HYPERTENSION

Many forms of unilateral and bilateral renal diseases are associated with hypertension. The classic example is renal artery stenosis: the major hypertensive effects of this and other situations such as renin-secreting tumours are directly or indirectly due to angiotensin II.

Angiotensin II receptor antagonists (e.g. losartan, valsartan, candesartan and irbesartan) are effective in hypertension and congestive cardiac failure, similar to angiotensin-converting enzyme inhibitors (ACEI). They produce much the same clinical effects, though with fewer side-effects (e.g. no cough and less hyperkalaemia). Direct renin inhibitors (e.g. aliskiren) have also been introduced for the treatment of hypertension.

Renal artery stenosis

This is discussed on page 607.

DISORDERS OF ALDOSTERONE SECRETION

Primary hyperaldosteronism

Pathophysiology

This condition is caused by excess aldosterone production leading to sodium retention, potassium loss and the combination of hypokalaemia and hypertension.

Causes (see Table 18.36)

Adrenal adenomas (Conn's syndrome) originally accounted for 60% of cases of primary hyperaldosteronism but represented a rare cause of hypertension. In recent series the increasing use of aldosterone:renin ratios in investigation of hypertension has led to an increased diagnosis of hyperaldosteronism due to bilateral adrenal hyperplasia of uncertain aetiology. Some claim that this is the cause of up to 10% of cases of 'essential' hypertension, but adenomas represent a minority of cases when diagnosed in this way.

Clinical features

The usual presentation is simply hypertension. Hypokalaemia (<3.5 mmol/L) was a hallmark of the condition as originally described but is now frequently absent. The few symptoms are non-specific; rarely muscle weakness, nocturia and tetany are seen. The hypertension may be severe and associated with renal, cardiac and retinal damage.

Adenomas, often very small, are more common in young females, while bilateral hyperplasia rarely occurs before age 40 years and is more common in males.

Investigations

Beta-blockers and other drugs may interfere with renin activity, and spironolactone, ACE inhibitors and AII-R antagonists will all affect results and all should be discontinued if possible. The characteristic features are as follows:

- **Plasma aldosterone:renin ratio (ARR)** is now most frequently used as a screening test for the condition, but raised ARR alone does not confirm the diagnosis (if the renin is low enough ARR will always be high). Note that normal ranges are highly assay dependent.
- **Elevated plasma aldosterone levels** that are not suppressed with 0.9% saline infusion (2 L over 4 hours) or fludrocortisone administration. Between 30% and 50% of patients with raised ARR on screening will suppress normally, excluding the diagnosis.
- **Suppressed plasma renin activity** or immunoreactivity.
- **Hypokalaemia** is often present but a normal serum potassium does not exclude the diagnosis.
- **Urinary potassium loss**. Levels >30 mmol daily during hypokalaemia are inappropriate.

Once a diagnosis of hyperaldosteronism is established, differentiation of adenoma from hyperplasia involves adrenal CT or MRI, but small adenomas may be missed and non-functioning incidentalomas also occur. Further information may be obtained from diurnal/postural changes in plasma aldosterone levels (which tend to rise with adenomas between 0900 h supine and 1300 h erect samples; in contrast, they fall with hyperplasia), measurement of 18-OH cortisol levels (raised in adenoma) and venous catheterization for aldosterone levels. All of these tests have their pitfalls and exceptions.

Glucocorticoid (or dexamethasone)-suppressible hyperaldosteronism is rare and is caused by a chimeric gene on chromosome 8. A fusion gene resulting from an unusual cross-over at meiosis between the genes encoding aldosterone synthase and adrenal 11β-hydroxylase produces aldosterone which is under ACTH control. Treatment with glucocorticoid resolves the problem.

Treatment

An adenoma can be removed surgically – usually laparoscopically; BP falls in 70% of patients. Those with hyperplasia should be treated with the aldosterone antagonist spironolactone (100–400 mg daily); frequent side-effects include nausea, rashes and gynaecomastia, and the pure aldosterone receptor antagonist eplerenone is a useful alternative (p. 657). Spironolactone metabolites have been linked with tumour development in animals. Amiloride and calcium-channel blockers are moderately effective in controlling the hypertension but do not correct the hyperaldosteronism.

Secondary hyperaldosteronism

This situation arises when there is excess renin (and hence angiotensin II) stimulation of the zona glomerulosa. Common causes are accelerated hypertension and renal artery stenosis, when the patient will be hypertensive. Causes associated with normotension include congestive cardiac failure and cirrhosis, where excess aldosterone production contributes to sodium retention.

Angiotensin-converting enzyme inhibitors (e.g. captopril, enalapril or lisinopril) and angiotensin II antagonists (e.g. losartan, candesartan) are effective in heart failure, both symptomatically and in increasing life expectancy (see p. 737). Spironolactone is of value in both situations, and 25 mg/day has been shown to improve survival in heart failure (see p. 739).

Syndrome of apparent mineralocorticoid excess

This causes the clinical syndrome of primary hyperaldosteronism but with low renin and aldosterone levels. Reduced activity of the enzyme 11β-hydroxysteroid dehydrogenase type 2 (11β-HSD2) prevents the normal conversion in the kidney of cortisol (which is active at the mineralocorticoid receptor) to cortisone (which is not) and therefore 'exposes' the mineralocorticoid receptor in the kidney to the usual molar excess of cortisol over aldosterone in the blood. While the inherited syndrome is rare, the same clinical syndrome can occur with excessive ingestion of liquorice, which inhibits the 11β-HSD2 enzyme.

Hypoaldosteronism

Except as part of primary hypoadrenalism (Addison's disease, see p. 1010), this is very uncommon. Causes include hyporeninaemic hypoaldosteronism, aldosterone biosynthetic defects, and drugs (e.g. ACE inhibitors, heparin).

THE ADRENAL MEDULLA

The major catecholamines, noradrenaline (norepinephrine) and adrenaline (epinephrine), are produced in the adrenal medulla (Fig. 18.28), although most noradrenaline is derived from sympathetic neuronal release. While noradrenaline and adrenaline undoubtedly produce hypertension when infused, they probably play little part in BP regulation in normal humans.

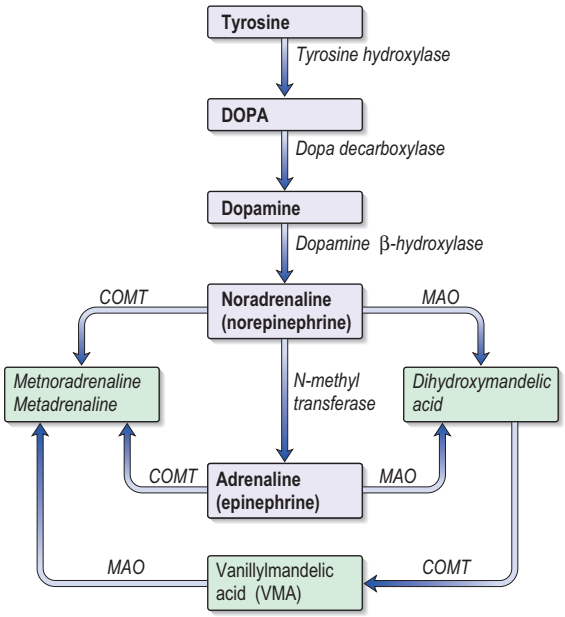


Fig. 18.28 **The synthesis and metabolism of catecholamines.** COMT, catechol-*O*-methyl transferase; MAO, monoamine oxidase.

Table 18.37 Symptoms and signs of phaeochromocytoma

Symptoms
Anxiety or panic attacks
Palpitations
Tremor
Sweating
Headache
Flushing
Nausea and/or vomiting
Weight loss
Constipation or diarrhoea
Raynaud's phenomenon
Chest pain
Polyuria/nocturia
Signs
Hypertension
Tachycardia/arrhythmias
Bradycardia
Orthostatic hypotension
Pallor or flushing
Glycosuria
Fever
(Signs of hypertensive damage)

Phaeochromocytoma

Phaeochromocytomas, tumours of the sympathetic nervous system, are very rare (less than 1 in 1000 cases of hypertension). Ninety per cent arise in the adrenal, while 10% occur elsewhere in the sympathetic chain. Some are associated with MEN 2 syndromes (see below) and the von Hippel–Lindau syndrome (p. 644). Most tumours release both noradrenaline (norepinephrine) and adrenaline (epinephrine) but large tumours and extra-adrenal tumours produce almost entirely norepinephrine.

Paragangliomas are related catecholamine-secreting tumours, typically in the head and neck but also found in the thorax, pelvis and bladder. They are more closely associated with other genetic associations of phaeochromocytoma. The association of paraganglioma, bilateral adrenal phaeochromocytomas, positive family history or young age at presentation is seen in multiple endocrine neoplasms (p. 1026).

Pathology

Oval groups of cells occur in clusters and stain for chromogranin A. Twenty-five per cent are multiple and 10% malignant, the latter being more frequent in the extra-adrenal tumours. Malignancy cannot be determined on simple histological examination alone.

Clinical features

The clinical features are those of catecholamine excess and are frequently, but not necessarily, intermittent (Table 18.37).

Diagnosis

Specific tests are:

- **Measurement of urinary catecholamines and metabolites** (metanephrines are most sensitive and specific – Fig. 18.28) is a useful screening test; normal levels on three 24-hour collections of metanephrines virtually exclude the diagnosis. Many drugs and dietary vanilla interfere with these tests.
- **Resting plasma catecholamines** are raised.
- **Plasma chromogranin A** (a storage vesicle protein) is raised.
- **Clonidine suppression test** may be appropriate, but should only be performed in specialist centres.
- **CT scans**, initially of the abdomen, are helpful to localize the tumours which are often large.
- **MRI** usually shows the lesion clearly.
- **Scanning with [^{131}I]metaiodobenzylguanidine (mIBG)** produces specific uptake in sites of sympathetic activity with about 90% success. It is particularly useful with extra-adrenal tumours.

Treatment

Tumours should be removed if this is possible; 5-year survival is about 95% for non-malignant tumours. Medical preoperative and perioperative treatment is vital and includes complete alpha- and beta-blockade with phenoxybenzamine (20–80 mg daily initially in divided doses), then propranolol (120–240 mg daily), plus transfusion of whole blood to re-expand the contracted plasma volume. The alpha-blockade must precede the beta-blockade, as worsened hypertension may otherwise result. Labetolol is not recommended. Surgery in the unprepared patient is fraught with dangers of both hypertension and hypotension; expert anaesthesia and an experienced surgeon are both vital, and sodium nitroprusside should be available in case sudden severe hypertension develops.

When operation is not possible, combined alpha- and beta-blockade can be used long term. Radionucleotide treatment with mIBG has been used but with limited success in malignant phaeochromocytoma.

Patients should be kept under clinical and biochemical review after tumour resection as over 10% recur or develop a further tumour. Catecholamine excretion measurements should be performed at least annually.

FURTHER READING

Lenders JW, Eisenhofer G, Mannelli M et al. Phaeochromocytoma. *Lancet* 2005; **366**: 665–675.

Taupenot L. The chromogranin-secretogranin family. *New England Journal of Medicine* 2003; **348**: 1134–1149.

OTHER ENDOCRINE DISORDERS

DISEASES OF MANY GLANDS

Multiple gland failure (polyglandular autoimmune syndromes)

These are caused by autoimmune disease as detailed in Table 18.2 on page 965. Most common are the associations of primary hypothyroidism and type 1 diabetes, and either of these with Addison's disease or pernicious anaemia.

Autoimmune polyendocrinopathy type 1 (APS-1) is an autosomal recessive disorder and is caused by *AIRE* gene mutations. The *AIRE* gene is present in the epithelium of the thymus and is involved in the presentation of self-antigens to thymocytes. Mutations will allow persistence of lymphocytes from the thymus to react against self antigens and cause development of autoimmune disorders. Mucocutaneous candidiasis often develops before the onset of endocrine deficiencies, such as Addison's disease, type 1 diabetes, hyperparathyroidism, nail dystrophy, vitiligo and dental enamel hypoplasia.

APS-2 is not associated with candidiasis and is also known as Schmidt's syndrome, typically when hypothyroidism, Addison's disease and type 1 diabetes are present in combination; coeliac disease is also an association.

Multiple endocrine neoplasia

This is the name given to the simultaneous or metachronous occurrence of tumours involving a number of endocrine glands (Table 18.38). The condition is inherited in an autosomal dominant manner and arise from the expression of recessive oncogenic mutations, most of which have been isolated. Affected persons may pass on the mutation to their offspring in the germ cell, but for the disease to become evident a somatic mutation must also occur, such as deletion or loss of a normal homologous chromosome. The defect in MEN 1 is in a novel gene (*menin*) on the long arm of chromosome 11 which encodes for a 610-amino-acid protein. *Menin* represses a transcription factor (*JunD*) and lack of *JunD* suppression leads to decreased apoptosis and oncogenesis. Patients with the *MEN1* gene carry one mutant gene and a wild type gene (i.e. are heterozygous). When the wild type gene undergoes a random somatic mutation during life, this leads to loss of heterozygosity and explains the late onset of tumours at any stage (the 'two hit' hypothesis). MEN 2a and 2b are caused by mutations of the *RET* proto-oncogene on chromosome 10 (see medullary thyroid cancer, p. 994). This gene encodes for a transmembrane glycoprotein receptor. For MEN 2a the mutation is in the extracellular domain; for 2b it is in the intracellular domain.

Table 18.38 Multiple endocrine neoplasia (MEN) syndromes

Organ	Frequency	Tumours/manifestations
Type 1		
Parathyroid	95%	Adenomas/hyperplasia
Pituitary	70%	Adenomas – prolactinoma, ACTH or growth hormone secreting (acromegaly)
Pancreas	50%	Islet cell tumours (secreting insulin, glucagon, somatostatin, VIP, pancreatic polypeptide, growth hormone-releasing factor) Zollinger–Ellison syndrome (gastrinoma) Non-functional tumour
Adrenal	40%	Non-functional adenoma
Thyroid	20%	Adenomas – multiple or single
Type 2a		
Adrenal	Most	Phaeochromocytoma (70% bilateral) Cushing's syndrome
Thyroid	Most	Medullary carcinoma (calcitonin producing)
Parathyroid	60%	Hyperplasia
Type 2b		
Type 2a with marfanoid phenotype and intestinal and visceral ganglioneuromas but not hyperparathyroidism Neuromas also present around lips and tongue		

Management

Treatment of established tumours in MEN is largely the same as treatment for similar tumours occurring sporadically. In MEN 1 four-gland parathyroidectomy is usually recommended when surgery is needed since all glands are typically involved. However, the essence of management in MEN is annual screening to detect tumours at an early, treatable stage.

Screening

A careful family history is essential. If the precise gene mutation has been identified in a particular family, then family members at risk can be offered genetic screening for the presence of the mutation, ideally in childhood. In affected individuals, biochemical screening and periodic imaging is then required.

Screening for type 1

Hyperparathyroidism is usually the first manifestation, and serum calcium is the simplest screening test in families with no identified mutation. In an established case (or gene-positive family member) other screening bloods include prolactin, GH/IGF-1 and 'gut hormones' (p. 272). Periodic imaging of pancreas, adrenals and pituitary is usually performed.

Screening for type 2

Serum calcium levels will easily detect hyperparathyroidism.

- *Medullary carcinoma of thyroid (MCT)* – with the known presence of the gene defect, total thyroidectomy is recommended in early childhood or as soon as the gene defect is identified. Calcitonin is a useful tumour marker.
- *Phaeochromocytoma* – metanephrine or catecholamine estimations.

McCune–Albright syndrome

This condition is associated with autonomous hypersecretion of a number of endocrine glands at a young age.

Gonadotrophin-independent puberty with Leydig cell hyperplasia in males and ovarian oestrogen production in girls occurs. Pituitary hypersecretion may lead to hyperprolactinaemia, acromegaly or gigantism. Cushing's syndrome due to nodular hyperplasia of the adrenal cortex is observed, as well as autonomous functioning thyroid nodules. Non-endocrine manifestations include café-au-lait patches and hypophosphataemic rickets. A point mutation of the *GNAS1* gene inhibits GTPase activity, leading to persistent activation of cAMP-mediated endocrine secretion.

ECTOPIC HORMONE SECRETION

This terminology refers to hormone synthesis, and normally secretion, from a neoplastic non-endocrine cell, most usually seen in tumours that have some degree of embryological resemblance to specialist endocrine cells. The clinical effects are those of the hormone produced, with or without manifestations of systemic malignancy. The most common situations seen are the following:

- *Hypercalcaemia of malignant disease*, often from squamous cell tumours of lung and breast, often with bone metastases. Where metastases are not present, most cases are mediated by secretion of PTH-related protein (PTHrP), which has considerable sequence homology to PTH; a variety of other factors may sometimes be involved, but very rarely PTH itself (see p. 1019). Treatment is also discussed on page 1020.
- *SIADH* (see p. 1018). Again, this is most common from a primary lung tumour.
- *Ectopic ACTH syndrome* (see p. 1012). Small-cell carcinoma of the lung, carcinoid tumours and medullary thyroid carcinomas are the most common causes, though many other tumours rarely cause it.
- *Production of insulin-like activity* may result in hypoglycaemia (see p. 1059).

ENDOCRINE TREATMENT OF OTHER MALIGNANCIES (see Ch. 9)

CHAPTER BIBLIOGRAPHY

Brook CGD, Hindmarsh PC. *Clinical Paediatric Endocrinology,* 4th edn. Oxford: Blackwell Science, 2001.

DeGroot L, Jameson JL. *Endocrinology.* Philadelphia: WB Saunders, 2001.

Wass JAH, Shalet SM (eds). *Oxford Textbook of Endocrinology and Diabetes.* Oxford: Oxford University Press, 2002.

SIGNIFICANT WEBSITES

http://www.endotext.org/
Online Endocrinology textbook

http://www.thyroidmanager.org/
Online Thyroid disease textbook

http://www.endocrinology.org
UK Society for Endocrinology

http://www.endo-society.org
Endocrine Society

http://www.niddk.nih.gov/health/endo/endo.htm
US National Institutes of Health, National Institute of Diabetes & Digestive & Kidney Diseases

http://www.endocrineweb.com
Endocrine web resources

http://www.addisons.org.uk/
The Addison's Self Help Group – information and guidelines on Addison's and steroid replacement

http://www.pituitary.org.uk
The Pituitary Foundation (UK charity) – comprehensive information for patients and GPs

http://www.tss.org.uk
UK Turner Syndrome Support Society

http://www.medicalert.org.uk
Emergency identification system for people with hidden medical conditions

This page intentionally left blank

19 Diabetes mellitus and other disorders of metabolism

DIABETES MELLITUS

Hyperglycaemia, insulin and insulin action 1029
Classification of diabetes 1031
Clinical presentation of diabetes 1036
Diagnosis and investigation of diabetes 1036
Treatment of diabetes 1037
Psychosocial implications of diabetes 1045
Diabetic metabolic emergencies 1046
Complications of diabetes 1049
Notes on special situations in diabetes 1058
Hypoglycaemia in the non-diabetic 1059
Disorders of lipid metabolism 1061
Hyperlipidaemia 1063
Classification, clinical features and investigation of primary hyperlipidaemias 1063
Therapies available to treat hyperlipidaemia 1065
Management of specific hyperlipidaemias 1065
Other lipid disorders 1068
Inborn errors of carbohydrate metabolism 1068
Inborn errors of amino acid metabolism 1068
Generalized aminoacidurias 1069
Specific aminoacidurias 1070
Lysosomal storage diseases 1071
Amyloidosis 1072
The porphyrias 1073

DIABETES MELLITUS

HYPERGLYCAEMIA, INSULIN AND INSULIN ACTION

Introduction

Diabetes mellitus (DM) is a syndrome of chronic hyperglycaemia due to relative insulin deficiency, resistance, or both. It affects more than 120 million people world-wide, and it is estimated that it will affect 370 million by the year 2030. Diabetes is usually irreversible and, although patients can have a reasonably normal lifestyle, its late complications result in reduced life expectancy and major health costs. These include macrovascular disease, leading to an increased prevalence of coronary artery disease, peripheral vascular disease and stroke, and microvascular damage causing diabetic retinopathy and nephropathy. Neuropathy is another major complication.

Insulin structure and secretion

Insulin is the key hormone involved in the storage and controlled release within the body of the chemical energy available from food. It is coded for on chromosome 11 and synthesized in the beta-cells of the pancreatic islets (Fig. 19.1). The synthesis, intracellular processing and secretion of insulin by the beta-cell is typical of the way that the body produces and manipulates many peptide hormones. Figure 19.2 illustrates the cellular events triggering the release of insulin-containing granules. After secretion, insulin enters the portal circulation and is carried to the liver, its prime target organ. About 50% of secreted insulin is extracted and degraded in the liver; the residue is broken down by the kidneys. C-peptide is only partially extracted by the liver (and hence provides a useful index of the rate of insulin secretion), but is mainly degraded by the kidneys.

An outline of glucose metabolism

Blood glucose levels are closely regulated in health and rarely stray outside the range of 3.5–8.0 mmol/L (63–144 mg/dL), despite the varying demands of food, fasting and exercise. The principal organ of glucose homeostasis is the liver, which absorbs and stores glucose (as glycogen) in the post-absorptive state and releases it into the circulation between meals to match the rate of glucose utilization by peripheral tissues. The liver also combines 3-carbon molecules derived from breakdown of fat (glycerol), muscle glycogen (lactate) and protein (e.g. alanine) into the 6-carbon glucose molecule by the process of gluconeogenesis.

Glucose production

About 200 g of glucose is produced and utilized each day. More than 90% is derived from liver glycogen and hepatic gluconeogenesis, and the remainder from renal gluconeogenesis.

Glucose utilization

The brain is the major consumer of glucose. Its requirement is 1 mg/kg bodyweight per minute, or 100 g daily in a 70 kg man. Glucose uptake by the brain is obligatory and is not dependent on insulin, and the glucose used is oxidized to carbon dioxide and water. Other tissues, such as muscle and fat, are facultative glucose consumers. The effect of insulin peaks associated with meals is to lower the threshold for glucose entry into cells; at other times, energy requirements are largely met by fatty-acid oxidation. Glucose taken up by

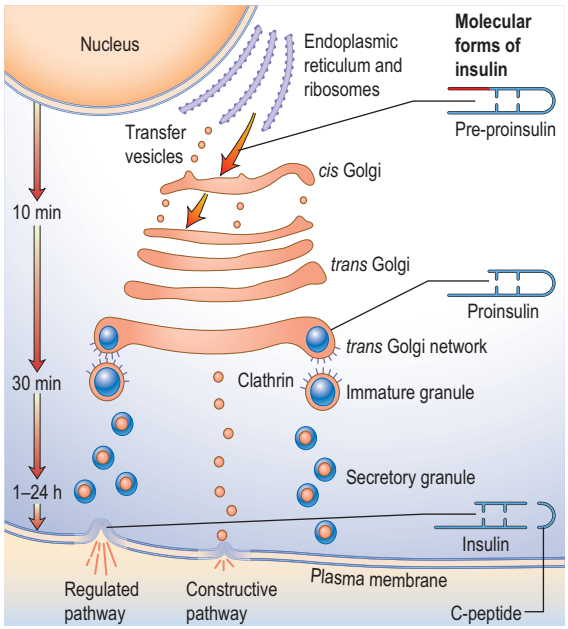


Fig. 19.1 Part of a beta-cell. The ribosomes manufacture pre-proinsulin from insulin mRNA. The hydrophobic 'pre' portion of pre-proinsulin allows it to transfer to the Golgi apparatus, and is subsequently enzymatically cleaved off. Proinsulin is parcelled into secretory granules in the Golgi apparatus. These mature and pass towards the cell membrane where they are stored before release. The proinsulin molecule folds back on itself and is stabilized by disulphide bonds. The biochemically inert peptide fragment known as connecting (C) peptide splits off from proinsulin in the secretory process, leaving insulin as a complex of two linked peptide chains. Equimolar quantities of insulin and C-peptide are released into the circulation. A small amount of insulin is secreted by the beta-cell directly via the 'constitutive pathway', which bypasses the secretory granules.

muscle is stored as glycogen or broken down to lactate, which re-enters the circulation and becomes a major substrate for hepatic gluconeogenesis. Glucose is used by fat tissue as a source of energy and as a substrate for triglyceride synthesis; lipolysis releases fatty acids from triglyceride together with glycerol, another substrate for hepatic gluconeogenesis.

Hormonal regulation

Insulin is a major regulator of intermediary metabolism, although its actions are modified in many respects by other hormones. Its actions in the fasting and postprandial states differ (Fig. 19.3). In the fasting state its main action is to regulate glucose release by the liver, and in the postprandial state it additionally facilitates glucose uptake by fat and muscle. The effect of counter-regulatory hormones (glucagon, epinephrine (adrenaline), cortisol and growth hormone) is to cause greater production of glucose from the liver and less utilization of glucose in fat and muscle for a given level of insulin.

Glucose transport

Cell membranes are not inherently permeable to glucose. A family of specialized glucose-transporter (GLUT) proteins carry glucose through the membrane into cells.

- GLUT-1 – enables basal non-insulin-stimulated glucose uptake into many cells (see Fig. 6.26).
- GLUT-2 – transports glucose into the beta-cell: a prerequisite for glucose sensing.
- GLUT-3 – enables non-insulin-mediated glucose uptake into brain neurones and placenta.
- GLUT-4 – enables much of the peripheral action of insulin. It is the channel through which glucose is taken up into muscle and adipose tissue cells following stimulation of the insulin receptor (Fig. 19.4).

The insulin receptor

This is a glycoprotein (400 kDa), coded for on the short arm of chromosome 19, which straddles the cell membrane of

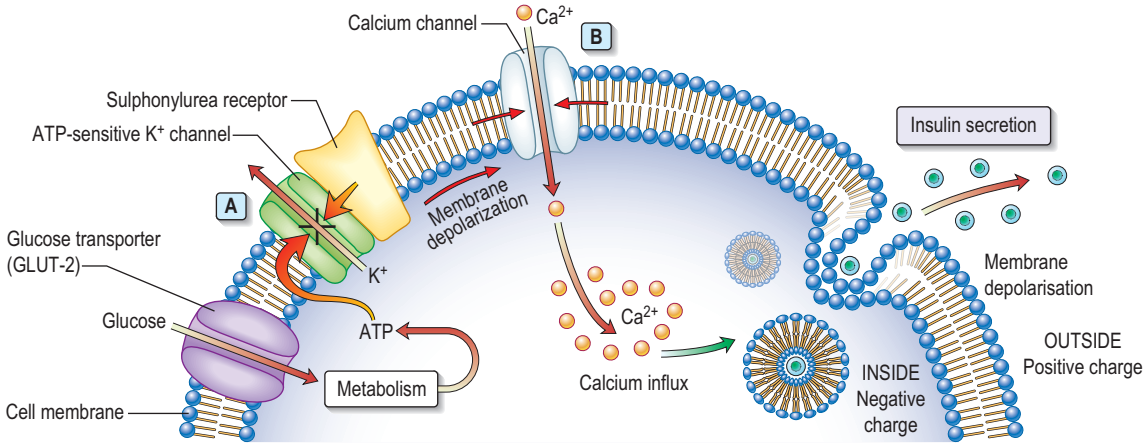


Fig. 19.2 Local forces regulating insulin secretion from beta-cells. Glucose enters the beta-cell via the GLUT-2 transporter protein, which is closely associated with the glycolytic enzyme glucokinase. Metabolism of glucose within the beta-cell generates ATP. ATP closes potassium channels in the cell membrane (A). If a sulphonylurea binds to its receptor, this also closes potassium channels. Closure of potassium channels predisposes to cell membrane depolarization, allowing calcium ions to enter the cell via calcium channels in the cell membrane (B). The rise in intracellular calcium triggers activation of calcium-dependent phospholipid protein kinase which, via intermediary phosphorylation steps, leads to fusion of the insulin-containing granules with the cell membrane and exocytosis of the insulin-rich granule contents. Similar mechanisms produce hormone-granule secretion in many other endocrine cells.

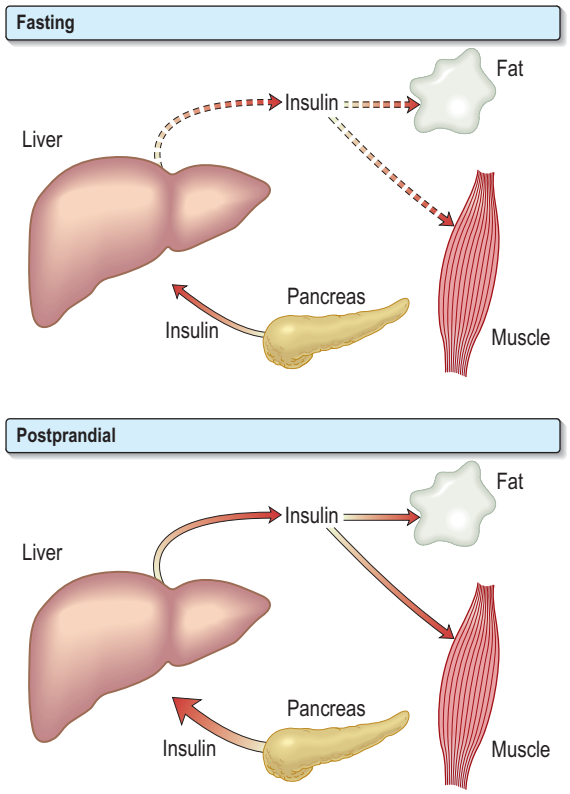


Fig. 19.3 **Fasting and postprandial effects of insulin.** In the fasting state insulin concentrations are low and it acts mainly as a hepatic hormone, modulating glucose production (via glycogenolysis and gluconeogenesis) from the liver. Hepatic glucose production rises as insulin levels fall. In the postprandial state insulin concentrations are high and it then suppresses glucose production from the liver and promotes the entry of glucose into peripheral tissues (increased glucose utilization).

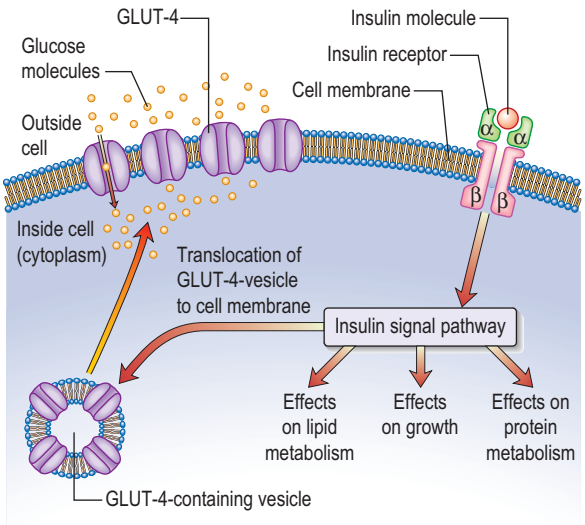


Fig. 19.4 **Insulin signalling in peripheral cells.** The insulin receptor consists of alpha- and beta-subunits linked by disulphide bridges (top right of figure). The beta-subunits straddle the cell membrane. The transporter protein GLUT-4 is stored in intracellular vesicles. The binding of insulin to its receptor initiates many intracellular actions including translocation of these vesicles to the cell membrane, carrying GLUT-4 with them.

many cells (Fig. 19.4). It consists of a dimer with two alpha-subunits, which include the binding sites for insulin, and two beta-subunits, which traverse the cell membrane. When insulin binds to the alpha-subunits it induces a conformational change in the beta-subunits, resulting in activation of tyrosine kinase and initiation of a cascade response involving a host of other intracellular substrates. One consequence of this is migration of the GLUT-4 glucose transporter to the cell surface and increased transport of glucose into the cell. The insulin–receptor complex is then internalized by the cell, insulin is degraded, and the receptor is recycled to the cell surface.

CLASSIFICATION OF DIABETES

Diabetes may be primary (idiopathic) or secondary (Table 19.1). Although secondary diabetes accounts for barely 1–2% of all new cases at presentation, it should not be missed because the cause can sometimes be treated. The classification of primary diabetes continues to evolve. Monogenic forms have been identified (see p. 1035), in some cases with significant therapeutic implications. The polygenic forms known as type 1 and type 2 diabetes are still distinguished to some extent by age of onset, or perceived need for insulin, but more formal classification is based around the balance between insulin deficiency and resistance, and the presence or absence of immunogenic markers associated with type 1 diabetes. Type 1 and type 2 diabetes represent two distinct diseases from the epidemiological point of view, but clinical distinction can sometimes be difficult. The two diseases should therefore be seen as a spectrum, distinct at the two ends but overlapping to some extent in the middle (Table 19.2). Hybrid forms are increasingly recognized, and patients with immune-mediated diabetes (type 1) may, for example, also be overweight and insulin resistant. It is more relevant to give the patient the right treatment on clinical grounds than to worry about how to label their diabetes. All forms of diabetes derive from inadequate insulin secretion relative to the needs of the body, and progressive insulin secretory failure is characteristic of both forms of diabetes. Thus, some patients with immune-mediated diabetes type 1 may not at first require insulin, whereas many with type 2 diabetes will eventually do so.

Table 19.1 Causes of secondary diabetes

Pancreatic disease
Cystic fibrosis
Chronic pancreatitis
Malnutrition-related pancreatic disease
Pancreatectomy
Hereditary haemochromatosis
Carcinoma of the pancreas

Endocrine disease
Cushing's syndrome
Acromegaly
Thyrotoxicosis
Phaeochromocytoma
Glucagonoma

Drug-induced disease
Thiazide diuretics
Corticosteroid therapy
Atypical antipsychotics
Antiretroviral protease inhibitors
Beta-blockers

Insulin-receptor abnormalities
Congenital lipodystrophy
Acanthosis nigricans

Genetic syndromes
Friedreich's ataxia
Dystrophia myotonica

Table 19.2 The spectrum of diabetes: a comparison of type 1 and type 2 diabetes mellitus

	Type 1	Type 2
Epidemiology	Younger (usually < 30 years of age) Usually lean Increased in those of Northern European ancestry Seasonal incidence	Older (usually > 30 years of age) Often overweight All racial groups. Increased in peoples of Asian, African, Polynesian and American-Indian ancestry
Heredity	HLA-DR3 or DR4 in > 90% 30–50% concordance in identical twins	No HLA links ~ 50% concordance in identical twins
Pathogenesis	Autoimmune disease: Islet cell autoantibodies Insulitis Association with other autoimmune diseases Immunosuppression after diagnosis delays beta-cell destruction	No immune disturbance Insulin resistance
Clinical	Insulin deficiency May develop ketoacidosis Always need insulin	Partial insulin deficiency initially May develop hyperosmolar state Many come to need insulin when beta-cells fail over time
Biochemical	Eventual disappearance of C-peptide	C-peptide persists

Type 1 diabetes mellitus

Epidemiology

Type 1 diabetes is a disease of insulin deficiency. In western countries almost all patients have the immune-mediated form of the disease (type 1A). Type 1 diabetes is prominent as a disease of childhood, reaching a peak incidence around the time of puberty, but can present at any age. A 'slow-burning' variant with slower progression to insulin deficiency occurs in later life and is sometimes called *latent autoimmune diabetes in adults (LADA)*. This may be difficult to distinguish from type 2 diabetes. Clinical clues are: hyperglycaemia which fails to correct with diet and tablet treatment, and autoantibody tests indicating autoimmune disease. The highest rates of type 1 diabetes in the world are seen in Finland and other Northern European countries, with the exception of the island of Sardinia, which for unknown reasons has the second highest rate in the world (Fig. 19.5). The incidence of type 1 diabetes appears to be increasing in most populations. In Europe the annual increase is of the order of 2–3%, and is most marked in children under the age of 5 years. WHO (1995) estimated that there are 19.4 million people with type 1 diabetes and that the number will rise to 57.2 million by 2025.

Causes

Type 1 diabetes belongs to a family of HLA-associated immune-mediated organ-specific diseases. Genetic susceptibility is polygenic, with the greatest contribution from the HLA region. Autoantibodies directed against pancreatic islet constituents appear in the circulation within the first few years of life, and often predate clinical onset by many years. Autoantibodies are also found in older patients with LADA and carry an increased risk of progression to insulin therapy.

Genetic susceptibility and inheritance

Increased susceptibility to type 1 diabetes is inherited, but the disease is not genetically predetermined. The identical twin of a patient with type 1 diabetes has a 30–50% chance of developing the disease, which implies that non-genetic factors must also be involved. The risk of developing diabetes by age 20, curiously, is greater with a diabetic father (3–7%) than with a diabetic mother (2–3%). If one child in a

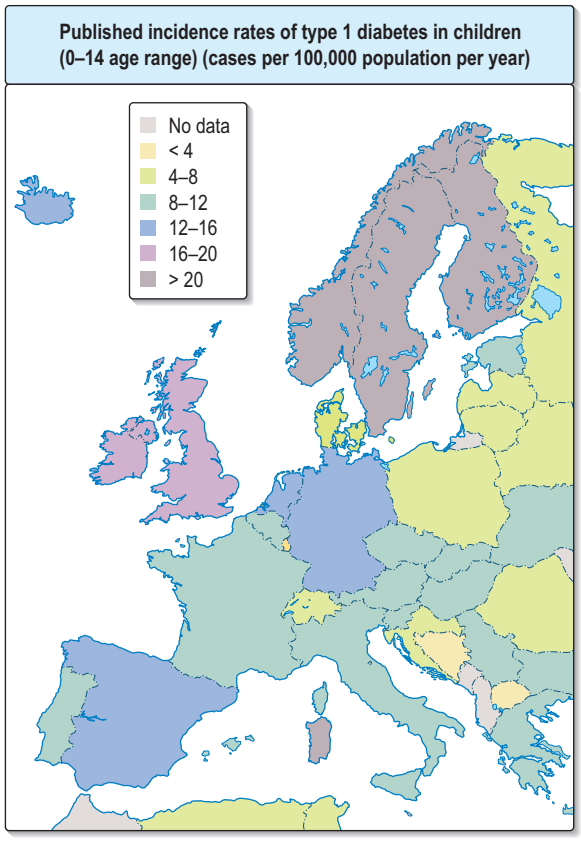


Fig. 19.5 **Age-standardized incidence rates of type 1 diabetes (onset 0–14 years) in Europe, per 100 000 per year.** Data courtesy of the International Diabetes Federation.

family has type 1 diabetes, each sibling has a ~ 6% risk of developing diabetes by age 20. This risk rises to about 20% in HLA-identical siblings who have the same HLA type as the proband. Since type 1 diabetes can present at any age, the lifetime risk for a sibling or child is at least double the risk by age 20.

HLA system

The HLA genes on chromosome 6 are highly polymorphic and modulate the immune defence system of the body. More than 90% of patients with type 1 diabetes carry HLA-DR3-DQ2, HLA-DR4-DQ8, or both, as compared with some 35% of the background population. All DQB1 alleles with an aspartic acid at residue 57 confer neutral to protective effects with the strongest effect from DQB1*0602 (DQ6), while DQB1 alleles with an alanine at the same position (i.e. DQ2 and DQ8) confer strong susceptibility. Genotypic combinations have a major influence upon risk of disease. For example, HLA DR3-DQ2/HLA DR4-DQ8 heterozygotes have a considerably increased risk of disease, and some HLA class I alleles also modify the risk conferred by class II susceptibility genes.

Other genes or gene regions

Genome-wide association studies have greatly broadened our understanding of the genetic background to type 1 diabetes and 11 genes or gene regions have been identified to date. These studies have confirmed that by far the greatest genetic contribution comes from the HLA region, but this is modulated by a large number of genes with small effects. These include the gene encoding insulin (*INS*) on chromosome 11 and a number of genes involved in immune responses, including the protein tyrosine phosphatase *PTPN22* which appears to be involved in a variety of HLA-associated autoimmune conditions.

Autoimmunity and type 1 diabetes

Type 1 diabetes overlaps with other organ-specific autoimmune diseases including autoimmune thyroid disease, coeliac disease, Addison's disease and pernicious anaemia. Autopsies of patients who died following diagnosis show infiltration of the pancreatic islets by mononuclear cells. This appearance, known as insulitis, resembles that in other autoimmune diseases such as thyroiditis. Autoantibodies directed against islet constituents precede the onset of clinical disease by many years and can be used to predict it; they are present in 95% of newly presenting patients. Several islet antigens have been characterized, and these include insulin itself, the enzyme glutamic acid decarboxylase (GAD), protein tyrosine phosphatase (IA-2) (Fig. 19.6) and the cation transporter ZnT8. Animal models of autoimmune diabetes such as the non-obese diabetic (NOD) mouse have been extensively studied. The observation that treatment with immunosuppressive agents such as ciclosporin prolongs beta-cell survival in newly diagnosed patients has confirmed that the disease is immune-mediated.

Environmental factors

The incidence of childhood type 1 diabetes is rising across Europe at the rate of 2–3% each year, suggesting that environmental factor(s) are involved in its pathogenesis. Islet autoantibodies (see above) appear in the first few years of life, indicating prenatal or early postnatal interactions with the environment. Exposures to dietary constituents, enteroviruses such as Coxsackie B4 or vaccinations have often been suspected, but their role in the causation of the disease has yet to be confirmed. It has also been suggested that a cleaner environment with less early stimulation of the immune system in childhood may increase susceptibility for type 1 diabetes, as for atopic/allergic conditions (the hygiene hypothesis) (see p. 847).

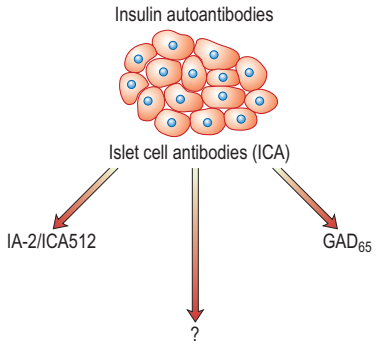


Fig. 19.6 **Islet autoantibodies.** Islet cell antibodies are detected by a fluorescent antibody technique which detects binding of autoantibodies to islet cells. Much of this staining reaction is due to antibodies specific for glutamic acid decarboxylase (GAD) and protein tyrosine phosphatase – IA-2 (also known as ICA512). Not all the staining seen with ICA is due to these two autoantibodies, so it is assumed that other islet autoantibodies are also involved. Insulin autoantibodies also appear in the circulation but do not contribute to the ICA reaction.

Pre-type 1 diabetes and prevention of type 1 diabetes

Prospective studies of first-degree relatives of children with type 1 diabetes show that islet autoantibodies appear early in life, and sometimes many years before diagnosis. Individuals who test positive for two or more autoantibodies have a >80% risk of progression to diabetes, and the risk approaches 100% in those who additionally lose their first phase insulin response to intravenous glucose. The ability to predict type 1 diabetes with this degree of precision has opened the way to disease prevention, but intervention trials using injected or oral insulin or other agents have so far proved unsuccessful.

Type 2 diabetes mellitus

Epidemiology

Type 2 diabetes is relatively common in all populations enjoying an affluent lifestyle. The four major determinants are increasing age, obesity, ethnicity and family history. Large differences in prevalence exist between populations based on these characteristics. In poor countries diabetes is a disease of the rich, but in rich countries it is a disease of the poor, obesity being the common factor. Diabetes may be present in a subclinical or undiagnosed form for years before diagnosis, and 25–50% of patients already have some evidence of vascular complications at the time of diagnosis. The onset may be accelerated by the stress of pregnancy, drug treatment or intercurrent illness. The overall prevalence within the UK is 2–3%, and the lifetime risk is around 15%. Type 2 diabetes is 2–4 times as prevalent in people of South Asian, African and Caribbean ancestry who live in the UK, and the

life-time risk in these groups exceeds 30%. High rates also affect people of middle Eastern and Latin or Hispanic American origin living western lifestyles. Obesity increases the risk of type 2 diabetes 80–100 fold, and this is reflected by the increasing prevalence of diabetes in different populations. The inhabitants of affluent countries gain on average nearly 1 gram in weight every day of their adult life between the ages of 25 and 55 years. This gain, due to a tiny excess in energy intake over expenditure – 90 kcal or one chocolate-coated digestive biscuit per day – is often due to reduced exercise rather than increased food intake. Further, our sedentary lifestyle means that the proportion of obese young adults is rising rapidly, and epidemic obesity will create a huge public health problem for the future. The increasing numbers of obese adolescents presenting with type 2 diabetes, particularly within high-risk ethnic groups, is a matter for concern.

Type 2 diabetes is associated with central obesity, hypertension, hypertriglyceridaemia, a decreased HDL-cholesterol, disturbed haemostatic variables and modest increases in a number of pro-inflammatory markers. Insulin resistance is strongly associated with many of these variables, as is increased cardiovascular risk. This group of conditions is referred to as the *metabolic syndrome* (see p. 231). The International Diabetes Federation has proposed criteria based on increased waist circumference (or BMI > 30) plus two of the following: diabetes (or fasting glucose > 6.0 mmol/L), hypertension, raised triglycerides or low HDL cholesterol. On this definition, about a third of the adult population have features of the syndrome, not necessarily associated with diabetes. Critics would argue that the metabolic syndrome is not a distinct entity, but one end of a continuum in the relationship between exercise, lifestyle and bodyweight on the one hand, and genetic make-up on the other, and that diagnosis adds little to standard clinical practice in terms of diagnosis, prognosis or therapy.

Causes of type 2 diabetes

Inheritance

Population-based studies show that identical twins of patients with type 2 diabetes have >50% chance of developing diabetes; the risk to non-identical twins or siblings is of the order of 25%. These observations confirm a strong genetic component to the disease. Type 2 diabetes is a polygenic disorder, and, as with type 1 diabetes, genome-wide association data have allowed identification of numerous susceptibility loci (16 at the time of writing). These loci, many of which are related to genes subserving beta-cell development or function, do not overlap with those identified for type 1 diabetes. Another difference is that there is no major gene susceptibility, equivalent to the HLA region, in type 2 diabetes. Thus the allelic odds ratio for TCLF7, the most common variant for type 2 diabetes, is a mere ~ 1.35, and other common variants have odds ratios which range between 1.1 and 1.2. Paradoxically, the new genes for type 2 diabetes account for a relatively small fraction of its observed heritability; nor do they allow subtypes of the condition to be identified with any confidence, as was previously anticipated.

Environmental factors: early and late

An association has been noted between low weight at birth and at 12 months of age and glucose intolerance later in life, particularly in those who gain excess weight as adults. The concept is that poor nutrition early in life impairs beta-cell development and function, predisposing to diabetes in later life. Low birthweight has also been shown to predispose to heart disease and hypertension.

Immunology and inflammation

There is no evidence of immune involvement in the pathogenesis of type 2 diabetes but, as noted earlier, a proportion of late-onset patients carry islet autoantibodies directed against GAD at diagnosis, and these are more likely to progress to insulin therapy. Such cases probably represent type 1 diabetes masquerading as type 2 diabetes.

Subclinical inflammatory changes are characteristic of both type 2 diabetes and obesity, and in diabetes, high-sensitivity C-reactive protein (CRP) levels are elevated in association with raised fibrinogen and increased plasminogen activator inhibitor-1 (PAI-1), and contribute to cardiovascular risk. Circulating levels of the proinflammatory cytokines TNF-α and IL-6 are elevated in both diabetes and obesity. Use of anti-inflammatory agents might potentially reduce the vascular risk associated with both conditions, but this has yet to be demonstrated.

Abnormalities of insulin secretion and action

Diabetes develops when the body can no longer secrete sufficient insulin to meet its requirements. The relative role of secretory failure versus insulin resistance in the pathogenesis of type 2 diabetes has been much debated, but even massively obese individuals with a fully functioning beta-cell mass do not necessarily develop diabetes, which implies that some degree of beta-cell dysfunction is necessary. Insulin can bind normally to its receptor on the surface of cells in type 2 diabetes, and the mechanisms of 'insulin resistance' are still poorly understood. Insulin resistance is however associated with central obesity and accumulation of intracellular triglyceride in muscle and liver in type 2 diabetes, and a high proportion of patients have non-alcoholic fatty liver disease (NAFLD), see page 344. It has long been maintained that patients with type 2 diabetes retain up to 50% of their beta-cell mass at the time of diagnosis, as compared with healthy controls, but the shortfall is greater than this when they are matched with healthy individuals who are equally obese. In addition, patients with type 2 diabetes almost all show islet amyloid deposition at autopsy, derived from a peptide known as amylin or islet amyloid polypeptide (IAPP) which is co-secreted with insulin. It is not known if this is a cause or consequence of beta-cell secretory failure.

Abnormalities of insulin secretion manifest early in the course of type 2 diabetes. One of the earliest signs is loss of the first phase of the normal biphasic response to intravenous insulin. The onset of diabetes is associated with hypersecretion of insulin by a depleted beta-cell mass. Circulating insulin levels are therefore higher than in healthy controls, although still inadequate to restore glucose homeostasis. Relative insulin lack is associated with increased glucose production from the liver (owing to inadequate suppression of gluconeogenesis) and reduced glucose uptake by peripheral tissues. Hyperglycaemia and lipid excess are toxic to beta-cells, at least in vitro, a phenomenon known as glucotoxicity, and this is thought to result in further beta-cell loss and further deterioration of glucose homeostasis. Circulating insulin levels thus peak with increasing duration of diabetes

and then decline due to secretory failure, an observation sometimes known as the 'Starling curve' of the pancreas. Type 2 diabetes is thus a condition in which relative insulin deficiency tends to progress towards absolute insulin deficiency requiring insulin therapy.

It has recently been shown that short-term intensive insulin therapy, given early in the course of type 2 diabetes, improves the recovery and maintenance of beta-cell function.

Overview and prevention

Genetic predisposition determines whether an individual is susceptible to type 2 diabetes; if and when diabetes develops largely depends upon lifestyle. A dramatic reduction in the incidence of new cases of adult-onset diabetes was documented in World War II when food was scarce, and clinical trials in individuals with impaired glucose tolerance have shown that diet, exercise or agents such as metformin have a marked effect in deferring the onset of type 2 diabetes. Established diabetes can be reversed, even if temporarily, by successful diet or bariatric surgery. Diabetes is therefore largely preventable, although the most effective measures would be directed at the whole population and implemented early in life. Prevention is well worth while, for diabetes diagnosed in a man between the ages of 40 and 59 reduces life expectancy by 5–10 years. By contrast, type 2 diabetes diagnosed after the age of 70 has little effect on life expectancy in men.

Monogenic diabetes mellitus

The genetic causes of some rare forms of diabetes are shown in Table 19.3. Considerable progress has been made in understanding these rare variants of diabetes. Genetic defects of beta-cell function (previously called 'maturity-onset diabetes of the young' (MODY)) are dominantly inherited, and seven variants have been described, each associated with different clinical phenotypes (Table 19.4). These should be considered in people presenting with early-onset diabetes in association with an affected parent and early-onset diabetes in ~ 50% of relatives. They can often be treated with a sulphonylurea.

Infants who develop diabetes before 6 months of age are likely to have a monogenic defect and not true type 1 diabetes. Transient neonatal diabetes mellitus (TNDM) occurs soon after birth, resolves at a median of 12 weeks, and 50% of cases ultimately relapse later in life. Most have an abnormality of imprinting of the *ZAC* and *HYMAI* genes on chromosome 6q. The commonest cause of permanent neonatal diabetes mellitus (PNDM) is mutations in the *KCNJ11* gene encoding the Kir6.2 subunit of the beta-cell potassium-ATP

Table 19.3 Rare genetic causes of type 2 diabetes

Disorder	Features
Insulin receptor mutations	Obesity, marked insulin resistance, hyperandrogenism in women, acanthosis nigricans (areas of hyperpigmented skin)
Maternally inherited diabetes and deafness (MIDD)	Mutation in mitochondrial DNA. Diabetes onset before age 40. Variable deafness, neuromuscular and cardiac problems, pigmented retinopathy
Wolfram syndrome (DIDMOAD – diabetes insipidus, diabetes mellitus, optic atrophy and deafness)	Recessively inherited. Mutation in the transmembrane gene, *WFS1*. Insulin-requiring diabetes and optic atrophy in the first decade. Diabetes insipidus and sensorineural deafness in the second decade progressing to multiple neurological problems. Few live beyond middle age
Severe obesity and diabetes	Alström, Bardet–Biedl and Prader–Willi syndromes. Retinitis pigmentosa, mental insufficiency and neurological disorders
Disorders of intracellular insulin signalling. All with severe insulin resistance	Leprechaunism, Rabson–Mendenhall syndrome, pseudoacromegaly, partial lipodystrophy: lamin A/C gene mutation
Genetic defects of beta-cell function	See Table 19.4

Table 19.4 Genetic defects of beta-cell function

	HNF-4a	Glucokinase	HNF-1a	IPF-1	HNF-1b
Chromosomal location	20q	7p	12q	13q	17q
Proportion of all cases	5%	15%	70%	< 1% (MODY)	2%
Onset	Teens to thirties	Present from birth	Teens/twenties	Teens to thirties	Teens/twenties
Progression	Progressive hyperglycaemia	Little deterioration with age	Progressive hyperglycaemia	Progression unclear	Progression unclear
Microvascular complications	Frequent	Rare	Frequent	Few data	Frequent
Other features	None	Reduced birthweight	Sensitivity to sulfonylureas	Pancreatic agenesis in homozygotes	Renal cysts, proteinuria, renal failure

The glucokinase gene is intimately involved in the glucose-sensing mechanism within the pancreatic beta-cell.
The hepatic nuclear factor (HNF) genes and the insulin promoter factor-1 (IPF-1) gene control nuclear transcription in the beta-cell where they regulate its development and function.
Abnormal nuclear transcription genes may cause pancreatic agenesis or more subtle progressive pancreatic damage.
A handful of families with autosomal dominant diabetes have been described with mutations in **NeuroD1** and recently the carboxyl ester lipase (CEL) gene.

Ask the authors
www.studentconsult.com

channel. Neurological features are seen in 20% of patients. Diabetes is due to defective insulin release rather than beta-cell destruction, and patients can be treated successfully with sulfonylureas, even after many years of insulin therapy.

FURTHER READING

Daneman D. Type 1 diabetes. *Lancet* 2006; **367**: 847–858.

EURODIAB ACE Study Group. Variation and trends in incidence of childhood diabetes in Europe. *Lancet* 2000; **355**: 873–876.

Gale EAM. The rise of childhood type 2 diabetes in the 20th century. *Diabetes* 2002; **51**: 3353–3361.

Haffner SM. Insulin resistance, inflammation and the prediabetic state. *American Journal of Cardiology* 2003; **92**(suppl): 18J–26J.

Retnakaran R, Zinman B. Type 1 diabetes, hyperglycaemia and the heart. *Lancet* 2008; **371**: 1790–1799.

Sturnvoll M, Goldstein BJ, van Haefken TW. Type 2 diabetes; pathogenesis and treatment. *Lancet* 2008; **371**: 2153–2156.

Taylor R. Causation of type 2 diabetes. *New England Journal of Medicine* 2004; **350**: 639–641.

CLINICAL PRESENTATION OF DIABETES

Presentation may be acute, subacute or asymptomatic.

Acute presentation

Young people often present with a 2- to 6-week history and report the classic triad of symptoms:

- polyuria – due to the osmotic diuresis that results when blood glucose levels exceed the renal threshold
- thirst – due to the resulting loss of fluid and electrolytes
- weight loss – due to fluid depletion and the accelerated breakdown of fat and muscle secondary to insulin deficiency.

Ketonuria is often present in young people and may progress to ketoacidosis if these early symptoms are not recognized and treated.

Subacute presentation

The clinical onset may be over several months or years, particularly in older patients. Thirst, polyuria and weight loss are typically present but patients may complain of such symptoms as lack of energy, visual blurring (owing to glucose-induced changes in refraction), or pruritus vulvae or balanitis that is due to *Candida* infection.

Complications as the presenting feature

These include:

- staphylococcal skin infections
- retinopathy noted during a visit to the optician
- a polyneuropathy causing tingling and numbness in the feet
- erectile dysfunction
- arterial disease, resulting in myocardial infarction or peripheral gangrene.

Asymptomatic diabetes

Glycosuria or a raised blood glucose may be detected on routine examination (e.g. for insurance purposes) in individuals who have no symptoms of ill-health. Glycosuria is not diagnostic of diabetes but indicates the need for further investigations. About 1% of the population have renal glycosuria. This is an inherited low renal threshold for glucose, transmitted either as a Mendelian dominant or recessive trait.

Physical examination at diagnosis

Evidence of weight loss and dehydration may be present, and the breath may smell of ketones. Older patients may present with established complications, and the presence of the characteristic retinopathy is diagnostic of diabetes. In occasional patients there will be physical signs of an illness causing secondary diabetes (Table 19.1). Patients with severe insulin resistance may have acanthosis nigricans, which is characterized by blackish pigmentation at the nape of the neck and in the axillae.

DIAGNOSIS AND INVESTIGATION OF DIABETES

Diabetes is easy to diagnose when overt symptoms are present, and a glucose tolerance test is hardly ever necessary for clinical purposes. The oral glucose tolerance test has,

Box 19.1 WHO diagnostic criteria – 1999

WHO criteria for the diagnosis of diabetes are:

- Fasting plasma glucose >7.0 mmol/L (126 mg/dL)
- Random plasma glucose >11.1 mmol/L (200 mg/dL)
- One abnormal laboratory value is diagnostic in symptomatic individuals; two values are needed in asymptomatic people.

The glucose tolerance test is only required for borderline cases and for diagnosis of gestational diabetes.

The glucose tolerance test – WHO criteria

	Normal	Impaired glucose tolerance	Diabetes mellitus
Fasting	<7.0 mmol/L	<7.0 mmol/L	>7.0 mmol/L
2 h after glucose	<7.8 mmol/L	7.8–11.0 mmol/L	≥11.1 mmol/L

- Adult: 75 g glucose in 300 mL water
- Child: 1.75 g glucose/kg body weight
- Only a fasting and a 120-min sample are needed
- Results are for venous plasma – whole blood values are lower.

Note: There is no such thing as mild diabetes. All patients who meet the criteria for diabetes are liable to disabling long-term complications.

however, allowed more detailed epidemiological characterization based on the existence of separate glucose thresholds for macrovascular and microvascular disease. These correspond with the levels for the diagnosis of impaired glucose tolerance (IGT) and diabetes as specified by the WHO criteria set out in Box 19.1. Epidemiological studies show that for every person with known diabetes, there is another undiagnosed in the population. A much larger proportion fall into the intermediate category of impaired glucose tolerance.

Impaired glucose tolerance (IGT)

This is not a clinical entity but a risk factor for future diabetes and cardiovascular disease. Obesity and lack of regular physical exercise make progression to frank diabetes more likely. The classification is complicated by the poor reproducibility of the key two-hour value in the oral glucose tolerance test. The group is heterogeneous; some patients are obese, some have liver disease, and others are on medication that impairs glucose tolerance. Individuals with IGT have the same risk of cardiovascular disease as in frank diabetes but do not develop the specific microvascular complications.

Impaired fasting glucose (IFG)

This diagnostic category (fasting plasma glucose between 6.1 and 6.9 mmol/L) was introduced by the American Diabetes Association (ADA), with the practical advantage that it avoids the need for a glucose tolerance test. This is not a clinical entity, but does predict future risk of frank diabetes and cardiovascular disease. This category only overlaps with IGT to a limited extent, and therefore the associated risks of cardiovascular disease and future diabetes are not directly comparable. A lower cut-off of 5.6 mmol/L (rather than 6.1 mmol/L) has been recommended by the ADA, which will of course greatly increase the number of those affected.

Other investigations

No further tests are needed to diagnose diabetes. Other routine investigations include screening the urine for protein, a full blood count, urea and electrolytes, liver biochemistry and random lipids. The latter test is useful to exclude an associated hyperlipidaemia and, if elevated, should be repeated fasting after diabetes has been brought under control. Diabetes may be secondary to other conditions (Table 19.1), may be precipitated by underlying illness and be associated with autoimmune disease or hyperlipidaemia. Hypertension is present in 50% of patients with type 2 diabetes and even more in African and Caribbean patients.

TREATMENT OF DIABETES

The role of patient education and community care

The care of diabetes is based on self-management by the patient, who is helped and advised by those with specialized knowledge. The quest for improved glycaemic control has made it clear that whatever the technical expertise applied, the outcome depends on willing cooperation by the patient. This in turn depends on an understanding of the risks of diabetes and the potential benefits of glycaemic control and other measures such as maintaining a lean weight, stopping smoking and taking care of the feet. If accurate information is not supplied, misinformation from friends and other patients will take its place. For this reason the best time to educate the patient is soon after diagnosis. Organized education programmes involve all healthcare workers, including nurse specialists, dieticians and podiatrists, and should include ongoing support and updates wherever possible.

Diet

The diet for people with diabetes is no different from that considered healthy for everyone. Table 19.5 lists recommendations on the ideal composition of this diet. To achieve this, food for people with diabetes should be:

Table 19.5 Recommended composition of the diet for people with diabetes, with comments on how this may be achieved

Component of diet	Comment
Protein	1 g per kg ideal bodyweight (approx.)
Total fat	<35% of energy intake. Limit: fat/oil in cooking, fried foods, processed meats (burgers, salami, sausages), high-fat snacks (crisps, cake, nuts, chocolate, biscuits, pastry). Encourage: lower-fat dairy products (skimmed milk, reduced-fat cheese, low-fat yoghurt), lean meat
Saturated and *trans*-unsaturated fat	<10% of total energy intake
n-6 polyunsaturated fat	<10% of total energy intake
n-3 polyunsaturated fat	No absolute quantity recommended. Eat fish, especially oily fish, once or twice weekly. Fish oil supplements not recommended
Cis-monounsaturated fat	10–20% of total energy intake (olive oil, avocado)
Total carbohydrate	40–60% of total energy intake Encourage: artificial (intense) sweeteners instead of sugar (sugar-free fizzy drinks, squashes and cordials). Limit: fruit juices, confectionery, cake, biscuits
Sucrose	Up to 10% of total energy intake, provided this is eaten in the context of a healthy diet (examples: fibre-rich breakfast cereals, baked beans)
Fibre	No absolute quantity recommended. Soluble fibre has beneficial effects on glycaemic and lipid metabolism. Insoluble fibre has no direct effects on glycaemic metabolism, but benefits satiety and gastrointestinal health
Vitamins and antioxidants	Best taken as fruit and vegetables (five portions per day) in a mixed diet. There is no evidence for the use of supplements
Alcohol	Not forbidden. Its energy content should be taken into account, as should its tendency to cause delayed hypoglycaemia in those treated with insulin
Salt	<6 g per day (lower in hypertension)

- low in sugar (though not sugar free)
- high in starchy carbohydrate (especially foods with a low glycaemic index) i.e. slower absorption
- high in fibre
- low in fat (especially saturated fat).

The overweight or obese should be encouraged to lose weight by a combination of changes in food intake and physical activity.

Carbohydrates

Slowly absorbed carbohydrates (i.e. those with a low glycaemic index) prevent rapid swings in circulating glucose. For example, the glucose peak seen in the blood after eating pasta is much flatter than that seen after eating the same amount of carbohydrate as white potato.

Prescribing a diet

Most people find it extremely difficult to modify their eating habits, and repeated advice and encouragement are needed if this is to be achieved. A diet history is taken, and the diet prescribed should involve the least possible interference with the person's lifestyle. Advice from dieticians is more likely to affect medium-term outcome than advice from doctors. People taking insulin or oral agents have traditionally been advised to eat roughly the same amount of food (particularly carbohydrate) at roughly the same time each day, so that treatment can be balanced against food intake and exercise. Knowledgeable and motivated patients with type 1 diabetes, who get feedback from regular blood glucose monitoring, can vary the amount of carbohydrate consumed, or meal times, by learning to adjust their exercise pattern and treatment. This is the basis of the DAFNE (Dose Adjustment For Normal Eating) regimen.

Exercise

Diet treatment is incomplete without exercise. Any increase in activity levels is to be encouraged, but participation in more formal exercise programmes is best. Where facilities for this exist, exercise should be prescribed for everyone with diabetes. Several trials have shown that regular exercise reduces the risk of progression to type 2 diabetes by 30–60%, and the lowest long-term morbidity and mortality is seen in those with established disease who have the highest levels of cardiorespiratory fitness. Both aerobic and resistance training improve insulin sensitivity and metabolic control in type 1 and type 2 diabetes, although reported effects on metabolic control are inconsistent. Patients on insulin or sulfonylureas should be warned that there is an increased risk of hypoglycaemia for up to 6–12 hours following heavy exertion.

Tablet treatment for type 2 diabetes

Diet and lifestyle changes are the key to successful treatment of type 2 diabetes, and no amount of medication will succeed where these have failed. The concept is that controlling diabetes is not just a matter of swallowing tablets, and these should in general never be prescribed until lifestyle changes have been implemented. Tablets will however be needed if satisfactory metabolic control (see 'Measuring control' below) is not established within 4–6 weeks. A consensus treatment pathway is shown in Figure 19.9 (p. 1041). The three main options are metformin, a sulfonylurea or a thiazolidinedione.

Biguanide (Metformin)

Metformin is the only biguanide currently in use, and remains the best validated primary treatment for type 2 diabetes. It activates the enzyme AMP-kinase, which is involved in GLUT4 metabolism and fatty acid oxidation, but its mechanism of action remains unclear at a cellular level. It reduces the rate of gluconeogenesis, and hence hepatic glucose output, and increases insulin sensitivity. It does not affect insulin secretion, does not induce hypoglycaemia and does not predispose to weight gain. It is thus particularly helpful in the overweight, although normal weight individuals also benefit, and may be given in combination with sulfonylureas or thiazolidinediones. Metformin was as effective as sulfonylurea or insulin in glucose control and reduction of microvascular risk in the United Kingdom Prospective Diabetic Study (UKPDS), but proved unexpectedly beneficial in reducing cardiovascular risk, an effect that could not be fully explained by its glucose-lowering actions. Metformin is currently the only oral agent to have demonstrated unequivocal cardiovascular protection within a randomized controlled trial. Adverse effects include anorexia, epigastric discomfort and diarrhoea, and these prohibit its use in 5–10% of patients. Diarrhoea should never be investigated in a diabetic patient without testing the effect of stopping metformin! Lactic acidosis has occurred in patients with severe hepatic or renal disease, and metformin is contraindicated when these are present. A Cochrane review showed little risk of lactic acidosis with standard clinical use, but most clinicians withdraw the drug when serum creatinine exceeds 150 μmol/L.

Sulfonylureas (Table 19.6)

These act upon the beta-cell to promote insulin secretion in response to glucose and other secretagogues. They are ineffective in patients without a functional beta-cell mass, and

Table 19.6 Properties of the most commonly used sulfonylureas

Drug	Features
Tolbutamide	Lower maximal efficacy than other sulfonylureas Short half-life – preferable in elderly Largely metabolized by liver – can use in renal impairment
Glibenclamide	Long biological half-life Severe hypoglycaemia Do not use in the elderly
Glipizide and Glimepiride	Active metabolites Renal excretion – avoid in renal impairment
Gliclazide	Intermediate biological half-life Largely metabolized by liver – can use in renal impairment More costly
Chlorpropamide	Very long biological half-life Renal excretion – avoid in renal impairment 1–2% develop inappropriate ADH-like syndrome Facial flush with alcohol Very inexpensive – major issue for developing countries Can produce fatal hypoglycaemia Not recommended in the elderly

they are usually avoided in pregnancy. Their action is to bind to the sulfonylurea receptor on the cell membrane, which closes ATP-sensitive potassium channels and blocks potassium efflux. The resulting depolarization promotes influx of calcium, a signal for insulin release (Fig. 19.2). Sulfonylureas are cheap and more effective than the other main agents in achieving short-term (1–3 years) glucose control, but their effect wears off as the beta-cell mass declines. There are theoretical concerns that they might hasten beta-cell apoptosis and promote weight gain, and are therefore best avoided in the overweight. They can also cause hypoglycaemia and although the episodes are generally mild, fatal hypoglycaemia may occur. Severe cases should always be admitted to hospital, monitored carefully, and treated with a continuous glucose infusion since rebound hypoglycaemia may occur after single injections of glucose. Sulfonylureas should be used with care in patients with liver disease. Patients with renal impairment should only be given those primarily excreted by the liver. Tolbutamide is the safest drug in the very elderly because of its short duration of action.

Thiazolidinediones

The thiazolidinediones (more conveniently known as the 'glitazones') reduce insulin resistance by interaction with peroxisome proliferator-activated receptor-gamma (PPAR-gamma), a nuclear receptor which regulates large numbers of genes including those involved in lipid metabolism and insulin action. The paradox that glucose metabolism should respond to a drug that binds to nuclear receptors mainly found in fat cells is still not fully understood. Fatty tissue is redistributed in patients taking a glitazone, with a reduction of central adiposity but an increase in peripheral fat. One suggestion as to their action is that they act indirectly via the glucose–fatty acid cycle, lowering free fatty acid levels and thus promoting glucose consumption by muscle. The glitazones lower circulating insulin relative to plasma glucose, but do not return glucose levels to normal. They can be used alone or in combination with other agents. The glitazones reduce hepatic glucose production, an effect that is synergistic with that of metformin, and also enhance peripheral glucose uptake. Like metformin, the glitazones potentiate the effect of endogenous or injected insulin. The glitazones have yet to demonstrate unique advantages in the treatment of diabetes, and their place in routine diabetes care remains uncertain. Long-term clinical trials show that the glucose-lowering effect of rosiglitazone in monotherapy persists longer than that of glibenclamide, but is comparable to that of metformin. *Unwanted effects* include weight gain of 5–6 kg, together with fluid retention and increased risk of heart failure, anaemia and osteoporosis. Rosiglitazone and pioglitazone should not be used in patients with heart failure. A meta-analysis of clinical trials with rosiglitazone has suggested that it might cause an increased cardiovascular risk; it should not be used in patients with ischaemic heart disease. Pioglitazone has not been shown to have this effect so far.

Meglitinides, e.g. *repaglinide* and *nateglinide*, are insulin secretagogues. Meglitinides are the non-sulfonylurea moiety of glibenclamide. As with the sulfonylureas, they act via closure of the K^+-ATP channel in the beta-cells (Fig. 19.2). They are short-acting agents that promote insulin secretion in response to meals. They are more expensive than standard sulfonylureas. Whether they have advantages over established short-acting sulfonylureas such as tolbutamide has not been addressed by clinical trials.

Injection therapies

Incretins

It has long been known that the insulin response to oral glucose is greater than the response to intravenous glucose. This is known as the incretin effect, and is due to release of two peptide hormones, glucose-dependent insulinotropic peptide (GIP) and glucagon-like peptide-1 (GLP-1) from the L cells in the intestine. The incretin effect is diminished in type 2 diabetes.

GLP-1 has a very short half-life, and *exenatide* and *liraglutide*, longer-acting analogues, are now available.

Exenatide, which must be given by twice-daily subcutaneous injection, promotes insulin release, inhibits glucagon release, reduces appetite and delays gastric emptying, thus blunting the postprandial rise in plasma glucose. Its main clinical disadvantage is the need for injection, and its advantage is that it improves glucose control whilst inducing useful weight reduction. *Side-effects* include nausea, and acute pancreatitis has been reported. At present it is used as an alternative to insulin, particularly in the overweight. A new long-acting preparation is now available and can be given once-weekly as additional therapy for type 2 patients on maximal oral medication.

The enzyme dipeptidyl peptidase 4 (DPP4) rapidly inactivates GLP-1 as this is released into the circulation, and inhibition of this enzyme thus potentiates the effect of endogenous GLP-1 secretion. DDP4 inhibition is an alternative approach to incretin-based therapy. The two agents currently available, *sitagliptin* and *vildagliptin*, are moderately effective in lowering blood glucose but do not induce weight loss. They are likely to be most effective in the early stages of type 2 diabetes when insulin secretion is relatively preserved. The main *side-effect* is nausea. Although incretin-based therapies are currently enjoying a vogue, their place in the management of type 2 diabetes has yet to be established.

Other therapies

- *Intestinal enzyme inhibitors* include *acarbose*, a sham sugar that competitively inhibits alpha-glucosidase enzymes situated in the brush border of the intestine, reducing absorption of dietary carbohydrate. Undigested starch may then enter the large intestine where it will be broken down by fermentation. Abdominal discomfort, flatulence and diarrhoea can result, and dosage needs careful adjustment to avoid these side-effects.
- *Orlistat* is a lipase inhibitor which reduces the absorption of fat from the diet. It benefits diabetes indirectly by promoting weight loss in patients under careful dietary supervision on a low fat diet. This is necessary to avoid unpleasant steatorrhoea.
- *Rimonobant* is a cannabinoid that promotes weight loss in people with type 2 diabetes. At least half of all patients using it experience a mood change whilst on the medication, and it should not be routinely used in those with a past history of depression. It can help those type 2 patients who gain a lot of weight when they go onto insulin, but its place in therapy is not yet clear.
- *Gastric banding and gastric bypass surgery* (see p. 233) have been used in those with marked obesity unresponsive to 6 months' intensive attempts at dieting and graded exercise. The risks of surgery are not insignificant, but need to be balanced against the risks

of patients staying as they are. About a third of patients become non-diabetic after gastric bypass.

Insulin treatment

Insulin is found in every creature with a backbone, and the key parts of the molecule show few species differences. Small differences in the amino acid sequence may alter the antigenicity of the molecule. The glucose and insulin profiles in normal subjects are shown in Figure 19.7.

Short-acting insulins

Insulins derived from beef or pig pancreas are still used but have largely been replaced in most developed countries by biosynthetic human insulin. This is produced by adding a DNA sequence coding for insulin or proinsulin into cultured yeast or bacterial cells. Short-acting insulins are used for pre-meal injection in multiple dose regimens, for continuous intravenous infusion in labour or during medical emergencies, and in patients using insulin pumps. Human insulin is absorbed slowly, reaching a peak 60–90 minutes after subcutaneous injection, and its action tends to persist too long after meals, predisposing to hypoglycaemia. Absorption is delayed because soluble insulin is in the form of stable hexamers (six insulin molecules around a zinc core) and needs to dissociate to monomers or dimers before it can enter the circulation. *Short-acting insulin analogues* have been engineered to dissociate more rapidly following injection without altering the biological effect. *Insulin analogues* (Fig. 19.8) such as *insulin lispro*, *insulin aspart* and *insulin glulisine* enter the circulation more rapidly than human soluble insulin, and also disappear more rapidly. Although widely used, the short-acting analogues have little effect upon overall glucose control in most patients, mainly because improved postprandial glucose is balanced by higher levels before the next meal. A Cochrane review has concluded that there is no evidence as to their benefit in type 2 diabetes.

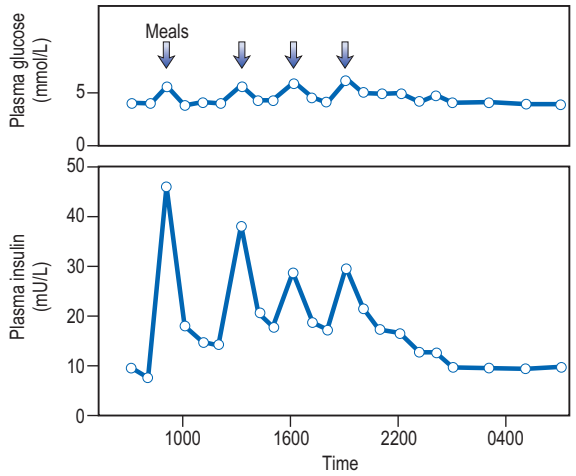


Fig. 19.7 **Glucose and insulin profiles in normal subjects.**

Longer-acting insulins

The action of human insulin can be prolonged by the addition of zinc or protamine derived from fish sperm. The most widely used form is NPH (isophane insulin), which has the advantage that it can be premixed with soluble insulin to form stable mixtures, of which the combination of 30% soluble with 70% NPH is most widely used. Long-acting analogues have their structure modified to delay absorption or to prolong their duration of action. *Insulin glargine* is soluble in the vial as a slightly acidic (pH 4) solution, but precipitates at subcutaneous pH, thus prolonging its duration of action. *Insulin detemir* has a fatty acid 'tail' which allows it to bind to serum albumin, and its slow dissociation from the bound state prolongs its duration of action. Although popular and widely used, these insulins have little advantage over NPH in most clinical situations, although useful in those on intensified therapy or with troublesome hypoglycaemia.

Inhaled insulin

The first inhaled insulin was withdrawn from the market in 2007, mainly because of limited clinical demand. An alterna-

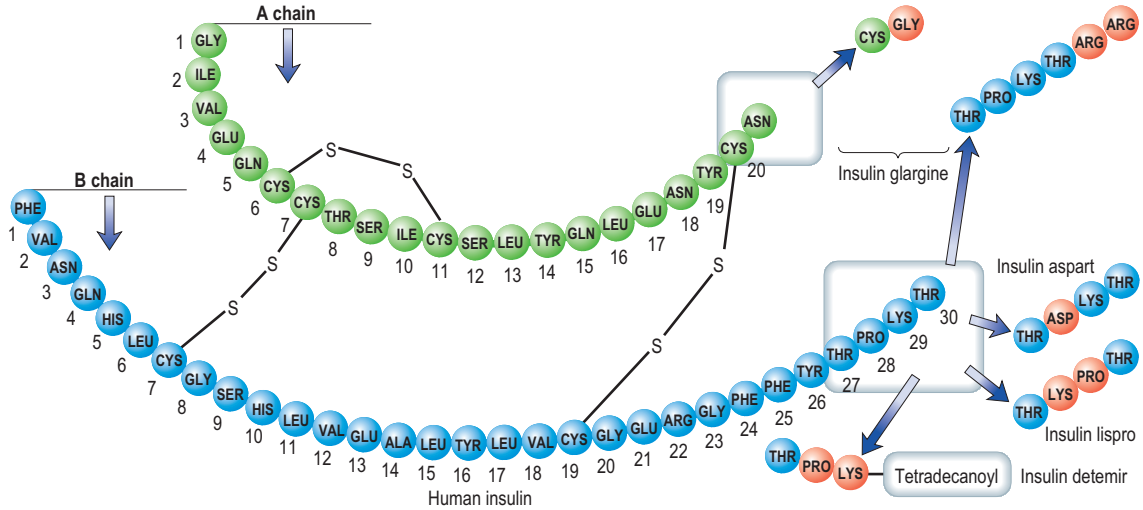


Fig. 19.8 **Amino acid structure of human insulin.** *Lispro* is a genetically engineered rapidly acting insulin analogue created by reversing the order of the amino acids proline and lysine in positions 28 and 29 of the B chain. *Insulin aspart* is a similar analogue created by replacing proline at position 28 of the B chain with an aspartic acid residue. *Insulin glargine* is a genetically engineered long-acting insulin created by replacing asparagine in position 21 of the A chain with a glycine residue and adding two arginines to the end of the B chain. *Insulin detemir* discards threonine in position 30 of the B chain and adds a fatty acyl chain to lysine in position B29.

tive may become available, but the main limitation is that only about 10% of the inhaled dose reaches the circulation, and this has cost implications. Variability of absorption and possible long-term adverse effects on the lungs are also of concern.

Practical management of diabetes

All patients with diabetes require advice about diet and lifestyle. Lifestyle changes, i.e. controlling weight, stopping smoking and taking regular exercise, can prevent or delay the onset of type 2 diabetes in people with glucose intolerance. Good glycaemic control is unlikely to be achieved with insulin or oral therapy when diet is neglected, especially when the patient is also overweight. Regular exercise helps to control weight and reduces cardiovascular risk. Blood pressure control is vital using an angiotensin converting enzyme (ACE) inhibitor or angiotension II receptor (AIIR) antagonist (see p. 804); Most patients will also benefit from a statin and low-dose aspirin (see p. 1050).

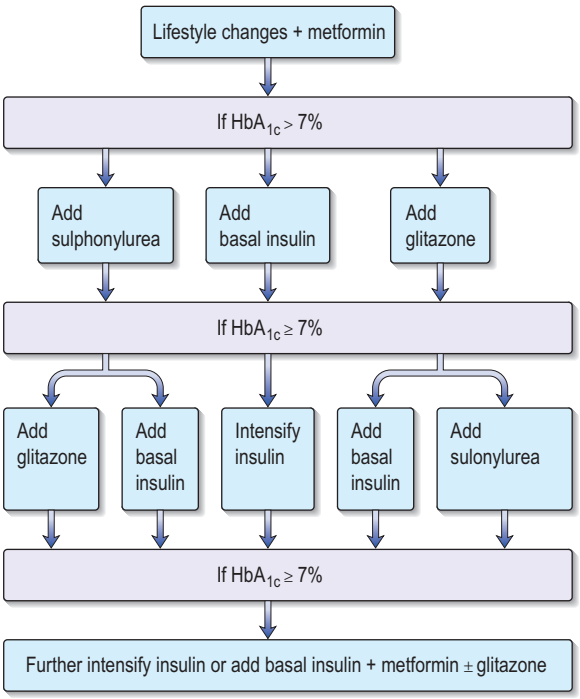
Type 2 diabetes

The great majority of patients presenting over the age of 40 will have type 2 diabetes, but do not miss the occasional type 1 patient presenting late. An approach to their management is illustrated in Figure 19.9. Recently it has been suggested that a GLP-1 agonist can be added to oral therapy. Goals of treatment are described on page 1044. Type 2 diabetes is characterized by progressive beta-cell failure, and glucose control deteriorates over time, requiring a progressive and pre-emptive escalation of diabetes therapy. Regular review is essential for this to be achieved. Most patients on tablets will eventually require insulin, and it is helpful to explain this from the outset. The most widespread error in management at this stage is procrastination; the patient whose control is inadequate on oral therapy should start insulin without undue delay. Insulin should also be considered in all type 2 patients with an HbA1c >8%, and in patients with high cardiovascular risk factors when HbA1c >7%.

There is little consensus regarding the optimal insulin regimen in type 2 diabetes, but an intermediate insulin given at night with metformin during the day is initially as effective as multidose insulin regimens in controlling glucose levels, and is less likely to promote weight gain. Metformin is a useful adjunct to insulin in those able to tolerate it. A second morning dose of insulin may become necessary to control postprandial hyperglycaemia. Twice-daily injections of premixed soluble and isophane insulins (i.e. biphasic isophane insulin) are widely used and reasonably effective (Fig. 19.10a). Once-daily insulin glargine showed good control of blood glucose in one study. More aggressive treatment, with multiple injections or continuous infusion pumps, is increasingly being used in younger patients with type 2 diabetes.

Type 1 diabetes

Insulin is always indicated in a patient who has been in ketoacidosis, and is usually indicated in lean patients who present under the age of 40 years.



Weight loss drugs not included

Fig. 19.9 **A treatment pathway for type 2 diabetes mellitus.** Note that discussion of lifestyle changes and compliance should be undertaken at every stage. All patients require BP control, statin therapy and low-dose aspirin. From American Diabetic Association/European Association for the Study of Diabetes guidelines on management of hyperglycemia in type 2 diabetes. *Diabetes Care* 2006; **29**: 1963–1972. Data courtesy of the International Diabetes Federation.

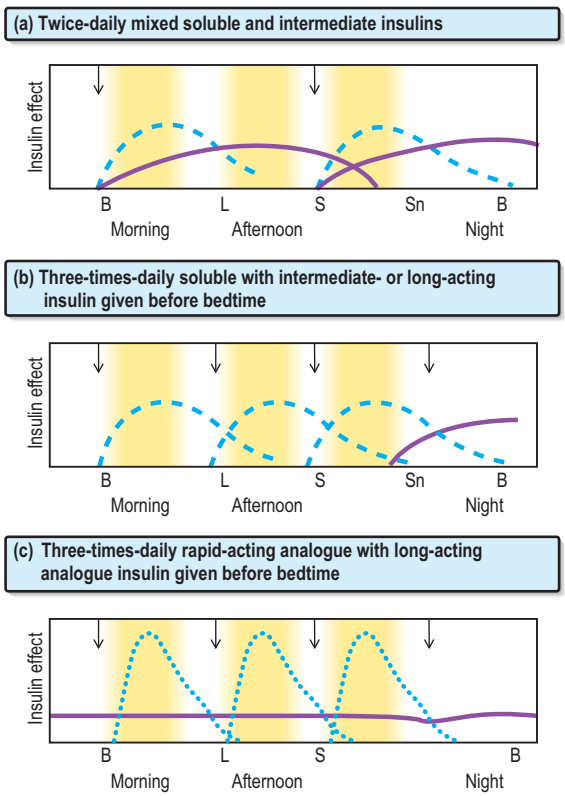


Fig. 19.10 **Insulin regimens.** Profiles of soluble insulins are shown as: dashed lines; intermediate- or long-acting insulin as solid lines (purple); and rapid-acting insulin as dotted lines (blue). The arrows indicate when the injections are given. B, breakfast; L, lunch; S, supper; Sn, snack (bedtime).

FURTHER READING

Nathan DS, Buse JB, Davidson MB et al. Medical management of hyperglycemia in type 2 diabetes; a consensus algorithm. *Diabetes Care* 2008; **31**: 1–11.

Principles of insulin treatment

Injections

The needles used to inject insulin are very fine and sharp. Even though most injections are virtually painless, patients are understandably apprehensive and treatment begins with a lesson in injection technique. Insulin is usually administered by a pen injection device but can be drawn up from a vial into special plastic insulin syringes marked in units (100 U in 1 mL). Injections are given into the fat below the skin on the abdomen, thighs or upper arm, and the needle is usually inserted to its full length. Slim adults and children usually use a 31 gauge 6 mm needle and fatter adults a 30 gauge 8 mm needle. Both reusable and disposable pen devices are available, together with a range of devices to aid injection. The injection site used should be changed regularly to prevent areas of lipohypertrophy (fatty lumps). The rate of insulin absorption depends on local subcutaneous blood flow, and is accelerated by exercise, local massage or a warm environment. Absorption is more rapid from the abdomen than from the arm, and is slowest from the thigh. All these factors can influence the shape of the insulin profile.

Insulin administration

In healthy individuals a sharp increase in insulin occurs after meals; this is superimposed on a constant background of secretion (Fig. 19.7). Insulin therapy attempts to reproduce this pattern, but ideal control is difficult to achieve for four reasons:

- In normal subjects, insulin is secreted directly into the portal circulation and reaches the liver in high concentration; about 50% of the insulin produced by the pancreas is cleared by the liver. By contrast, insulin injected subcutaneously passes into the systemic circulation before passage to the liver. Insulin-treated patients therefore have lower portal levels of insulin and higher systemic levels relative to the physiological situation.
- Subcutaneous soluble insulin takes 60–90 minutes to achieve peak plasma levels, so the onset and offset of action are too slow.
- The absorption of subcutaneous insulin into the circulation is variable.
- Basal insulin levels are constant in normal people, but injected insulin invariably peaks and declines in people with diabetes, with resulting swings in metabolic control.

A multiple injection regimen with short-acting insulin and a longer-acting insulin at night is appropriate for most younger patients (Fig. 19.10b). The advantages of multiple injection regimens are that the insulin and the food go in at roughly the same time so that meal times and sizes can vary, without greatly disturbing metabolic control. The flexibility of multiple injection regimens is of great value to patients with busy jobs, shift workers and those who travel regularly. Some recovery of endogenous insulin secretion may occur over the first few months (the 'honeymoon period') in type 1 patients and the insulin dose may need to be reduced or even stopped for a period. Requirements rise thereafter. Strict glucose control from diagnosis in type 1 diabetes prolongs beta-cell function, resulting in better glucose levels and less hypoglycaemia. Some type 1 diabetes patients will opt for twice-daily mixed insulin injections (Fig. 19.10a) and put up with the lifestyle restrictions that this imposes (with twice-daily regimens, the size and timing of meals are fixed more rigidly). Target blood glucose values should normally be 4–7 mmol/L before meals, 4–10 mmol/L after meals.

All patients need careful training for a life with insulin. This is best achieved outside hospital, provided that adequate facilities exist for outpatient diabetes education. A scheme for adjusting insulin regimens is given in Table 19.7. DAFNE is described on p. 1038.

When to use insulin analogues

Hypoglycaemia between meals and particularly at night is the limiting factor for many patients on multiple injection regimens. The more expensive rapid-acting insulin analogues (Fig. 19.10c) are a useful substitute for soluble insulin in some patients. They reduce the frequency of nocturnal hypoglycaemia due to reduced carry-over effect from the day-time. They are often used on grounds of convenience, since patients can inject shortly before meals but this is illogical since standard insulins injected at the same time give an equivalent overall control. High or erratic morning blood sugar readings can prove a problem for about a quarter of all patients on conventional multiple injection regimens, because the bedtime intermediate-acting insulin falls and the absorption is variable. The long-acting insulin analogues insulin glargine and insulin detemir help to overcome these problems and reduce the risk of nocturnal hypoglycaemia

Infusion devices

CSII (continuous subcutaneous insulin infusion) is delivered by a small pump strapped around the waist that infuses a constant trickle of insulin via a needle in the subcutaneous tissues. Meal-time doses are delivered when the patient touches a button on the side of the pump.

This approach is particularly useful in the overnight period, since the basal overnight infusion rate can be programmed to fit each patient's needs. Disadvantages include the nuisance of being attached to a gadget, skin infections, the risk of ketoacidosis if the flow of insulin is broken (since these patients have no protective reservoir of depot insulin) and cost. Infusion pumps should only be used by specialized centres able to offer a round-the-clock service to their patients. This form of treatment has revolutionized the lives of some patients.

Table 19.7 Guide to adjusting insulin dosage according to blood glucose test results

	Blood glucose persistently too high	Blood glucose persistently too low
Before breakfast	Increase evening long-acting insulin	Reduce evening long-acting insulin
Before lunch	Increase morning short-acting insulin	Reduce morning short-acting insulin or increase mid-morning snack
Before evening meal	Increase morning long-acting insulin or lunch short-acting insulin	Reduce morning long-acting insulin or lunch short-acting insulin or increase mid-afternoon snack
Before bed	Increase evening short-acting insulin	Reduce evening short-acting insulin

Complications of insulin therapy

At the injection site

Shallow injections result in intradermal insulin delivery and painful, reddened lesions or even scarring. Injection site abscesses occur but are extremely rare.

Local allergic responses sometimes occur early in therapy but usually resolve spontaneously. Generalized allergic responses are exceptionally rare. Fatty lumps, known as lipohypertrophy, may occur as the result of overuse of a single injection site with any type of insulin.

Insulin resistance

The most common cause of mild insulin resistance is obesity. Occasional unstable patients require massive insulin doses, often with a fluctuating requirement. There are often associated behavioural problems. Insulin resistance associated with antibodies directed against the insulin receptor has been reported in patients with acanthosis nigricans (Table 19.3).

Weight gain

Many patients show weight gain on insulin treatment, especially if the insulin dose is increased inappropriately – insulin makes you feel hungry! Sticking to a careful dietary regimen is vital when on insulin. Patients who are in poor control when insulin is started tend to gain most weight.

Hypoglycaemia during insulin treatment

This is the most common complication of insulin therapy, and limits what can be achieved with insulin treatment. It is a major cause of anxiety for patients and relatives. It results from an imbalance between injected insulin and a patient's normal diet, activity and basal insulin requirement. The times of greatest risk are before meals and during the night. Irregular eating habits, unusual exertion and alcohol excess may precipitate episodes; other cases appear to be due simply to variation in insulin absorption.

Symptoms develop when the blood glucose level falls below 3 mmol/L and typically develop over a few minutes, with most patients experiencing 'adrenergic' features of sweating, tremor and a pounding heartbeat. Virtually all patients experience intermittent symptoms and one in three will go into a coma at some stage in their lives. A small minority suffer attacks that are so frequent and severe as to be virtually disabling.

Physical signs include pallor and a cold sweat. Many patients with longstanding diabetes report loss of these warning symptoms and are at a greater risk of progressing to more severe hypoglycaemia. Such patients appear pale, drowsy or detached, signs that their relatives quickly learn to recognize. Behaviour is clumsy or inappropriate, and some become irritable or even aggressive. Others slip rapidly into hypoglycaemic coma. Occasionally, patients develop convulsions during hypoglycaemic coma, especially at night. This must not be confused with idiopathic epilepsy. Another presentation is with a hemiparesis that resolves when glucose is administered. Hypoglycaemia is a common problem.

Hypoglycaemic unawareness. People with diabetes have an impaired ability to counter-regulate glucose levels after hypoglycaemia. The glucagon response is invariably deficient, even though the alpha-cells are preserved and respond normally to other stimuli. The epinephrine (adrenaline) response may also fail in patients with a long duration of diabetes, and this is associated with loss of warning symptoms. Recurrent hypoglycaemia may itself induce a state of hypoglycaemia unawareness, and the ability to recognize the condition may sometimes be restored by relaxing control for a few weeks.

Nocturnal hypoglycaemia. Basal insulin requirements fall during the night but increase again from about 4 a.m. onwards, at a time when levels of injected insulin are falling. As a result many patients wake with high blood glucose levels, but find that injecting more insulin at night increases the risk of hypoglycaemia in the early hours of the morning. The problem may be helped by the following:

- checking that a bedtime snack is taken regularly
- for patients taking twice-daily mixed insulin to separate their evening dose and take the intermediate insulin at bedtime rather than before supper
- reducing the dose of soluble insulin before supper, since the effects of this persist well into the night
- changing patients on a multiple injection regimen with soluble insulin to a rapid-acting insulin analogue
- changing to a long-lasting insulin analogue at night.

Mild hypoglycaemia

Any form of rapidly absorbed carbohydrate will relieve the early symptoms, and sufferers should always carry glucose or sweets. Drowsy individuals will be able to take carbohydrate in liquid form (e.g. Lucozade). All patients and their close relatives need training about the risks of hypoglycaemia. Patients should not take more carbohydrate than necessary during the recovery period, since this causes a rebound to hyperglycaemia. Alcohol excess increases the risk of hypoglycaemia, and this requires careful explanation, together with the need to guard against hypoglycaemia while driving.

Severe hypoglycaemia

The diagnosis of severe hypoglycaemia resulting in confusion or coma is simple and can usually be made on clinical grounds, backed by a bedside blood test. If real doubt exists, blood should be taken for glucose estimation before treatment is given. Patients should carry a card or wear a bracelet or necklace identifying themselves as diabetic, and these should be looked for in unconscious patients.

Unconscious patients should be given either intramuscular glucagon (1 mg) or intravenous glucose (25–50 mL of 50% dextrose solution) followed by a flush of 0.9% saline to preserve the vein (since 50% dextrose scleroses veins). Glucagon acts by mobilizing hepatic glycogen, and works almost as rapidly as glucose. It is simple to administer and can be given at home by relatives. It does not work after a prolonged fast. Oral glucose is given to replenish glycogen reserves once the patient revives.

Whole pancreas and pancreatic islet transplantation

Whole pancreas transplantation has been performed for some 30 years, usually in diabetic patients who require immunosuppression for a kidney transplant. Surgical

advances have greatly improved the outcome of this procedure. In experienced hands graft function lasts longer with considerable improvement in quality of life. The procedure, however, adds to the risks of renal transplantation. Whole pancreas transplants survive better if grafted together with a kidney than when the pancreas alone is transplanted. There is evidence of protection against or reversal of some complications of diabetes.

Islet transplantation is performed by harvesting pancreatic islets from cadavers (two or three pancreata are usually needed); these are then injected into the portal vein and seed themselves into the liver. This form of treatment was attempted for many years with poor results, but better results have been achieved by an improved protocol developed in Edmonton, Canada. The main indication is disabling hypoglycaemia, and the main disadvantage is the need for powerful immunosuppressive therapy, with associated costs and complications.

Measuring the metabolic control of diabetes

Urine tests

Some patients will not perform capillary blood glucose monitoring at home. Urine tests using dipsticks give such patients some feedback on whether their diet and treatment are achieving reasonable metabolic control. If they have consistently negative tests with no symptoms of hypoglycaemia, one can assume that they are fairly well controlled. Nevertheless, the correlation between urine tests and simultaneous blood glucose is poor for three reasons:

- Changes in urine glucose lag behind changes in blood glucose.
- The mean renal threshold is around 10 mmol/L but the range is wide (7–13 mmol/L). The threshold also rises with age.
- Urine tests can give no guidance concerning blood glucose levels below the renal threshold.

Home blood glucose testing

The home provides the best place for assessment of day-to-day control and this is an essential aid to good glycaemic control. The fasting blood glucose concentration is a useful guide to therapy in tablet-treated type 2 diabetes, but random blood glucose testing (e.g. in the clinic) is of limited value. Patients can be easily taught to provide their own profiles by testing finger-prick blood samples with reagent strips and reading these with the aid of a meter. Blood is taken from the side of a finger tip (not from the tip, which is densely innervated) using a special lancet usually fitted to a spring-loaded device. Patients can talk through the use of various devices with a diabetes nurse specialist to select the most appropriate devices for them. Patients should take regular profiles (e.g. four daily samples on two days each week) and note these in a diary or record book. Patients on insulin are encouraged to adjust their insulin dose as appropriate (Table 19.7) and should ideally be able to obtain advice over the telephone when needed.

Glycosylated haemoglobin (HbA_1 or HbA_{1c}) and fructosamine

Glycosylation of haemoglobin occurs as a two-step reaction, resulting in the formation of a covalent bond between the glucose molecule and the terminal valine of the β chain of the haemoglobin molecule. The rate at which this reaction occurs is related to the prevailing glucose concentration. Glycosylated haemoglobin is expressed as a percentage of the normal haemoglobin (standardized range 4–6.2%). This test provides an index of the average blood glucose concentration over the life of the haemoglobin molecule (approximately 6 weeks). The figure will be misleading if the life-span of the red cell is reduced or if an abnormal haemoglobin or thalassaemia is present. There is considerable interindividual variation in HbA_{1c} levels, even in health. Although the glycosylated haemoglobin test provides a rapid assessment of the level of glycaemic control in a given patient, blood glucose testing is needed before the clinician can know what to do about it.

Glycosylated plasma proteins ('fructosamine') may also be measured as an index of control. Glycosylated albumin is the major component, and fructosamine measurement relates to glycaemic control over the preceding 2–3 weeks. It is useful in patients with anaemia or haemoglobinopathy and in pregnancy (when haemoglobin turnover is changeable) and other situations where changes of treatment need a swift means of assessing progress.

Targets

Data from the UK Prospective Diabetic Study (UKPDS) and the Diabetes Control and Complications Trial (DCCT) suggest that under ideal circumstances patients with both type 1 diabetes and type 2 diabetes should aim to run their glycosylated haemoglobin readings below 7.5% in order to reduce the risk of long-term microvascular complications (Table 19.8). Hypoglycaemia, patterns of eating and lifestyle, weight problems and problems accepting and coping with diabetes limit what can be achieved (see p. 1046). Some will, but most will not, be able to reach these target values, particularly as their duration of diabetes increases. Realistic goals should be set for each patient, taking into account what is likely to be achievable.

Statin therapy reduces cardiovascular events in both primary and secondary prevention studies. More intensive statin therapy to lower LDL-C to <1.8 mmol/L is advocated by some. Ezetimibe can be added to statin therapy.

Does good metabolic control of diabetes matter?

Blood glucose is just one measure of the diverse metabolic consequence of diabetes which not only affects carbohydrate metabolism but also the metabolism of lipids and proteins.

Table 19.8 Target goals of risk factors for diabetic patients

Parameter	Ideal	Reasonable but not ideal
HbA_{1c}	<7%	<8%
Blood pressure (mmHg)	<130/80	<140/80
Total cholesterol (mmol/L)	<4.0	<5.0
LDL	<2.0	<3.0
HDL*	>1.1	>0.8
Triglycerides	<1.7	<2.0

Standards from American Diabetes Association (2003).
*In women > 1.3 mmol/L.

The DCCT in the USA compared standard and intensive insulin therapy in a large prospective controlled trial of young patients with type 1 diabetes. Despite intensive therapy, mean blood glucose levels were still 40% above the non-diabetic range, but even at this level of control, the risk of progression to retinopathy was reduced by 60%, nephropathy by 30% and neuropathy by 20% over the 7 years of the study. Near-normoglycaemia should, therefore, be the goal for all young patients with type 1 diabetes. The unwanted effects of this policy include weight gain and a two- to three-fold increase in the risk of severe hypoglycaemia. Control should be less strict in those with a history of recurrent severe hypoglycaemia.

The UKPDS compared standard and intensive treatment in a large prospective controlled trial of type 2 diabetes patients. There was a 25% overall reduction in microvascular disease end-points, a 33% reduction in albuminuria and a 30% reduction in the need for laser treatment for retinopathy in the more intensively treated patients. This study also showed blood pressure control to be equally necessary in the prevention of retinopathy. However, a recent UKPDS study shows that good blood pressure control needs to be continued to maintain the benefits (i.e. there is no 'legacy' effect). There appeared to be little difference in outcome between the agents used to achieve good metabolic control (metformin, sulphonylurea or insulin). Intensive blood pressure control very considerably reduced the cardiovascular risk.

Can established complications be halted or reversed by intensive insulin therapy?

Insulin infusion devices have made near-normal blood glucose control possible for closely supervised groups of patients. Studies in patients with established retinopathy have shown that patients with early retinopathy benefit from intensive therapy, but that patients with more advanced retinal changes generally do not. These observations suggest that microvascular lesions are self-perpetuating once a threshold level of damage has been reached.

Regular checks for patients with diabetes

Box 19.2 is modified from the guidelines set out in *The European Patients' Charter* published by the St Vincent Declaration Steering Committee of the WHO. The charter sets out goals for both the healthcare team and the patient.

Box 19.2 Regular checks for patients with diabetes

Checked each visit

- Review of self-monitoring results and current treatment
- Talk about targets and change where necessary
- Talk about any general or specific problems
- Continued education

Checked at least once a year

- Biochemical assessment of metabolic control (e.g. glycosylated Hb test)
- Measure bodyweight
- Measure blood pressure
- Measure plasma lipids (except in extreme old age)
- Measure visual acuity
- Examine state of retina (ophthalmoscope or retinal photo)
- Test urine for proteinuria/microalbuminuria
- Test blood for renal function (creatinine)
- Check condition of feet, pulses and neurology
- Review cardiovascular risk factors
- Review self-monitoring and injection techniques
- Review eating habits

FURTHER READING

DCCT/Epidemiology of Diabetes Interventions and Complications Research Group. Retinopathy and nephropathy in patients with type 1 diabetes four years after a trial of intensive insulin therapy. *New England Journal of Medicine* 2000; **342**: 381–389.

Eckel RH, Grundy SM, Zimmet PZ. The metabolic syndrome. *Lancet* 2005; **365**: 1415–1428.

Hirsch IB. Insulin analogues. *New England Journal of Medicine* 2005; **352**: 174–183.

Holman RR, Paul SK, Bethel MA et al. Long-term follow-up after tight control of blood pressure in type 2 diabetes. *New England Journal of Medicine* 2008; **359:** 1565–1576.

Holman RR, Paul SK, Bethel MA et al. 10 year follow-up of intensive glucose control in type 2 diabetes. *New England Journal of Medicine* 2008; **359:** 1577–1589.

Johnston D. NICE guidance on type 1 diabetes in adults. *Clinical Medicine* 2004; **4**: 491–493.

Mazzone T, Chait A, Plutzky J. Cardiovascular disease risk in type 2 diabetics. *Lancet* 2008; **371**: 1800–1809.

Nutrition Subcommittee. The implementation of nutritional advice for people with diabetes. *Diabetic Medicine* 2003; **20**: 786–807.

Robertson RP. Islet transplantation as treatment for diabetes. *New England Journal of Medicine* 2004; **350**: 694–705.

Stumvoll M, Goldstein BJ, von Haeften TW. Type 2 diabetes; pathogenesis and therapy. *Lancet* 2005; **365**: 1333–1346.

PSYCHOSOCIAL IMPLICATIONS OF DIABETES

Patients starting tablet or insulin treatment should be encouraged to live as normal a life as possible, but this is not easy. Tact, empathy, encouragement and practical support are needed from all members of the clinical team. Diabetes, like any chronic disease, has psychological sequelae. Most patients will experience periods of not coping, of helplessness, of denial and of acceptance often fluctuating over time. Other problems include:

- *The impossibility to take a 'holiday' from diabetes* – yet the human psyche is poorly developed to cope with unremitting adversity.
- *Concessions or sympathy are often denied* to the person with diabetes, since its presence is invisible.
- *The treatment is complex and demanding*, and the person with diabetes is expected to make trade-offs between short-term and long-term well-being.
- *Embarrassing loss of control over personal behaviour* or consciousness can occur in insulin-treated patients after only a slight miscalculation.
- *Risk-taking behaviour* is indulged in by all humans when emotion is in conflict with logical thought, but its effects can be much greater for the person with diabetes (particularly the risks of unplanned pregnancy, alcohol and tobacco).
- *Poor self-image* is a very common problem.
- *Eating disorders* are more common in people with diabetes – 30–40% of young women with diabetes will exhibit a clinically significant eating disorder.

- *Omission of tablets or insulin* is common since non-adherence to treatment regimens is universal in all illness. Between one in four and one in five tablets prescribed for diabetes is not consumed within the designated treatment period. Insulin omission is very common in young women where the pressure to lose weight overcomes concerns about long-term complications.

Adolescence

Adolescence is a difficult time of life. A period of poor metabolic control, or dropping out of medical care for a time and re-emerging with complications, is very common. Diabetic Summer Camps (e.g. run by Diabetes UK) help prevent a feeling of isolation and not knowing anyone else with the same problem. Separate adolescent clinics allow:

- *treatment without marginalization* in a larger group of older people
- *meeting peers with similar problems* in the waiting room
- *gradual separation from parents* and assumption of personal responsibility for the illness
- *age-appropriate literature* to be available.

Practical aspects

Patients need to inform the driving and vehicle licensing authority and their insurance companies after diagnosis. They would also be wise to inform their family, friends and employers in case unexpected hypoglycaemia occurs. Insulin treatment can be undertaken by people in most walks of life. A few jobs are unsuitable; these include driving Heavy Goods or Public Service vehicles, working at heights, piloting aircraft or working close to dangerous machinery in motion. Certain professions such as the police and the armed forces are barred to all diabetic patients. There are few other limitations, although a considerable amount of ill-informed prejudice exists. Doctors can sometimes help support patients in the face of misinformed work practices.

Table 19.9 Terms used in uncontrolled diabetes

Ketonuria	Detectable ketone levels in the urine; it should be appreciated that ketonuria occurs in fasted non-diabetics and may be found in relatively well-controlled patients with insulin-dependent diabetes mellitus
Ketosis	Elevated plasma ketone levels in the absence of acidosis
Diabetic ketoacidosis	A metabolic emergency in which hyperglycaemia is associated with a metabolic acidosis due to greatly raised (> 5 mmol/L) ketone levels
Hyperosmolar hyperglycaemic state	A metabolic emergency in which uncontrolled hyperglycaemia induces a hyperosmolar state in the absence of significant ketosis
Lactic acidosis	A metabolic emergency in which elevated lactic acid levels induce a metabolic acidosis. In diabetic patients it is rare and associated with biguanide therapy

DIABETIC METABOLIC EMERGENCIES

The main terms used are defined in Table 19.9.

Diabetic ketoacidosis

Diabetic ketoacidosis is the hallmark of type 1 diabetes. It is usually seen in the following circumstances:

- previously undiagnosed diabetes
- interruption of insulin therapy
- the stress of intercurrent illness.

The majority of cases reaching hospital could have been prevented by earlier diagnosis, better communication between patient and doctor, and better patient education. The most common error of management is for patients to reduce or omit insulin because they feel unable to eat owing to nausea or vomiting. This is a factor in at least 25% of all hospital admissions. Insulin may need adjusting up or down but should never be stopped.

Pathogenesis

Ketoacidosis is a state of uncontrolled catabolism associated with insulin deficiency. Insulin deficiency is a necessary precondition since only a modest elevation in insulin levels is sufficient to inhibit hepatic ketogenesis, and stable patients do not readily develop ketoacidosis when insulin is withdrawn. Other factors include counter-regulatory hormone excess and fluid depletion. The combination of insulin deficiency with excess of its hormonal antagonists leads to the parallel processes shown in Figure 19.11. In the absence of insulin, hepatic glucose production accelerates, and periph-

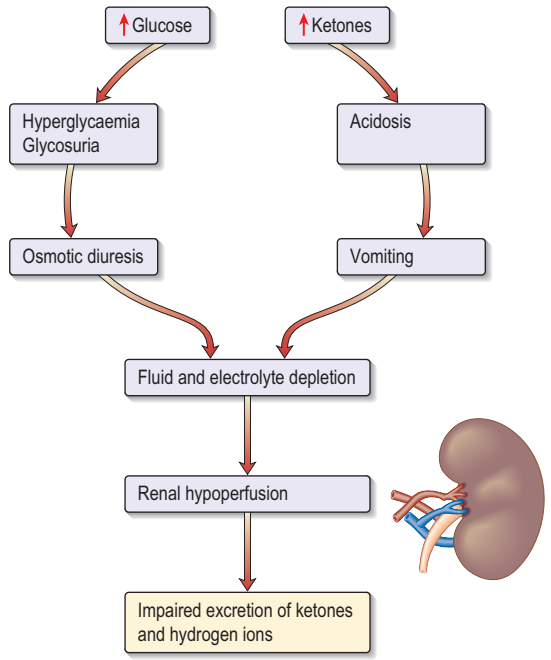


Fig. 19.11 **Dehydration occurs during ketoacidosis as a consequence of two parallel processes.** Hyperglycaemia results in osmotic diuresis, and hyperketonaemia results in acidosis and vomiting. Renal hypoperfusion then occurs and a vicious circle is established as the kidney becomes less able to compensate for the acidosis.

eral uptake by tissues such as muscle is reduced. Rising glucose levels lead to an osmotic diuresis, loss of fluid and electrolytes, and dehydration. Plasma osmolality rises and renal perfusion falls. In parallel, rapid lipolysis occurs, leading to elevated circulating free fatty-acid levels. The free fatty acids are broken down to fatty acyl-CoA within the liver cells, and this in turn is converted to ketone bodies within the mitochondria (Fig. 19.12). Accumulation of ketone bodies produces a metabolic acidosis. Vomiting leads to further loss of fluid and electrolytes. The excess ketones are excreted in the urine but also appear in the breath, producing a distinctive smell similar to that of acetone. Respiratory compensation for the acidosis leads to hyperventilation, graphically described as 'air hunger'. Progressive dehydration impairs renal excretion of hydrogen ions and ketones, aggravating the acidosis. As the pH falls below 7.0 ($[H^+]$ > 100 nmol/L), pH-dependent enzyme systems in many cells function less effectively. Untreated, severe ketoacidosis is invariably fatal.

Clinical features

The features of ketoacidosis are those of uncontrolled diabetes with acidosis, and include prostration, hyperventilation (Kussmaul respiration), nausea, vomiting and, occasionally, abdominal pain. The latter is sometimes so severe that it can be confused with a surgical acute abdomen.

Some patients are mentally alert at presentation, but confusion and stupor are present in more severe cases. Up to 5% present in coma. Evidence of marked dehydration is present and the eyeball is lax to pressure in severe cases. Hyperventilation is present but becomes less marked in very severe acidosis owing to respiratory depression. The smell of ketones on the breath allows an instant diagnosis to be made by those able to detect the odour. The skin is dry and the body temperature is often subnormal, even in the presence of infection; in such cases, pyrexia may develop later.

Diagnosis

This is confirmed by demonstrating hyperglycaemia with ketonaemia or heavy ketonuria, and acidosis. No time should be lost, and treatment is started as soon as the first blood sample has been taken. Hyperglycaemia is demonstrated by dipstick, while a blood sample is sent to the laboratory for confirmation. Ketonaemia is confirmed by centrifuging a blood sample and testing the plasma with a dipstick that measures ketones. Hand-held sensors measuring β-hydroxybutyrate in 30 seconds are available. An arterial blood sample is taken for blood gas analysis.

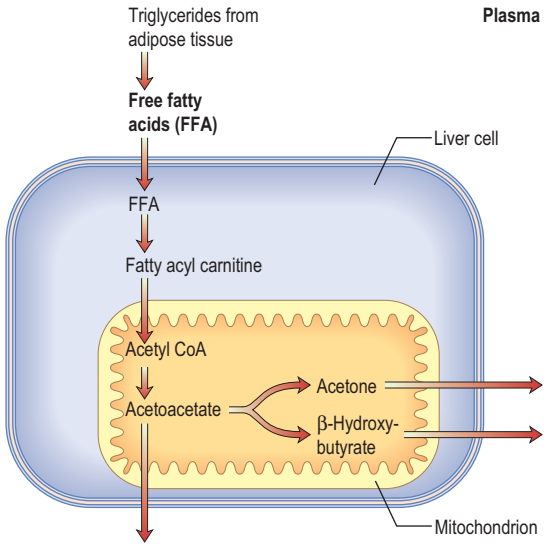


Fig. 19.12 **Ketogenesis.** During insulin deficiency, lipolysis accelerates and free fatty acids taken up by liver cells form the substrate for ketone formation (acetoacetate, acetone and β-hydroxybutyrate) within the mitochondrion. These ketones pass into the blood, producing acidosis.

Management

The principles of management are as follows (Emergency box 19.1); these should be carried out in a high dependency area.

- Replace the *fluid losses* with 0.9% saline. Average loss of water is 5–7 litres with a sodium loss of 500 mmol.
- Replace the *electrolyte losses*. Potassium levels need to be monitored. Patients have a total body potassium deficit of 350 mmol, although initial plasma levels may not be low. Insulin therapy leads to uptake of potassium by the cells with a consequent fall in plasma K^+ levels. Potassium is therefore given as soon as insulin is started.
- *Restore the acid–base balance*. A patient with healthy kidneys will rapidly compensate for the metabolic acidosis once the circulating volume is restored. Bicarbonate is seldom necessary and is only considered if the pH is below 7.0 ($[H^+]$ > 100 nmol/L), and is best given as an isotonic (1.26%) solution.
- *Replace the deficient insulin*. Modern treatment is with relatively modest doses of insulin, which lower blood glucose by suppressing hepatic glucose output rather than by stimulating peripheral uptake, and are therefore much less likely to produce hypoglycaemia. Soluble insulin is given as an intravenous infusion where facilities for adequate supervision exist, or as hourly intramuscular injections. The subcutaneous route is avoided because subcutaneous blood flow is reduced in shocked patients.
- *Monitor blood glucose closely*. Hourly measurement is needed in the initial phases of treatment.
- *Replace the energy losses*. When plasma glucose falls to near-normal values (12 mmol/L), saline infusion should be replaced with 5% dextrose containing 20 mmol/L of potassium chloride. The insulin infusion rate is reduced and adjusted according to blood glucose.
- *Seek the underlying cause*. Physical examination may reveal a source of infection (e.g. a perianal abscess). Two common markers of infection are misleading: fever is unusual even when infection is present, and polymorpholeucocytosis is present even in the absence of infection. Relevant investigations include a chest X-ray, urine and blood cultures, and an ECG (to exclude myocardial infarction). If infection is suspected, broad-spectrum antibiotics are started once the appropriate cultures have been taken.

Problems of management

- *Hypotension*. This may lead to renal shutdown. Plasma expanders (or whole blood) are therefore given if the systolic blood pressure is below 80 mmHg. A central venous pressure line is useful in this situation. A bladder catheter is inserted if no urine is produced within 2

Emergency Box 19.1

Guidelines for the diagnosis and management of diabetic ketoacidosis

Diagnosis
- Hyperglycaemia: measure blood glucose.
- Ketonaemia: test plasma with ketostix. Finger prick sample for β-hydroxybutyrate measurement.
- Acidosis: measure pH and blood gases.

Investigations
- Blood glucose.
- Urea and electrolytes.
- Full blood count.
- Blood gases.
- Blood and urine culture.
- Chest X-ray.
- ECG.
- Cardiac enzymes.

Phase 1 Management
- Admit to HDU.
- Insulin: soluble insulin i.v. 6 units/h by infusion, or 20 units i.m. stat. followed by 6 units i.m. hourly.
- Fluid replacement: 0.9% sodium chloride with 20 mmol KCl per litre. An average regimen would be 1 L in 30 minutes, then 1 L in 1 hour, then 1 L in 2 hours, then 1 L in 4 hours, then 1 L in 6 hours.
- Adjust KCl concentration depending on results of 2 hourly blood K^+ measurement.

IF:
- Blood pressure below 80 mmHg, give plasma expander.
- pH below 7.0 give 500 mL of sodium bicarbonate 1.26% plus 10 mmol KCl. Repeat if necessary to bring pH up to 7.0.

Phase 2 Management
- When blood glucose falls to 10–12 mmol/L change infusion fluid to 1 L 5% dextrose plus 20 mmol KCl 6-hourly. Continue insulin with dose adjusted according to hourly blood glucose test results (e.g. i.v. 3 units/h glucose 15 mmol/L; 2 units/h when glucose 10 mmol/L).

Phase 3 Management
- Once stable and able to eat and drink normally, transfer patient to four times daily subcutaneous insulin regimen (based on previous 24 hours' insulin consumption, and trend in consumption).

Special measures
- Broad-spectrum antibiotic if infection likely.
- Bladder catheter if no urine passed in 2 hours.
- Nasogastric tube if drowsy.
- Consider CVP pressure monitoring if shocked or if previous cardiac or renal impairment.
- Give s.c. prophylactic LMW heparin.

Subsequent management
- Monitor glucose hourly for 8 hours.
- Monitor electrolytes 2-hourly for 8 hours.
- Adjust K replacement according to results.

Note: The regimen of fluid replacement set out above is a guide for patients with severe ketoacidosis. Excessive fluid can precipitate pulmonary and cerebral oedema; inadequate replacement may cause renal failure. Fluid replacement must therefore be tailored to the individual and monitored carefully throughout treatment.

hours, but routine catheterization is not necessary.
- *Coma*. The usual principles apply (see p. 1124). It is essential to pass a nasogastric tube to prevent aspiration, since gastric stasis is common and carries the risk of aspiration pneumonia if a drowsy patient vomits.
- *Cerebral oedema*. This is a rare, but serious complication and has mostly been reported in children or young adults. Excessive rehydration and use of hypertonic fluids such as 8.4% bicarbonate may sometimes be responsible. The mortality is high.
- *Hypothermia*. Severe hypothermia with a core temperature below 33°C may occur and can be overlooked unless a rectal temperature is taken with a low-reading thermometer.
- *Late complications*. These include pneumonia and deep-vein thrombosis (DVT prophylaxis, see p. 446 is essential) and occur especially in the comatose or elderly patient.
- *Complications of therapy*. These include hypoglycaemia and hypokalaemia, due to loss of K^+ in the urine from osmotic diuresis. Overenthusiastic fluid replacement may precipitate pulmonary oedema in the very young or the very old. Hyperchloraemic acidosis may develop in the course of treatment since patients have lost a large variety of negatively charged electrolytes, which are replaced with chloride. The kidneys usually correct this spontaneously within a few days.

Subsequent management

Intravenous dextrose and insulin are continued until the patient feels able to eat and keep food down. The drip is then taken down and a similar amount of insulin is given as four injections of soluble insulin subcutaneously at meal times and a dose of intermediate-acting insulin at night.

Sliding-scale regimens are unnecessary and may even delay the establishment of stable blood glucose levels. The treatment of diabetic ketoacidosis is incomplete without a careful enquiry into the causes of the episode and advice as to how to avoid its recurrence.

Hyperosmolar hyperglycaemic state

This condition, in which severe hyperglycaemia develops without significant ketosis, is the metabolic emergency characteristic of uncontrolled type 2 diabetes. Patients present in middle or later life, often with previously undiagnosed diabetes. Common precipitating factors include consumption of glucose-rich fluids, concurrent medication such as thiazide diuretics or steroids, and intercurrent illness. The hyperosmolar hyperglycaemic state and ketoacidosis represent two ends of a spectrum rather than two distinct disorders (Box 19.3). The biochemical differences may partly be explained as follows:

- *Age*. The extreme dehydration characteristic of hyperosmolar hyperglycaemic state may be related to age. Old people experience thirst less acutely, and more readily become dehydrated. In addition, the mild renal impairment associated with age results in increased urinary losses of fluid and electrolytes.
- *The degree of insulin deficiency*. This is less severe in the hyperosmolar hyperglycaemic state. Endogenous insulin levels are sufficient to inhibit hepatic ketogenesis, whereas glucose production is unrestrained.

Box 19.3 Electrolyte changes in diabetic ketoacidosis and the hyperosmolar hyperglycaemic state

Examples of blood values

	Severe ketoacidosis	Hyperosmolar hyperglycaemic state
Na^+ (mmol/L)	140	155
K^+ (mmol/L)	5	5
Cl^- (mmol/L)	100	110
HCO_3^- (mmol/L)	5	25
Urea (mmol/L)	8	15
Glucose (mmol/L)	30	50
Arterial pH	7.0	7.35

The normal range of osmolality is 285–300 mOsm/kg. It can be measured directly, or can be calculated approximately from the formula:

Osmolality = $2(Na^+ + K^+)$ + glucose + urea.

For instance, in the example of severe ketoacidosis given above:

Osmolality = 2(140 + 5) + 30 + 8 = 328 mOsm/kg

and in the example of the hyperosmolar hyperglycaemic state:

Osmolality = 2(155 + 5) + 50 + 15 = 385 mOsm/kg.

The normal anion gap is less than 17. It is calculated as $(Na^+ + K^+) - (Cl^- + HCO_3^-)$. In the example of ketoacidosis the anion gap is 40, and in the example of the hyperosmolar hyperglycaemic state the anion gap is 25. Mild hyperchloraemic acidosis may develop in the course of therapy. This will be shown by a rising plasma chloride and persistence of a low bicarbonate even though the anion gap has returned to normal.

Clinical features

The characteristic clinical features on presentation are dehydration and stupor or coma. Impairment of consciousness is directly related to the degree of hyperosmolality. Evidence of underlying illness such as pneumonia or pyelonephritis may be present, and the hyperosmolar state may predispose to stroke, myocardial infarction or arterial insufficiency in the lower limbs.

Investigations and treatment

These are (with some exceptions) according to the guidelines for ketoacidosis. The plasma osmolality is usually extremely high. It can be measured directly or calculated as $(2(Na^+ + K^+)$ + glucose + urea), all in mmol/L. Many patients are extremely sensitive to insulin, and the glucose concentration may plummet. The resultant change in osmolality may cause cerebral damage. It is sometimes useful to infuse insulin at a rate of 3 U per hour for the first 2–3 hours, increasing to 6 U/h if glucose is falling too slowly. The standard fluid for replacement is 0.9% physiological saline. Avoid 0.45% saline, since rapid dilution of the blood may cause more cerebral damage than a few hours of exposure to hypernatraemia.

Prognosis

The reported mortality ranges as high as 20–30%, mainly because of the advanced age of the patients and the frequency of intercurrent illness. Unlike ketoacidosis, the hyperosmolar hyperglycaemia state is not an absolute indication for subsequent insulin therapy, and survivors may do well on diet and oral agents.

Lactic acidosis

Lactic acidosis may occur in diabetic patients on biguanide therapy. The risk in patients taking metformin is extremely low provided that the therapeutic dose is not exceeded and the drug is withheld in patients with advanced hepatic or renal dysfunction.

Patients present in severe metabolic acidosis with a large anion gap (normally less than 17 mmol/L), usually without significant hyperglycaemia or ketosis. Treatment is by rehydration and infusion of isotonic 1.26% bicarbonate. The mortality is in excess of 50%.

COMPLICATIONS OF DIABETES

Insulin-treated patients still have a considerably reduced life expectancy. The major cause of death in treated patients is due to cardiovascular problems (70%) followed by renal failure (10%) and infections (6%). There is no doubt that the duration and degree of hyperglycaemia play a major role in the production of complications. Better diabetic control can reduce the rate of progression of both nephropathy and retinopathy and the DCCT showed a 60% reduction in developing complications over 9 years when the HbA_{1c} was kept at around 7% in type 1 diabetes.

Pathophysiology

The mechanisms leading to damage are ill defined. The following are consequences of hyperglycaemia and may play a role:

- *Non-enzymatic glycosylation* of a wide variety of proteins, e.g. haemoglobin, collagen, LDL and tubulin in peripheral nerves. This leads to an accumulation of advanced glycosylated end-products causing injury and inflammation via stimulation of pro-inflammatory factors, e.g. complement, cytokines.
- *Polyol pathway*. The metabolism of glucose by increased intracellular aldose reductase leads to accumulation of sorbitol and fructose. This causes changes in vascular permeability, cell proliferation and capillary structure via stimulation of protein kinase C and TGF-β.
- *Abnormal microvascular blood flow* impairs supply of nutrients and oxygen. Microvascular occlusion is due to vasoconstrictors, e.g. endothelins and thrombogenesis, and leads to endothelial damage.
- *Other factors* include the formation of reactive oxygen species and growth factors stimulation (TGF-β) and vascular endothelial growth factor (VEGF). These growth factors are released by ischaemic tissues and cause endothelial cells to proliferate.
- *Haemodynamic changes*, e.g. in kidney (see p. 1053).

It has been proposed that all of the above mechanisms stem from a single hyperglycaemia-induced process of overproduction of superoxide by the mitochondrial electron chain. This new paradigm offers an integrated explanation of how complications of diabetes develop.

Macrovascular complications (Table 19.10)

Diabetes is a risk factor in the development of atherosclerosis. This risk is related to that of the background population. For example, Japanese diabetics are much less likely than patients in Europe to develop atherosclerosis, but are much more likely to develop it than are non-diabetic Japanese. The excess risk to diabetics compared with the general population increases as one moves down the body:

- Stroke is twice as likely.
- Myocardial infarction is three to five times as likely, and women with diabetes lose their premenopausal protection from coronary artery disease.
- Amputation of a foot for gangrene is 50 times as likely.

The UKPDS and DCCT studies have shown that intensive treatment of diabetes has only a small effect upon the cardiovascular risk of type 2 or type 1 diabetes patients.

Cardiovascular risk factors occurring together tend to have a multiplicative effect on the overall level of cardiovascular risk. It is vital to tackle all cardiovascular risk factors together in diabetes, and not just to focus on glucose levels.

- *Hypertension*. The UKPDS demonstrated that aggressive treatment of hypertension produces a marked reduction in adverse cardiovascular outcomes, both microvascular and macrovascular. To achieve the target for blood pressure (Table 19.8), the UKPDS found that one-third of patients needed three or more antihypertensive drugs in combination, and two-thirds of treated patients needed two or more.
- *Smoking*: the avoidable risk factor (see p. 829). Never give up efforts to help diabetic patients stop smoking.
- *Lipid abnormalities*. Clinical trials suggest that there is no 'safe' cut-off for serum cholesterol. The lowest achievable level seems best to aim for, and in practice this means that almost all people with type 2 diabetes will be treated with a statin.
- *Low-dose aspirin* can reduce macrovascular risk, but is associated with a morbidity and mortality from bleeding. The benefits of aspirin outweigh the bleeding risk when the risk of a cardiovascular end-point is >30% in the next 10 years. This risk is reached in patients aged under 45 with three strong additional cardiovascular risk factors, aged 45–54 with three additional risk factors, aged 54–65 with two additional risk factors or aged over 65 with just one additional risk factor.
- *ACE inhibitors/angiotensin II receptor antagonists*. Treating people with diabetes and at least one other major cardiovascular risk factor with an ACE-inhibitor produces a 25–35% lowering of the risk of heart attack, stroke, overt nephropathy or cardiovascular death. Angiotensin II receptor antagonists are sometimes preferred initially and are also used for those intolerant to ACE inhibitors.

Table 19.10 Diabetic risk factors for macrovascular complications

Duration
Increasing age
Systolic hypertension
Hyperinsulinaemia due to insulin resistance associated with obesity and syndrome X (metabolic syndrome)
Hyperlipidaemia, particularly hypertriglyceridaemia/low HDL
Proteinuria (including microalbuminuria)
Other factors are the same as for the general population

Microvascular complications

In contrast to macrovascular disease, which is prevalent in the West as a whole, microvascular disease is specific to diabetes. Small blood vessels throughout the body are affected but the disease process is of particular danger in three sites:

- retina
- renal glomerulus
- nerve sheaths.

Diabetic retinopathy, nephropathy and neuropathy tend to manifest 10–20 years after diagnosis in young patients. They present earlier in older patients, probably because these have had unrecognized diabetes for months or even years prior to diagnosis. Genetic factors appear to contribute to the susceptibility to microvascular disease. Diabetic siblings of diabetic patients with renal and eye disease have a three- to fivefold increased risk of the same complication in both type 1 and type 2 patients. There are racial differences in the overall prevalence of nephropathy. In the USA prevalence is: Pima American Indians > Hispanic/Mexican > US black > US white patients.

Diabetic eye disease

About one in three young people in this patient population is likely to develop visual problems, and in the UK 5% have in the past become blind after 30 years of diabetes. Diabetes is still the commonest cause of blindness in under 65 year old people despite the widespread use of laser photocoagulation and control of blood pressure.

It affects the eye in a variety of ways:

- *Diabetic retinopathy* with lesions developing in the retina and in the iris.
- *Cataract* which develops earlier in diabetes than in the general population. Fluctuations in blood sugar can in addition give rise to some alterations in the refractive error as a result of osmotic changes within the lens. With the absorption of water into the lens the eye becomes more hypermetropic. This change occurs most commonly in young type 1 diabetics resulting in difficulty in reading. With control of blood sugar this is frequently reversed. However with high levels of blood sugar often associated with a degree of ketosis an acute cataract may develop (snowflake cataract) which comes on very rapidly and does not clear.
- *External ocular* palsies (p. 1104).
- The sixth and the third nerve are the most commonly affected. Third nerve palsy is not associated with pain. These nerve palsies usually recover spontaneously within a period of 3–6 months.

Natural history

Cataracts

Diabetic cataracts occur at an earlier stage than in non-diabetics.

Ask the authors
www.studentconsult.com

Diabetic retinopathy

Diabetic retinopathy (Fig. 19.13) increases with the length of diabetes, 20% will have retinal changes after 10 years, rising to 80% after 20 years (see Table 19.11).

The effects on the retina initially differ with the different types of diabetes. In type 1 diabetes the progression may be fairly rapid to proliferative retinopathy (Fig. 19.14), often accelerated with poor control of diabetes and blood pressure. With good control the progression is slower (DCCT study). In type 2 diabetes the progression is much slower and initially in the macular and paramacular region. These changes are associated with macular oedema and consequent loss of vision. This maculopathy again can be slowed by good control of diabetes and blood pressure (UKPDS). Patients with type 2 diabetes may go on to develop proliferative retinopathy and patients with type 1 may go on to develop maculopathy although this is less common.

Background diabetic retinopathy. The first sign is the development of *microaneurysms* (small red dots) within the retina around the posterior pole. On fluorescein angiography they are associated with areas of capillary non-perfusion. Other features are *superficial haemorrhages* in the ganglion cell layer and outer plexiform layer. These haemorrhages are blotch-like in appearance (like ink on a blotter) and are the consequence of capillary closure. *Cotton wool spots* are micro-infarcts within the retina and the spot itself is the accumulation of axoplasmic debris. This debris is removed by macrophages and as this occurs there may be white dots at the site of the previous cotton wool spot (cytoid bodies).

Cotton wool spots may be associated with high blood pressure but may be entirely due to the diabetes. The cotton wool spots of hypertension tend to resolve quickly whereas those in diabetes can last over 3–6 months.

Pre-proliferative and proliferative retinopathy Proliferative retinopathy is preceded by the widespread development of capillary non-perfusion. This ischaemia induces new blood vessels to grow. Some of these blood vessels are inside the retina and therefore represent a healing process. These intraretinal new vessels (IRMA) do not give rise to any symptoms.

Preretinal new vessels (when the vessel has come through the retina) lie on the surface of the retina and usually occur at the margin of the area of capillary closure. They represent aborted attempts at revascularization: they are induced to grow because of the accumulation of VEGF (vascular endothelial growth factor). However these new vessels are prone to bleed, particularly if there is vitreous traction. Small haemorrhages give rise to the preretinal haemorrhage (boat-shaped haemorrhage) that lies on the inferior half of the retina. With further bleeding or traction, the blood seeps into the vitreous with the consequent loss of vision. Once new vessels have developed they undergo evolution with collagen tissue growing along the margins of the vessel and giving rise to fibrotic traction bands. These bands may pull on the retina and cause further haemorrhaging and possible retinal detachment. Abnormalities in the veins may be a strong indication of capillary non-perfusion. Beading occurs as the vein passes through an area of ischaemia and venous loops result from closure of the vein at the margin of an area of capillary non-perfusion. These two are clinical signs which are a good indication of advanced capillary closure.

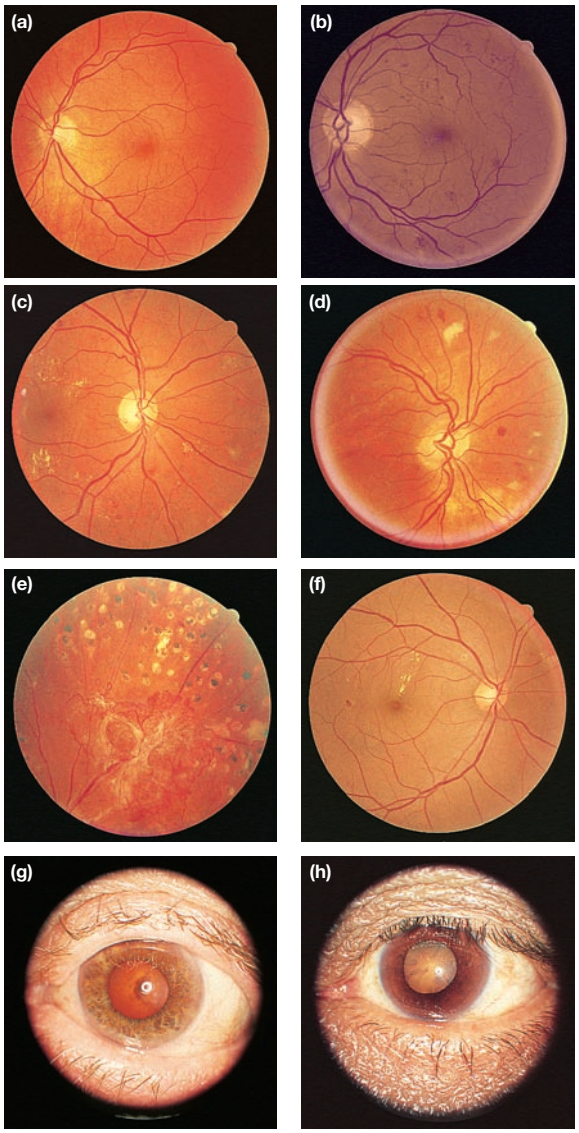


Fig. 19.13 **Features of diabetic eye disease. (a)** The normal macula (centre) and optic disc. **(b)** Dot and blot haemorrhages (early background retinopathy). **(c)** Hard exudates are present in addition in background retinopathy. **(d)** Multiple cotton-wool spots indicate pre-proliferative retinopathy requiring routine ophthalmic referral. **(e)** Multiple frond-like new vessels, the hallmark of proliferative retinopathy. White fibrous tissue is forming near the new vessels, a feature of advanced retinopathy (this eye also illustrates multiple xenon arc laser burns superiorly). **(f)** Exudates appearing within a disc-width of the macula are a feature of an exudative maculopathy. **(g)** Central and **(h)** cortical cataracts can be seen against the red reflex with the ophthalmoscope.

If the capillary non-perfusion is rapidly developing and widespread then vessels may be induced to grow on the pupil margin and then ultimately in the angle of the anterior chamber giving rise to rapid increase in intraocular pressure (thrombotic glaucoma). If the capillary closure should extend across the macular area, there will be associated loss of central vision.

Maculopathy. In type 2 diabetics the aneurysms increase in number and leak, giving rise to fluid accumulation within the

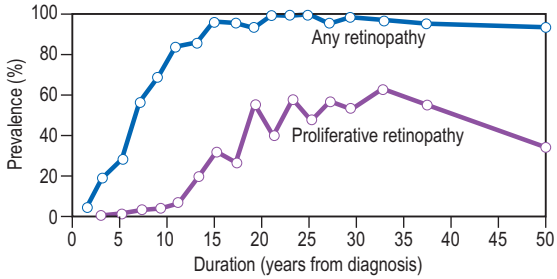


Fig. 19.14 **Prevalence of retinopathy in relation to duration of the disease in patients with type 1 diabetes mellitus diagnosed under the age of 33 years.** Almost all eventually develop background change and 60% progress to proliferative retinopathy. From *Archives of Ophthalmology* 1984; **102**: 520.

Box 19.4 Criteria for a successful local screening scheme for sight-threatening diabetic retinopathy

- Clearly defined geographical area for the screening programme
- Adequate number of people with diabetes for viability (>12 000)
- An identified screening programme manager
- An identified clinical screening lead
- An identified hospital eye service for diagnosis and laser treatment
- Computer software capable of supporting call/recall of patients and image grading
- Centralized appointment administration
- Single collated list of all people with diabetes in the area over the age of 12
- Equipment to obtain adequate disc and macula centred images of each eye
- Single image grading centre
- Process to manage people with poor quality images
- Clear route of referral for treatment, and for feedback from treatment centre to screening unit
- Accreditation of screening staff
- Annual reporting of service performance

retina. If these aneurysms are localized they are associated with the deposition of fat and protein in the surrounding area (circinate retinopathy). If the oedema extends into the macula then the retina becomes thickened, visual function deteriorates and there is loss of central vision. Later new vessels may develop giving rise again to recurrent progressive vitreous haemorrhages. The final outcome in patients with maculopathy is for a large exudative plaque to be present in the central macular area.

Mixed retinopathy. Although proliferative and maculopathy are fairly distinct, frequently there is a combination of both types of retinopathy.

Examination

Examination of the eye. Visual acuity should be checked with the best distance glasses using a pinhole. The *ocular movements* are assessed to detect any ocular motor palsies. The *iris* is examined for rubeosis and then the pupils dilated with 1% tropicamide. About 20 minutes later the eye is examined for the presence of a cataract by looking at the lens with a +10.00 lens in the ophthalmoscope and viewing the lens against the red reflex. If there is any opacity in the central area then this will be the cause of visual loss. The *retina* is then examined and a systematic examination should always be carried out, looking at the disc, in all quadrants, and then finally the macula. The *macula* is examined last because this induces the greatest pupillary constriction.

Screening for sight. Threatening eye disease with universal access is now seen as offering the best hope of displacing diabetes as the commonest cause of blindness in those under 65 years of age. The National Screening Committee in the UK has helped establish digital photography-based screening across the country based on a national set of standards. All patients with diabetes, over the age of 12 years, have annual measurements of their acuity, and photographs of their retina taken. Box 19.4 shows standardized criteria for screening schemes; these are regularly inspected.

Management of diabetic eye disease (Table 19.11)

- *Cataract* extraction is indicated if the cataract is causing visual disability to the patient or is giving rise to inability to view the retina adequately. Cataract extraction is straightforward if there is no retinopathy present. However, if there is retinopathy and/or iris neovascularization, both of these can progress. In addition the development of macular oedema if not pre-existing is more common and in the presence of existing macular oedema this can deteriorate fairly rapidly. Cataract extraction is indicated for visual reasons and also if it obstructs the view of the retina.
- *Retinopathy.* The DCCT and UKPDS show that the risk of developing diabetic eye disease can be reduced by striving for aggressive metabolic control of the diabetes. There is no specific medical treatment for background retinopathy. Smoking and hypertension worsen the rate of progression, and blood pressure control is of particular importance. Development or progression of retinopathy may be accelerated by rapid improvement in glycaemic control, pregnancy and in those with nephropathy, and these groups need frequent monitoring. An ophthalmologist may perform fluorescein angiography (a fluorescent dye is injected into an arm vein and photographed in transit through the retinal vessels) to define the extent of the potentially sight-threatening diabetic retinopathy.
- *Proliferative retinopathy.* Once new vessels have developed this is an indication for laser therapy. The laser should be directed at the new vessels and, in addition, to the associated areas of capillary non-perfusion (ischaemia). If the proliferative retinopathy has progressed to new vessels developing on the optic disc then a technique known as *panretinal photocoagulation (PRP)* is carried out. This involves multiple laser burns to the peripheral retinopathy again to the areas of capillary non-perfusion. The presence of new vessels on the disc carries the worst prognosis and therefore laser therapy should be carried out as rapidly as possible. If bleeding has occurred and there is a good view then laser treatment should be applied. If the bleeding is recurrent and affecting vision and/or if there is the presence of fibrous tissue giving rise to traction on the retina, then

Table 19.11 Grading of pathological changes in the retina in diabetic retinopathy: the action needed

Retinopathy grade	Retinal abnormality (cause)	Action needed
Peripheral retina		
Non-proliferative/background	Dot haemorrhages (capillary microaneurysms) (usually appear first)	Annual screening only
	Blot haemorrhages (leakage of blood into deeper retinal layers)	
	Hard exudates (exudation of plasma rich in lipids and protein)	
Pre-proliferative	Venous beading/loops	Non-urgent referral to an ophthalmologist
	Intraretinal microvascular abnormalities – IRMAs	
	Multiple cotton wool spots	
Proliferative	New blood vessel formation/neovascularization	Urgent referral to an ophthalmologist
	Preretinal or subhyaloid haemorrhage	
	Vitreous haemorrhage	
Advanced retinopathy	Retinal fibrosis Traction retinal detachment	Urgent referral to an ophthalmologist – but much vision already lost
Central retina		
Maculopathy	Hard exudates within one disc-width of macula	Referral to an ophthalmologist soon

Note: Hard exudates have a bright yellowish white colour and are often irregular in outline with a sharply defined margin. Cotton-wool spots are greyish white, have indistinct margins and a dull matt surface, unlike the glossy appearance of hard exudates.

vitrectomy is indicated. The presence of rubeosis indicates that there is widespread peripheral closure and panretinal photocoagulation should be performed.

- *Maculopathy*. Extrafoveal lesions can be watched. However if they are beginning to encroach on the fovea then the centre of the circinate exudate rings, where the aneurysms lie, should be treated by laser photocoagulation. If the oedema has already spread into the centre, then a technique known as grid photocoagulation is used, where a number of laser burns are scattered around the macula in the area of oedema. Although this on occasion does result in some improvement in vision and some improvement of the oedema, it is not always successful.

The future

The control of diabetes and blood pressure, and argon laser photocoagulation, has been a major advance in the management of patients with diabetic retinopathy. New laser systems are undergoing clinical trials and are already showing promise as they result in less damage to the retina and therefore less loss of function. At the moment there are no drugs available for the treatment of diabetic retinopathy but it is hoped that in the future anti-VEGF drugs (see p. 464) will be used more to control retinopathy.

Vitreoretinal surgery can be performed to try to salvage some vision after vitreous haemorrhage and to treat traction retinal detachment in advanced retinopathy.

The diabetic kidney

The kidney may be damaged by diabetes in three main ways:

- glomerular damage
- ischaemia resulting from hypertrophy of afferent and efferent arterioles
- ascending infection.

Diabetic nephropathy

Epidemiology

Clinical nephropathy secondary to glomerular disease usually manifests 15–25 years after diagnosis of diabetes and affects 25–35% of patients diagnosed under the age of 30 years. It is the leading cause of premature death in young diabetic patients. Older patients also develop nephropathy, but the proportion affected is smaller. Some centres have reported a falling incidence rate of diabetic nephropathy in type 1 diabetes. This may reflect good-quality local care for diabetes rather than a change in the natural history of the disease itself. As people with type 2 diabetes develop diabetes progressively younger, a rising incidence rate of nephropathy is seen.

Pathophysiology

The earliest functional abnormality in the diabetic kidney is renal hypertrophy associated with a raised glomerular filtration rate. This appears soon after diagnosis and is related to poor glycaemic control. As the kidney becomes damaged by diabetes, the afferent arteriole (leading to the glomerulus) becomes vasodilated to a greater extent than the efferent glomerular arteriole. This increases the intraglomerular filtration pressure, further damaging the glomerular capillaries. This increased intraglomerular pressure also leads to increased shearing forces locally which are thought to contribute to mesangial cell hypertrophy and increased secretion of extracellular mesangial matrix material. This process eventually leads to glomerular sclerosis. The initial structural lesion in the glomerulus is thickening of the basement membrane. Associated changes result in disruption of the protein cross-linkages which normally make the membrane an effective filter. In consequence, there is a progressive leak of large molecules (particularly protein) into the urine.

Albuminuria

The earliest evidence of this is 'microalbuminuria' – amounts of urinary albumin so small as to be undetectable by standard dipsticks (see p. 579). Microalbuminuria may be tested for by radioimmunoassay or by using special dipsticks. It is a predictive marker of progression to nephropathy in type 1 diabetes, and of increased cardiovascular risk in type 2 diabetes. Microalbuminuria may, after some years, progress to intermittent albuminuria followed by persistent proteinuria. Light-microscopic changes of glomerulosclerosis become manifest; both diffuse and nodular glomerulosclerosis can occur. The latter is sometimes known as the Kimmelstiel–Wilson lesion. At the later stage of glomerulosclerosis, the glomerulus is replaced by hyaline material.

At the stage of persistent proteinuria, the plasma creatinine is normal but the average patient is only some 5–10 years from end-stage kidney disease. The proteinuria may become so heavy as to induce a transient nephrotic syndrome, with peripheral oedema and hypoalbuminaemia.

Patients with nephropathy typically show a normochromic normocytic anaemia and a raised erythrocyte sedimentation rate (ESR). Hypertension is a common development and may itself damage the kidney still further. A rise in plasma creatinine is a late feature that progresses inevitably to renal failure, although the rate of progression may vary widely between individuals.

The natural history of this process is shown in Figure 19.15.

Ischaemic lesions

Arteriolar lesions, with hypertrophy and hyalinization of the vessels, can occur in patients with diabetes. The appearances are similar to those of hypertensive disease and lead to ischaemic damage to the kidneys.

Infective lesions

Urinary tract infections are relatively more common in women with diabetes, but this does not apply to men. Ascending infection may occur because of bladder stasis resulting from autonomic neuropathy, and infections more easily become established in damaged renal tissue. Autopsy material frequently reveals interstitial changes suggestive of infection, but ischaemia may produce similar changes and the true frequency of pyelonephritis in diabetes is uncertain. Untreated infections in diabetics can result in renal papillary necrosis, in which renal papillae are shed in the urine, but this complication is rare.

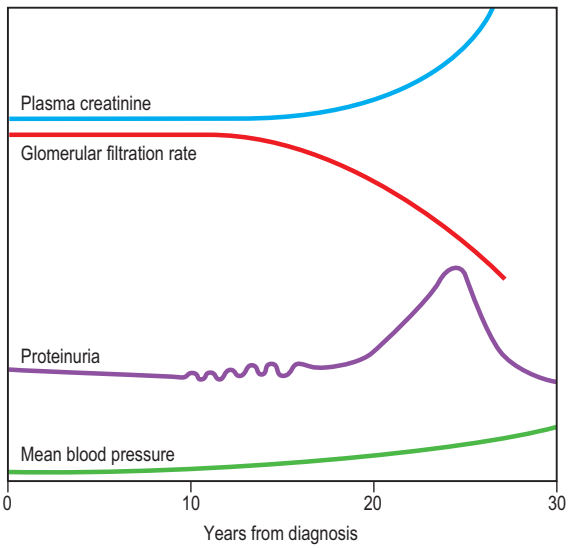


Fig. 19.15 **Schematic representation of the natural history of nephropathy.** The typical onset is 15 years after diagnosis. Intermittent proteinuria leads to persistent proteinuria. In time, the plasma creatinine rises as the glomerular filtration rate falls.

Diagnosis

The urine of all diabetic patients should be checked regularly for the presence of protein. Many centres also screen younger patients for microalbuminuria since there is evidence that meticulous glycaemic control or early antihypertensive treatment at this stage may delay the onset of frank proteinuria. Once proteinuria is present, other possible causes for this should be considered (see below), but once these are excluded, a presumptive diagnosis of diabetic nephropathy can be made. For practical purposes this implies inevitable progression to end-stage kidney disease, although the time course can be markedly slowed by early aggressive antihypertensive therapy. Clinical suspicion of a non-diabetic cause of nephropathy may be provoked by an atypical history, the absence of diabetic retinopathy (usually but not invariably present with diabetic nephropathy) and the presence of red-cell casts in the urine. Renal biopsy should be considered in such cases, but in practice is rarely necessary or helpful. A 24-hour urine collection is performed to quantify protein loss. Regular measurement is made of the plasma creatinine level with estimated glomerular filtration rate (GFR).

Management

The management of diabetic nephropathy is similar to that of other causes of chronic kidney disease, with the following provisos:

- Aggressive treatment of blood pressure with a target below 130/80 mmHg has been shown to slow the rate of deterioration of renal failure considerably. Angiotensin-converting enzyme inhibitors or an angiotensin receptor II antagonist are the drugs of choice (see p. 804). These drugs should also be used in normotensive patients with persistent microalbuminuria. Reduction in albuminuria occurs with this treatment.
- Oral hypoglycaemic agents partially excreted via the kidney (e.g. glibenclamide and metformin) should be avoided.
- Insulin sensitivity increases and drastic reductions in insulin dosage may be needed.
- Associated diabetic retinopathy tends to progress rapidly, and frequent ophthalmic supervision is essential.

Management of end-stage disease is made more difficult by the fact that patients often have other complications of diabetes such as blindness, autonomic neuropathy or peripheral vascular disease. Vascular shunts tend to calcify rapidly and hence chronic ambulatory peritoneal dialysis may be preferable to haemodialysis. The failure rate of renal transplants is somewhat higher than in non-diabetic patients. A segmental pancreatic graft is sometimes performed at the same time as a renal graft. Although pancreatic transplants

have a limited viability, owing to progressive fibrosis within the graft, they may give the patient a year or so of freedom from insulin injections.

Diabetic neuropathy

Diabetes can damage peripheral nervous tissue in a number of ways. The vascular hypothesis postulates occlusion of the vasa nervorum as the prime cause. This seems likely in isolated mononeuropathies, but the diffuse symmetrical nature of the common forms of neuropathy implies a metabolic cause. Since hyperglycaemia leads to increased formation of sorbitol and fructose in Schwann cells, accumulation of these sugars may disrupt function and structure.

The earliest functional change in diabetic nerves is delayed nerve conduction velocity; the earliest histological change is segmental demyelination, caused by damage to Schwann cells. In the early stages axons are preserved, implying prospects of recovery, but at a later stage irreversible axonal degeneration develops.

The following varieties of neuropathy occur (Fig. 19.16):

- symmetrical mainly sensory polyneuropathy (distal)
- acute painful neuropathy
- mononeuropathy and mononeuritis multiplex
  (a) cranial nerve lesions
  (b) isolated peripheral nerve lesions
- diabetic amyotrophy (asymmetrical motor diabetic neuropathy)
- autonomic neuropathy.

Symmetrical mainly sensory polyneuropathy

This is often unrecognized by the patient in its early stages. Early clinical signs are loss of vibration sense, pain sensation (deep before superficial) and temperature sensation in the feet. At later stages patients may complain of a feeling of 'walking on cotton wool' and can lose their balance when washing the face or walking in the dark owing to impaired proprioception. Involvement of the hands is much less common and should prompt a search for non-diabetic causes. Complications include unrecognized trauma, beginning as blistering due to an ill-fitting shoe or a hot-water bottle, and leading to ulceration.

Sequelae of neuropathy. Involvement of motor nerves to the small muscles of the feet gives rise to interosseous wasting. Unbalanced traction by the long flexor muscles leads to a characteristic shape of the foot, with a high arch and clawing of the toes, which in turn leads to abnormal distribution of pressure on walking, resulting in callus formation under the first metatarsal head or on the tips of the toes and perforating neuropathic ulceration. *Neuropathic arthropathy (Charcot's joints)* may sometimes develop in the ankle. The hands show small-muscle wasting as well as sensory changes, but these signs and symptoms must be differentiated from those of the carpal tunnel syndrome, which occurs with increased frequency in diabetes and may be amenable to surgery.

Acute painful neuropathy

A diffuse, painful neuropathy is less common. The patient describes burning or crawling pains in the feet, shins and anterior thighs. These symptoms are typically worse at night, and pressure from bedclothes may be intolerable. It may present at diagnosis or develop after sudden improvement in glycaemic control (e.g. when insulin is started). It usually remits spontaneously after 3–12 months if good control is maintained. A more chronic form, developing later in the course of the disease, is sometimes resistant to almost all forms of therapy. Neurological assessment is difficult because of the hyperaesthesia experienced by the patient, but muscle wasting is not a feature and objective signs can be minimal.

Management is firstly to explore for non-diabetic causes (see p. 1175). Explanation and reassurance about the high likelihood of remission within months may be all that is needed. Tricyclics, gabapentin, pregabalin, duloxetine, mexiletine, valproate and carbamazepine all reduce the perception of neuritic pain somewhat, but usually not as much as patients hope for. Transepidermal nerve stimulation (TENS) benefits some patients. Topical capsaicin-containing creams help occasionally. A few report that acupuncture has helped.

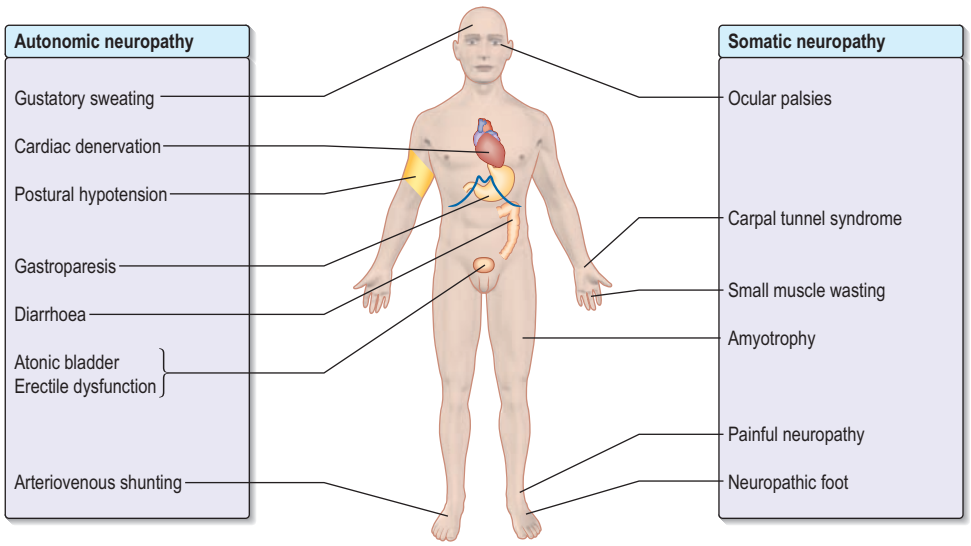


Fig. 19.16 **The neuropathic man.**

Mononeuritis and mononeuritis multiplex (multiple mononeuropathy)

Any nerve in the body can be involved in diabetic mononeuritis; the onset is typically abrupt and sometimes painful. Radiculopathy (i.e. involvement of a spinal root) may also occur.

Isolated palsies of nerves to the external eye muscles, especially the third and sixth nerves, are more common in diabetes. A characteristic feature of diabetic third nerve lesions is that pupillary reflexes are retained owing to sparing of pupillomotor fibres. Full spontaneous recovery is the rule for most episodes of mononeuritis over 3–6 months. Lesions are more likely to occur at common sites for external pressure palsies or nerve entrapment (e.g. the median nerve in the carpal tunnel, see p. 1172).

Diabetic amyotrophy

This condition is usually seen in older men with diabetes. Presentation is with painful wasting, usually asymmetrical, of the quadriceps muscles or occasionally in the shoulders. The wasting may be very marked and knee reflexes are diminished or absent. The affected area is often extremely tender. Extensor plantar responses sometimes develop and CSF protein content is elevated. Diabetic amyotrophy is usually associated with periods of poor glycaemic control and may be present at diagnosis. It often resolves in time with careful metabolic control of the diabetes.

Autonomic neuropathy

Asymptomatic autonomic disturbances can be demonstrated on laboratory testing in many patients, but symptomatic autonomic neuropathy is rare. It affects both the sympathetic and parasympathetic nervous systems and can cause disabling postural hypotension.

The cardiovascular system

Vagal neuropathy results in tachycardia at rest and loss of sinus arrhythmia. At a later stage, the heart may become denervated (resembling a transplanted heart). Cardiovascular reflexes such as the Valsalva manoeuvre are impaired. Postural hypotension occurs owing to loss of sympathetic tone to peripheral arterioles. A warm foot with a bounding pulse is sometimes seen in a polyneuropathy as a result of peripheral vasodilatation.

Gastrointestinal tract

Vagal damage can lead to gastroparesis, often asymptomatic, but sometimes leading to intractable vomiting. Implantable devices which stimulate gastric emptying, and injections of botulinum toxin into the pylorus (to partly paralyse the sphincter), have each shown benefit in cases of this previously intractable problem. Autonomic diarrhoea often occurs at night accompanied by urgency and incontinence. Diarrhoea and steatorrhoea may occur owing to small bowel bacterial overgrowth; treatment is with antibiotics such as tetracycline.

Bladder involvement

Loss of tone, incomplete emptying and stasis (predisposing to infection) can occur, and may ultimately result in an atonic, painless, distended bladder. Treatment is with intermittent self-catheterization, permanent catheterization if that fails and prophylactic antibiotic therapy for those prone to recurrent infection.

Male erectile dysfunction

This is common. The first manifestation is incomplete erection which may in time progress to total failure; retrograde ejaculation also occurs in patients with autonomic neuropathy. Erectile dysfunction in diabetes has many causes including anxiety, depression, alcohol excess, drugs (e.g. thiazides and beta-blockers), primary or secondary gonadal failure, hypothyroidism, and inadequate vascular supply owing to atheroma in pudendal arteries. The history and examination should focus on these possible causes. Blood is taken for LH, FSH, testosterone, prolactin and thyroid function. Treatment should ideally include sympathetic counselling of both partners.

Phosphodiesterase type-5 inhibitors (sildenafil, tadalafil, vardenafil), which enhance the effects of nitric oxide on smooth muscle and increase penile blood flow, are used in those who do not take nitrates for angina. Sixty per cent of diabetic patients can be expected to benefit from this therapy.

Alternatives for those who fail to improve with a phosphodiesterase inhibitor, who dislike the side-effects (headache and a green tinge to vision the next day), or those in whom it is contraindicated, are:

- Apomorphine 2 or 3 mg sublingually 20 minutes before sexual activity.
- Alprostadil (prostaglandin E1 preparation) given as a small pellet inserted with a device into the urethra (125 μg initially with a maximum of 500 μg). If the partner is pregnant, barrier contraception must be used to keep prostaglandin away from the fetus.
- Intracavernosal injection of alprostadil urethra (2.5 μg initially with a maximum of 40 μg). Side-effects include priapism which needs urgent treatment should erection last more than 3 hours.
- Vacuum devices.

The diabetic foot

Ten to fifteen per cent of diabetic patients develop foot ulcers at some stage in their lives. Diabetic foot problems are responsible for nearly 50% of all diabetes-related hospital admissions. Many diabetic limb amputations could be delayed or prevented by more effective patient education and medical supervision. Ischaemia, infection and neuropathy combine to produce tissue necrosis. Although all these factors may coexist, the ischaemic and the neuropathic foot (Table 19.12) can be distinguished. In rural India foot ulcers

Table 19.12 Distinguishing features between ischaemia and neuropathy in the diabetic foot

	Ischaemia	Neuropathy
Symptoms	Claudication	Usually painless
	Rest pain	Sometimes painful neuropathy
Inspection	Dependent rubor	High arch
	Trophic changes	Clawing of toes
		No trophic changes
Palpation	Cold	Warm
	Pulseless	Bounding pulses
Ulceration	Painful	Painless
	Heels and toes	Plantar

are common due to neuropathic and infective causes rather than vascular causes.

Management

Many diabetic foot problems are avoidable, so patients need to learn the principles of foot care (Table 19.13). Older patients should visit a chiropodist regularly and should not cut their own toe-nails. Once tissue damage has occurred in the form of ulceration or gangrene, the aim is preservation of viable tissue. The four main threats to the skin and subcutaneous tissues are:

- *Infection*. This can take hold rapidly in a diabetic foot. Early antibiotic treatment is essential, with antibiotic therapy adjusted in the light of culture results. The organisms grown from the skin surface may not be the organism causing deeper infection. Collections of pus are drained and excision of infected bone is needed if osteomyelitis develops and does not respond to appropriate antibiotic therapy. Regular X-rays of the foot are needed to check on progress.
- *Ischaemia*. The blood flow to the feet is assessed clinically and with Doppler ultrasound. Femoral angiography is used to localize areas of occlusion amenable to bypass surgery or angioplasty. Relatively few patients fall into this category, and the risks (deterioration leading to amputation) and benefits of surgical intervention are finely balanced.
- *Abnormal pressure*. An ulcerated site must be kept non-weight-bearing. Resting the affected leg may need to be supplemented with special deep shoes and insoles to move pressure away from critical sites, or by removable or non-removable casts of the leg. After healing, special shoes and insoles are likely to continue to be needed to protect the feet and prevent abnormal pressure repeating damage to a healed area. In neuropathic feet particularly, sharp surgical debridement by a chiropodist is necessary to prevent callus distorting the local wound architecture and causing damage through abnormal pressure on normal skin nearby.
- *Wound environment*. Dressings are used to absorb or remove exudate, maintain moisture, and protect the wound from contaminating agents, and should be easily removable. Expensive new dressings containing growth factors and other biologically active agents may have a role to play in future, but their place is still being assessed.

Good liaison between physician, chiropodist and surgeon is essential if periods in hospital are to be used efficiently. When irreversible arterial insufficiency is present, it is often quicker and kinder to opt for an early major amputation rather than subject the patient to a debilitating sequence of conservative procedures.

Table 19.13 Principles of diabetic foot care

Inspect feet daily
Seek early advice for any damage
Check shoes inside and out for sharp bodies/areas before wearing
Use lace-up shoes with plenty of room for the toes
Keep feet away from sources of heat (hot sand, hot-water bottles, radiators, fires)
Check the bath temperature before stepping in

Infections

There is no evidence that diabetic patients with good glycaemic control are more prone to infection than normal subjects. However, poorly controlled diabetes entails increased susceptibility to the following infections:

- Skin
  - (a) staphylococcal infections (boils, abscesses, carbuncles)
  - (b) mucocutaneous candidiasis
- Gastrointestinal tract
  - (a) chronic periodontitis
  - (b) rectal and ischiorectal abscess formation (when control very poor)
- Urinary tract
  - (a) urinary tract infections (in women)
  - (b) pyelonephritis
  - (c) perinephric abscess
- Lungs
  - (a) staphylococcal and pneumococcal pneumonia
  - (b) Gram-negative bacterial pneumonia
  - (c) tuberculosis.

One reason why poor control leads to infection is that chemotaxis and phagocytosis by polymorphonuclear leucocytes are impaired because at high blood glucose concentrations neutrophil superoxide generation is impaired.

Conversely, infections may lead to loss of glycaemic control, and are a common cause of ketoacidosis. Insulin-treated patients need to increase their dose by up to 25% in the face of infection, and non-insulin-treated patients may need insulin cover while the infection lasts. Patients should be told never to omit their insulin dose, even if they are nauseated and unable to eat; instead they should test their blood glucose frequently and seek urgent medical advice. Diabetic patients should receive pneumococcal vaccine and yearly influenza vaccine.

Skin and joints (see also p. 1254)

Joint contractures in the hands are a common consequence of childhood diabetes. The sign may be demonstrated by asking the patient to join the hands as if in prayer; the metacarpophalangeal and interphalangeal joints cannot be opposed. Thickened, waxy skin can be noted on the backs of the fingers. These features may be due to glycosylation of collagen and are not progressive. The condition is sometimes referred to as diabetic cheiroarthropathy.

Osteopenia in the extremities is also described in type 1 diabetes but rarely leads to clinical consequences.

FURTHER READING

Boulton AJM, Kirsener RS, Vileikyte L. Neuropathic diabetic foot ulcers. *New England Journal of Medicine* 2004; **351**: 48–55.

Brownlee M. Biochemistry and molecular cell biology of diabetic complications. *Nature* 2001; **44**: 813–820.

Camilleri M. Diabetic gastroparesis. *New England Journal of Medicine* 2007; **356**: 820–829.

Colhoun HM, Betteridge PJ, Burrington PM et al. Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the collaborative atorvastatin diabetes study (CARDS). *Lancet* 2004; **364**: 685–696.

Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *New England Journal of Medicine* 1993; **329**: 977–986.

Frank RN. Diabetic retinopathy. *New England Journal of Medicine* 2004; **350**: 48–58.

McVary KT. Erectile dsyfunction. *New England Journal of Medicine* 2007; **357**: 2473–2481.

Mitch WE. Treating diabetic neuropathy. *New England Journal of Medicine* 2004; **351**: 1934–1936.

NOTES ON SPECIAL SITUATIONS IN DIABETES

Surgery

Smooth control of diabetes minimizes the risk of infection and balances the catabolic response to anaesthesia and surgery. The procedure for insulin-treated patients is simple:

- Long-acting and/or intermediate insulin should be stopped the day before surgery, with soluble insulin substituted.
- Whenever possible, diabetic patients should be first on the morning theatre list.
- An infusion of glucose, insulin and potassium is given during surgery. The insulin can be mixed into the glucose solution or administered separately by syringe pump. A standard combination is 16 U of soluble insulin with 10 mmol of KCl in 500 mL of 10% glucose, infused at 100 mL/h.
- Postoperatively, the infusion is maintained until the patient is able to eat. Other fluids needed in the perioperative period must be given through a separate intravenous line and must not interrupt the glucose/insulin/potassium infusion. Glucose levels are checked every 2–4 hours and potassium levels are monitored. The amount of insulin and potassium in each infusion bag is adjusted either upwards or downwards according to the results of regular monitoring of the blood glucose and serum potassium concentrations.

The same approach is used in the emergency situation, with the exception that a separate variable-rate insulin infusion may be needed to bring blood glucose under control before surgery.

Non-insulin-treated patients should stop medication 2 days before the operation. Patients with mild hyperglycaemia (fasting blood glucose below 8 mmol/L) can be treated as non-diabetic. Those with higher levels are treated with soluble insulin prior to surgery, and with glucose, insulin and potassium during and after the procedure, as for insulin-treated patients.

Pregnancy and diabetes

Meticulous metabolic control of the diabetes and careful medical and obstetric management is required.

Metabolic control of diabetes in pregnancy

The patient should perform daily home blood glucose profiles, recording blood tests before and 2 hours after meals. The renal threshold for glucose falls in pregnancy, and urine tests are therefore of little or no value. Insulin requirements rise progressively, and intensified insulin regimens are generally used. The aim is to maintain blood glucose and fructosamine (or HbA_{1c}) levels as close to the normal range as can be tolerated. Oral anti-diabetic therapy should not be used.

General management

The patient is seen at intervals of 2 weeks or less at a clinic managed jointly by physician and obstetrician. Circumstances permitting, the aim should be outpatient management with a spontaneous vaginal delivery at term. Retinopathy and nephropathy may deteriorate during pregnancy. Digital photographic eye screening and urine testing for protein should be undertaken at booking, at 28 weeks and before delivery.

Obstetric problems associated with diabetes

Poorly controlled diabetes is associated with stillbirth, mechanical problems in the birth canal owing to fetal macrosomia, hydramnios and pre-eclampsia. Ketoacidosis in pregnancy carries a 50% fetal mortality, but maternal hypoglycaemia is relatively well tolerated.

Neonatal problems

Maternal diabetes, especially when poorly controlled, is associated with fetal macrosomia. The infant of a diabetic mother is more susceptible to hyaline membrane disease than non-diabetic infants of similar maturity. In addition, neonatal hypoglycaemia may occur. The mechanism is as follows: maternal glucose crosses the placenta, but insulin does not; the fetal islets hypersecrete to combat maternal hyperglycaemia, and a rebound to hypoglycaemic levels occurs when the umbilical cord is severed.

These complications are due to hyperglycaemia in the third trimester. Poor glycaemic control around the time of conception carries an increased risk of major congenital malformations. When a pregnancy is planned, optimal metabolic control should be sought before conception.

Gestational diabetes

This term refers to glucose intolerance that develops in the course of pregnancy and usually remits following delivery. The condition is typically asymptomatic. Women who have a previous history of gestational diabetes, older or overweight women, those with a history of large for gestational age babies and women from certain ethnic groups are at particular risk, but many cases occur in women who are not in any of these categories. For this reason some advocate screening of all pregnant women on the basis of random plasma glucose testing in each trimester and by oral glucose tolerance testing if the glucose concentration is, for example, 7 mmol/L or more. There is no consensus concerning the level of blood glucose which is harmful for the baby, and therefore no consensus concerning cut-off levels for screening and intervention.

Treatment is with diet in the first instance, but most patients require insulin cover during the pregnancy. Insulin does not cross the placenta. Many oral agents cross the placenta and are usually avoided because of the potential risk to the fetus.

Gestational diabetes has been associated with all the obstetric and neonatal problems described above for pre-

existing diabetes, except that there is no increase in the rate of congenital abnormalities. Gestational diabetes is often the harbinger of type 2 diabetes in later life.

Not all diabetes presenting in pregnancy is gestational. True type 1 diabetes may develop, and swift diagnosis is essential to prevent the development of ketoacidosis. Hospital admission is required if the patient is symptomatic, or has ketonuria or a markedly elevated blood glucose level.

Unstable diabetes

This term is used to describe patients with recurrent ketoacidosis and/or recurrent hypoglycaemic coma. Of these, the largest group is made up of those who experience recurrent severe hypoglycaemia.

Recurrent severe hypoglycaemia

This affects 1–3% of insulin-dependent patients. Most are adults who have had diabetes for more than 10 years. By this stage, endogenous insulin secretion is negligible in the great majority of patients. Pancreatic alpha-cells are still present in undiminished numbers, but the glucagon response to hypoglycaemia is virtually absent. Long-term patients are thus subject to fluctuating hyperinsulinaemia owing to erratic absorption of insulin from injection sites, and lack a major component of the hormonal defence against hypoglycaemia. In this situation adrenaline (epinephrine) secretion becomes vital, but this too may become impaired in the course of diabetes. Loss of adrenaline (epinephrine) secretion has been attributed to autonomic neuropathy, but this is unlikely to be the sole cause; central adaptation to recurrent hypoglycaemia may also be a factor.

The following factors may also predispose to recurrent hypoglycaemia:

- *Overtreatment with insulin*. Frequent biochemical hypoglycaemia lowers the glucose level at which symptoms develop. Symptoms often reappear when overall glucose control is relaxed.
- *An unrecognized low renal threshold for glucose*. Attempts to render the urine sugar-free will inevitably produce hypoglycaemia.
- *Excessive insulin doses*. A common error is to increase the dose when a patient needs more frequent injections to overcome a problem of timing.
- *Endocrine causes*. These include pituitary insufficiency, adrenal insufficiency and premenstrual insulin sensitivity.
- *Alimentary causes*. These include exocrine pancreatic failure and diabetic gastroparesis.
- *Chronic kidney disease*. Clearance of insulin is diminished.
- *Patient causes*. Patients may be unintelligent, uncooperative or may manipulate their therapy.

Recurrent ketoacidosis

This usually occurs in adolescents or young adults, particularly girls. Metabolic decompensation may develop very rapidly. A combination of chaotic food intake and insulin omission, whether conscious or unconscious, is now regarded as the primary cause of this problem. It almost always occurs in the context of considerable psychosocial problems, particularly eating disorders. This area needs careful and sympathetic exploration in any patient with recurrent ketoacidosis. It is perhaps not surprising that in an illness where much of one's life is spent thinking of and controlling food intake, 30% of women with diabetes have had some features of an eating disorder at some time. Other causes include:

- *Iatrogenic*. Inappropriate insulin combinations may be a cause of swinging glycaemic control. For example, a once-daily regimen may cause hypoglycaemia during the afternoon or evening and pre-breakfast hyperglycaemia due to insulin deficiency.
- *Intercurrent illness*. Unsuspected infections, including urinary tract infections and tuberculosis, may be present. Thyrotoxicosis can also manifest as unstable glycaemic control.

HYPOGLYCAEMIA IN THE NON-DIABETIC

Hypoglycaemia develops when hepatic glucose output falls below the rate of glucose uptake by peripheral tissues. Hepatic glucose output may be reduced by:

- the inhibition of hepatic glycogenolysis and gluconeogenesis by insulin
- depletion of hepatic glycogen reserves by malnutrition, fasting, exercise or advanced liver disease
- impaired gluconeogenesis (e.g. following alcohol ingestion).

In the first of these categories, insulin levels are raised, the liver contains adequate glycogen stores, and the hypoglycaemia can be reversed by injection of glucagon. In the other two situations, insulin levels are low and glucagon is ineffective. Peripheral glucose uptake is accelerated by high insulin levels and by exercise, but these conditions are normally balanced by increased glucose output.

The most common symptoms and signs of hypoglycaemia are neurological. The brain consumes about 50% of the total glucose produced by the liver. This high energy requirement is needed to generate ATP used to maintain the potential difference across axonal membranes.

Insulinomas

Insulinomas are pancreatic islet cell tumours that secrete insulin. Most are sporadic but some patients have multiple tumours arising from neural crest tissue (multiple endocrine neoplasia). Some 95% of these tumours are benign. The classic presentation is with fasting hypoglycaemia, but early symptoms may also develop in the late morning or afternoon. Recurrent hypoglycaemia is often present for months or years before the diagnosis is made, and the symptoms may be atypical or even bizarre; the presenting features in one series are given in Table 19.14. Common misdiagnoses include psychiatric disorders, particularly pseudodementia in

Table 19.14 Presenting features of insulinoma

Diplopia
Sweating, palpitations, weakness
Confusion or abnormal behaviour
Loss of consciousness
Grand mal seizures

FURTHER READING

Ecker JL, Greene MF. Gestational diabetes mellitus. *New England Journal of Medicine* 2008; **358**: 2061–2063.

elderly people, epilepsy and cerebrovascular disease. Whipple's triad remains the basis of clinical diagnosis. This is satisfied when:

- symptoms are associated with fasting or exercise
- hypoglycaemia is confirmed during these episodes
- glucose relieves the symptoms.

A fourth criterion – demonstration of inappropriately high insulin levels during hypoglycaemia – may usefully be added to these.

The diagnosis is confirmed by the demonstration of hypoglycaemia in association with inappropriate and excessive insulin secretion. Hypoglycaemia is demonstrated by:

- Measurement of overnight fasting (16 hours) glucose and insulin levels on three occasions. About 90% of patients with insulinomas will have low glucose and non-suppressed (normal or elevated) insulin levels.
- A prolonged 72-hour supervised fast if overnight testing is inconclusive and symptoms persist.

Autonomous insulin secretion is demonstrated by lack of the normal feedback suppression during hypoglycaemia. This may be shown by measuring insulin, C-peptide or proinsulin during a spontaneous episode of hypoglycaemia.

Treatment of insulinoma

The most effective therapy is surgical excision of the tumour, but insulinomas are often very small and difficult to localize. Many techniques can be used to attempt to localize insulinomas. Sensitivity and specificity vary between centres and between operators. These include highly selective angiography, contrast-enhanced high-resolution CT scanning, scanning with radio-labelled somatostatin (some insulinomas express somatostatin receptors), and endoscopic and intraoperative ultrasound scanning. Venous sampling for the detection of 'hot spots' of high insulin concentration in the various intra-abdominal veins is still used occasionally.

Medical treatment with diazoxide is useful when the insulinoma is malignant, in patients in whom a tumour cannot be located, and in elderly patients with mild symptoms. Symptoms may also remit on treatment with a somatostatin analogue (octreotide or lanreotide).

Hypoglycaemia with other tumours

Hypoglycaemia may develop in the course of advanced neoplasia and cachexia, and has been described in association with many tumour types. Certain massive tumours, especially sarcomas, may produce hypoglycaemia owing to the secretion of insulin-like growth factor-1. True ectopic insulin secretion is extremely rare.

Postprandial hypoglycaemia

If frequent venous blood glucose samples are taken following a prolonged glucose tolerance test, about one in four subjects will have at least one value below 3 mmol/L. The arteriovenous glucose difference is quite marked during this phase, so that very few are truly hypoglycaemic in terms of arterial (or capillary) blood glucose content. Failure to appreciate this simple fact led some authorities to believe that postprandial (or reactive) hypoglycaemia was a potential 'organic' explanation for a variety of complaints that might otherwise have been considered psychosomatic. An epidemic of false 'hypoglycaemia' followed, particularly in the USA. Later work showed a poor correlation between symptoms and biochemical hypoglycaemia. Even so, a number of otherwise normal people occasionally become pale, weak and sweaty at times when meals are due, and report benefit from advice to take regular snacks between meals.

True postprandial hypoglycaemia may develop in the presence of alcohol, which 'primes' the cells to produce an exaggerated insulin response to carbohydrate. The person who substitutes alcoholic beverages for lunch is particularly at risk. Postprandial hypoglycaemia sometimes occurs after gastric surgery, owing to rapid gastric emptying and mismatching of nutrient absorption and insulin secretion. This is referred to as 'dumping' but it is now rarely encountered (see p. 262).

Hepatic and renal causes of hypoglycaemia

The liver can maintain a normal glucose output despite extensive damage, and hepatic hypoglycaemia is uncommon. It is particularly a problem with fulminant hepatic failure.

The kidney has a subsidiary role in glucose production (via gluconeogenesis in the renal cortex), and hypoglycaemia is sometimes a problem in terminal renal failure.

Hereditary fructose intolerance occurs in 1 in 20 000 live births and can cause hypoglycaemia (see p. 1068).

Endocrine causes of hypoglycaemia

Deficiencies of hormones antagonistic to insulin are rare but well-recognized causes of hypoglycaemia. These include hypopituitarism, isolated adrenocorticotrophic hormone (ACTH) deficiency and Addison's disease.

Drug-induced hypoglycaemia

Many drugs have been reported to produce isolated cases of hypoglycaemia, but usually only when other predisposing factors are present:

- Sulfonylureas may be used in the treatment of diabetes or may be taken by non-diabetics in suicide attempts.
- Quinine may produce severe hypoglycaemia in the course of treatment for falciparum malaria.
- Salicylates may cause hypoglycaemia; usually accidental ingestion by children.
- Propranolol can induce hypoglycaemia in the presence of strenuous exercise or starvation.
- Pentamidine used in the treatment of resistant pneumocystis pneumonia (see p. 200).

Alcohol-induced hypoglycaemia

Alcohol inhibits gluconeogenesis. Alcohol-induced hypoglycaemia occurs in poorly nourished chronic alcoholics, binge drinkers and in children who have taken relatively small amounts of alcohol, since they have a diminished hepatic glycogen reserve. They present with coma and hypothermia.

Factitious hypoglycaemia

This is a relatively common variant of self-induced disease and is more common than an insulinoma. Hypoglycaemia is produced by surreptitious self-administration of insulin or sulfonylureas. Many patients in this category have been extensively investigated for an insulinoma. Measurement of C-peptide levels during hypoglycaemia should identify patients who are injecting insulin; sulphonylurea abuse can be detected by chromatography of plasma or urine.

DISORDERS OF LIPID METABOLISM

Lipid physiology

Lipids are insoluble in water, and are transported in the bloodstream as macromolecular complexes. In these complexes, lipids (principally triglyceride, cholesterol and cholesterol esters) are surrounded by a stabilizing coat of phospholipid. Proteins (called apoproteins) embedded into the surface of these 'lipoprotein' particles exert a stabilizing function and allow the particles to be recognized by receptors in the liver and the peripheral tissues. The structure of a chylomicron (one type of lipoprotein particle) is illustrated in Figure 19.17.

Five principal types of lipoprotein particles are found in the blood (Fig. 19.18). They are structurally different and can be separated in the laboratory by their density and electrophoretic mobility. The larger particles give postprandial plasma its cloudy appearance. More than half of all patients aged under 60 with angiographically confirmed coronary artery disease have a lipoprotein disorder.

The genes for all the major apoproteins and that for the low-density lipoprotein (LDL) receptor have been isolated, sequenced and their chromosomal sites mapped. Production of abnormal apoproteins is known to produce, or predispose to, several types of lipid disorder, and it is likely that others will be discovered.

Chylomicrons (Fig. 19.18)

Chylomicrons are synthesized in the small intestine postprandially, passing initially into the intestinal lymphatic drainage, then along the thoracic duct into the bloodstream. They contain triglyceride and a small amount of cholesterol and its ester, and provide the main mechanism for transporting the digestion products of dietary fat to the liver and peripheral tissues. Each newly formed chylomicron contains several different apoproteins (B-48, A-I, A-II), and acquires apoproteins C-II and E by transfer from high-density lipoprotein (HDL) particles in the bloodstream. Apoprotein C-II binds to specific receptors in adipose tissue and skeletal muscle and the liver and allows the endothelial enzyme, lipoprotein lipase, to remove most of the triglyceride from the particle. The remaining chylomicron remnant particle, which contains the bulk of the original cholesterol, is taken up by the liver by mechanisms which are not fully understood, possibly mediated by apoprotein E.

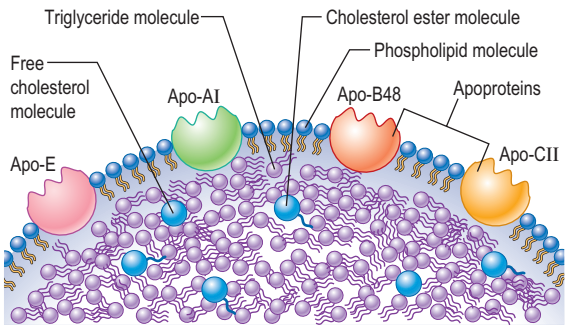


Fig. 19.17 **Schematic diagram of a chylomicron particle (75–1200 nm) showing apoproteins lying in the surface membrane.**

Very-low-density lipoprotein (VLDL) particles (Fig. 19.18)

These are synthesized continuously in the liver and contain most of the body's endogenously synthesized triglyceride and a smaller quantity of cholesterol. They are the body's main source of energy during prolonged fasting. Apoprotein B-100 is an essential component of VLDL. Apoproteins C-II and E are incorporated later into VLDL by transfer from HDL particles. As they pass round the circulation, VLDL particles bind through apoprotein C-II allowing triglyceride to be progressively removed by lipoprotein lipase in the capillary endothelium. This leaves a particle, now depleted of triglyceride and apoprotein C-II, called an intermediate-density lipoprotein (IDL) particle.

Intermediate-density lipoprotein (IDL) particles (Fig. 19.18)

These have apoprotein B-100 and apoprotein E molecules on the particle surface. Most IDL particles bind to liver LDL receptors through the apoprotein E molecule and are then

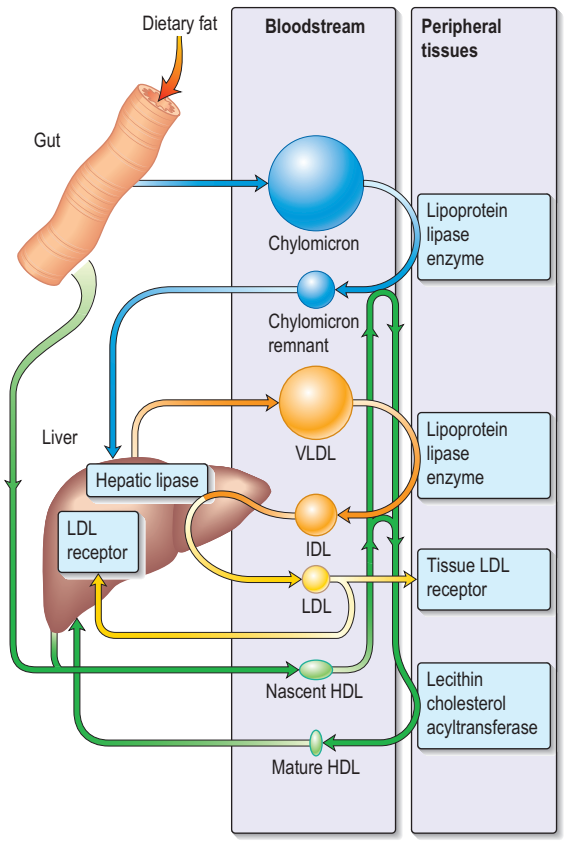


Fig. 19.18 **Schematic representation of the sites of origin, interaction between, and fate of, the major lipoprotein particles.**

catabolized. Some IDL particles have further triglyceride removed (by the enzyme hepatic lipase), producing LDL particles.

Low-density lipoprotein (LDL) particles

(Fig. 19.18)

LDL particles are the main carrier of cholesterol, and deliver it both to the liver and to peripheral cells. The surface of the LDL particle contains apoprotein B-100, and also apoprotein E. The apoprotein B-100 is the principal ligand for the LDL clearance receptor. This receptor lies within coated pits on the surface of the hepatocyte. Once bound to the receptor, the coated pit invaginates and fuses with liposomes which destroy the LDL particle (Fig. 19.19). The number of hepatic LDL clearance receptors regulates the circulating LDL concentration, which is also influenced by controlling the activity of the rate-limiting enzyme in the cholesterol synthetic pathway, hydroxymethylglutaryl coenzyme A (HMG-CoA) reductase (Fig. 7.4).

LDL particles can deposit lipid into the walls of the peripheral vasculature. Not all the cholesterol synthesized by the liver is packaged immediately into lipoprotein particles. Some is oxidized into bile salts. Both bile salts and cholesterol are excreted in the bile: both are then reabsorbed through the terminal ileum and recirculated (enterohepatic circulation).

LDL particles become Lp(a) lipoproteins as a result of the linkage of apoprotein (a) to aproprotein B-100 with a single disulphide bond. Raised levels of Lp(a) lipoprotein are a risk factor for cardiovascular disease.

High-density lipoprotein (HDL) particles

(Fig. 19.18)

Nascent HDL particles are produced in both the liver and intestine. They are disc shaped, seemingly inert and contain apoprotein A-I. They are transmuted into mature particles by the acquisition of phospholipids, and the E and C apoproteins from chylomicrons and VLDL particles in the circulation. The more mature HDL particles take up cholesterol from cells in the peripheral tissues aided by cholesterol-efflux regulatory protein – a product of the ATP-binding cassette transporter 1 gene (ABC1 gene). As it is taken up, the enzyme lecithin-cholesterol acyl-transferase (LCAT), activated by the apoprotein A on the particle's surface, esterifies the sequestered cholesterol. The HDL particle transports cholesterol away from the periphery and may transfer it indirectly to other particles such as VLDL in the circulation or deliver its cholesterol directly to the liver (reverse cholesterol transport) and steroid-synthetic tissues (ovaries, testes, adrenal cortex).

This direct delivery takes place through scavenger-receptor B1. In experimental animals the absence of scavenger-receptor B1 dramatically accelerates the development of atheroma, and genetically programmed overproduction suppresses atheroma formation.

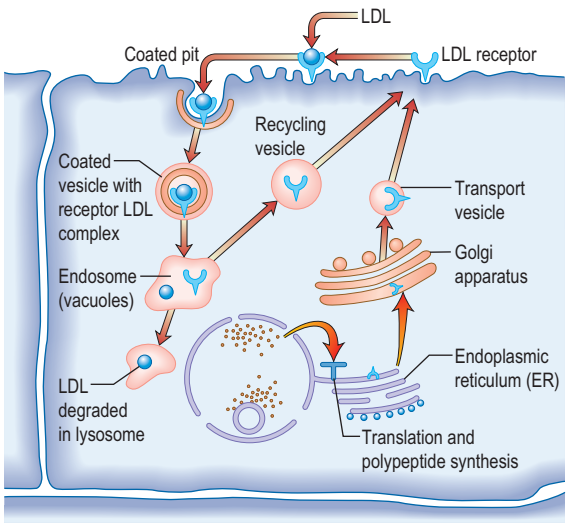


Fig. 19.19 **Receptor-mediated endocytosis.** LDL receptors are formed in the endoplasmic reticulum and transported via the Golgi apparatus to the surface of a hepatocyte. LDLs bind to these receptors, are internalized and taken up by the endosome. The receptor is recycled back to the surface, while the LDL is broken down by the lysosomes, freeing cholesterol needed for membrane synthesis.

Measurement

When a laboratory measures fasting serum lipids, the majority of the total cholesterol concentration consists of LDL particles with a 20–30% contribution from HDL particles. The triglyceride concentration largely reflects the circulating number of VLDL particles, since chylomicrons are not normally present in the fasted state. If the patient is not fasted, the total triglyceride concentration will be raised owing to the additional presence of triglyceride-rich chylomicrons.

Epidemiology and lipids

LDL and total cholesterol

Population studies have repeatedly demonstrated a strong association between both total and LDL cholesterol concentration and coronary heart risk. There is a strong link between mean fat consumption, mean serum cholesterol concentration and the prevalence of coronary heart disease between countries. The exception is France where the cardiovascular risk is only moderate – perhaps owing to high alcohol consumption. Studies of migrants, particularly of Japanese men migrating to Hawaii, have shown that as diet changes, and cholesterol concentrations rise, so does the cardiovascular risk. Such studies show the role of the environment rather than the genetic make-up of a population.

The Multiple Risk Factor Intervention Trial (MRFIT) screened one-third of a million American men for various cardiovascular risk factors and then followed them for 6 years. Data from this study have shown that although cardiovascular risk rises progressively as total cholesterol concentration increases (Fig. 19.20), the risk increase is modest for individuals with no other cardiovascular risk factors. With each additional risk factor the effect produced by the same difference in cholesterol concentration becomes greatly magnified. The Framingham Study has reproduced these findings in a separate population.

HDL cholesterol

Epidemiological studies have shown that higher HDL concentrations protect against cardiovascular disease. Raising HDL by pharmacological means does not, however, necessarily reduce cardiovascular risk. HDL also has effects on the function of platelets and of the haemostatic cascade. These properties may favourably influence thrombogenesis.

VLDL particles (triglycerides)

There is a relatively weak independent link between raised concentrations of (triglyceride-rich) VLDL particles and car-

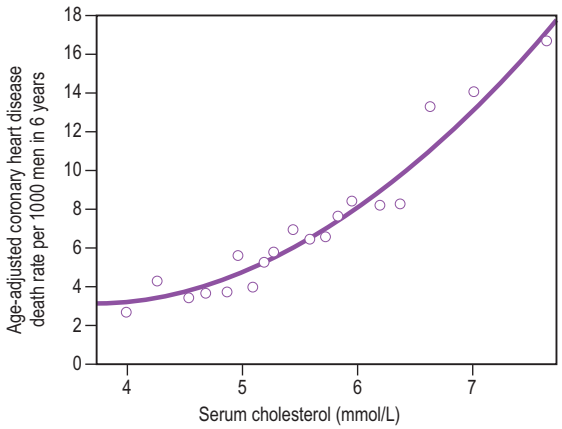


Fig. 19.20 **The Multiple Risk Factor Intervention Trial.** Relationship between levels of serum cholesterol and risks of fatal coronary artery disease in a longitudinal study of more than 361 000 men screened for entry into the trial. From Stamler J, Wentworth D, Neaton JD. Is relationship between serum cholesterol and risk of death from coronary heart disease continuous and graded? Findings in 356 222 primary screenees of the Multiple Risk Factor Intervention Trial (MRFIT). *Journal of the American Medical Association* 1986; **256**: 2823.

diovascular risk. Very raised triglyceride concentrations (>6 mmol/L) cause a greatly increased risk of acute pancreatitis and retinal vein thrombosis. Hypertriglyceridaemia tends to occur in association with a reduced HDL concentration. Much of the cardiovascular risk associated with 'hypertriglyceridaemia' turns out on multivariate analysis to be due to the associated low HDL levels and not to the hypertriglyceridaemia itself.

Chylomicrons

Excess chylomicrons do not confer an excess cardiovascular risk, but raise the total plasma triglyceride concentration.

HYPERLIPIDAEMIA

Hyperlipidaemia results from genetic predisposition interacting with an individual's diet.

Secondary hyperlipidaemia

If a lipid disorder has been detected it is vital to carry out a clinical history, examination and simple special investigations to detect causes of secondary hyperlipidaemia (Table 19.15), which may need treatment in their own right. Hypothyroidism, diabetes, renal disease and abnormal liver function can all raise plasma lipid levels.

Table 19.15 Causes of secondary hyperlipidaemia

Hypothyroidism
Diabetes mellitus (when poorly controlled)
Obesity
Renal impairment
Nephrotic syndrome
Dysglobulinaemia
Hepatic dysfunction
Drugs:
Oral contraceptives in susceptible individuals
Retinoids, thiazide diuretics, corticosteroids, op'DDD (used in the treatment of Cushing's syndrome), sirolimus (and other immunosuppressive agents)

CLASSIFICATION, CLINICAL FEATURES AND INVESTIGATION OF PRIMARY HYPERLIPIDAEMIAS

We have used the functional/genetic classification which has the advantage that the genetic disorders may be grouped by the results of simple lipid biochemistry into causes of:

- disorders of VLDL and chylomicrons – hypertriglyceridaemia alone
- disorders of LDL – hypercholesterolaemia alone
- disorders of HDL
- combined hyperlipidaemia.

Disorders of VLDL and chylomicrons – hypertriglyceridaemia alone

The majority of cases appear to be due to multiple genes acting together to produce a modest excess of circulating concentration of VLDL particles, such cases being termed polygenic hypertriglyceridaemia.

In a proportion of cases there will be a family history of a lipid disorder or its effects (e.g. pancreatitis). Such cases are often classified as familial hypertriglyceridaemia. The defect underlying the vast majority of such cases is not understood. The main clinical feature is a history of attacks of pancreatitis or retinal vein thrombosis in some individuals.

Lipoprotein lipase deficiency and apoprotein C-II deficiency

These are rare diseases which produce greatly elevated triglyceride concentrations owing to the persistence of chylomicrons (and not VLDL particles) in the circulation. The chylomicrons persist because the triglyceride within cannot be metabolized if the enzyme lipoprotein lipase is defective, or because the triglycerides cannot gain access to the normal enzyme owing to deficiency of the apoprotein C-II on their surface. Patients present in childhood with eruptive xanthomas, lipaemia retinalis and retinal vein thrombosis, pancreatitis and hepatosplenomegaly. If not identified in childhood, it can present in adults with gross hypertriglyceridaemia resistant to simple measures. The presence of chylomicrons floating like cream on top of fasting plasma suggests this diagnosis. It is confirmed by plasma electrophoresis or ultracentrifugation. An abnormality of apoprotein C can be deduced if the hypertriglyceridaemia improves temporarily after infusing fresh frozen plasma, and lipoprotein lipase deficiency is likely if it does not.

Disorders of LDL – hypercholesterolaemia alone

Heterozygous familial hypercholesterolaemia is an autosomal dominant monogenic disorder present in 1 in 500 of the normal population. The average primary care physician would therefore be expected to have four such patients, but because of clustering within families the prevalence varies.

There is an increased prevalence in some racial groups (e.g. French Canadians, Finns, South Africans). Surprisingly, most individuals with this disorder remain undetected. Patients may have no physical signs, in which case the diagnosis is made on the presence of very high plasma cholesterol concentrations which are unresponsive to dietary modification and are associated with a typical family history of early cardiovascular disease. Diagnosis can more easily be made if typical clinical features are present. These include xanthomatous thickening of the Achilles tendons and xanthomas over the extensor tendons of the fingers. Xanthelasma may be present, but is not diagnostic of familial hypercholesterolaemia.

The genetic defect is the underproduction or malproduction of the LDL cholesterol uptake receptor in the liver (Table 19.16). Over 150 different mutations in the LDL receptor have been described to date. Fifty per cent of men with the disease will die by the age of 60, most from coronary artery disease, if untreated.

Homozygous familial hypercholesterolaemia is very rare indeed. Affected children have no LDL receptors in the liver. They have a hugely elevated LDL cholesterol concentration, and massive deposition of lipid in arterial walls, the aorta and the skin. The natural history is for death from ischaemic heart disease in late childhood or adolescence. Repeated plasmapheresis has been used to remove LDL cholesterol with some success. Liver transplantation is a 'cure'. Plasma lipids normalize and xanthomata regress after transplantation, but the numbers of patients having undergone this procedure is small.

Mutations in the apoprotein B-100 gene cause another relatively common single gene disorder. Since LDL particles bind to their clearance receptor in the liver through apoprotein B-100, this defect also results in high LDL concentrations in the blood, and a clinical picture which closely resembles classical heterozygous familial hypercholesterolaemia. The two disorders can be distinguished clearly only by genetic tests. The approach to treatment is the same.

Polygenic hypercholesterolaemia is a term used to lump together patients with raised serum cholesterol concentrations, but without one of the monogenic disorders above. They exist in the right-hand tail of the normal distribution of cholesterol concentration. The precise nature of the polygenic variation in plasma cholesterol concentration remains unknown. Variations in the apoprotein E gene (chromosome 19), and in Sterol-regulatory element-binding protein (SREBP)-2 gene, are involved in some individuals in this heterogeneous group.

Disorders of HDL (very low HDL, low total cholesterol)

Tangier disease is an autosomal recessive disorder characterized by a low HDL cholesterol concentration. Cholesterol accumulates in reticuloendothelial tissue and arteries causing enlarged orange-coloured tonsils and hepatosplenomegaly. Cardiovascular disease, corneal opacities and a polyneuropathy also occur. It is due to a gene mutation (ABC1 gene – see HDL physiology above and Table 19.16) which normally promotes cholesterol uptake from cells by HDL particles.

Other mutations in this gene have been found in a few families with autosomal dominant HDL deficiency. It is as yet unknown whether abnormalities of this gene contribute to the low HDL cholesterol concentrations commonly seen in cardiovascular disease patients.

Combined hyperlipidaemia (hypercholesterolaemia and hypertriglyceridaemia)

The most common patient group is a polygenic combined hyperlipidaemia. Patients have an increased cardiovascular risk due to both high LDL concentrations and suppression of HDL by the hypertriglyceridaemia.

Familial combined hyperlipidaemia

This is relatively common, affecting 1 in 200 of the general population. The genetic basis for the disorder has not yet been characterized. It is diagnosed by finding raised cholesterol and triglyceride concentrations in association with a typical family history. There are no typical physical signs.

Remnant hyperlipidaemia

This is a rare (1 in 5000) cause of combined hyperlipidaemia. It is due to accumulation of LDL remnant particles and is associated with an extremely high risk of cardiovascular disease. It may be suspected in a patient with raised total cholesterol and triglyceride concentrations by finding xan-

Table 19.16 The genetic defects underlying some lipoprotein disorders

Disorder	Affected gene	Chromosome	Frequency
Common disorders			
Heterozygous familial hypercholesterolaemia	LDL receptor	19	1:500
Familial defective apoprotein B	Apo B-100	2	1:700
Hypobetalipoproteinaemia	Apo B-100	2	1:1000
Familial combined hyperlipidaemia	As yet unknown	As yet unknown	1:200
Familial hypertriglyceridaemia	As yet unknown	As yet unknown	1:200
Some rarer disorders			
Homozygous familial hypercholesterolaemia	LDL receptor	19	1:1000000
Lipoprotein lipase deficiency	As yet unknown	8	1:1000000 (homozygous)
Apoprotein C-II deficiency	Apo C-II	19	40 cases
Tangier disease	ATP-binding cassette	9	Very rare

Fig. 19.21 **Tuberous xanthomata behind the elbow, and lipid deposits in the hand creases in a patient with remnant hyperlipidaemia.**

thomas in the palmar creases (diagnostic) and the presence of tuberous xanthomas, typically over the knees and elbows (Fig. 19.21). Remnant hyperlipidaemia is almost always due to the inheritance of a variant of the apoprotein E allele (apoprotein E2) together with an aggravating factor such as another primary hyperlipidaemia. When suspected clinically the diagnosis can be confirmed using ultracentrifugation of plasma, or phenotyping apoprotein E.

THERAPIES AVAILABLE TO TREAT HYPERLIPIDAEMIA

The lipid-lowering diet

Studies have shown that dieticians helping patients to adjust their own diet to meet the nutritional targets set out below produce a better lipid-lowering effect than does the issuing of standard diet sheets and advice from a doctor. The main elements of a lipid-lowering diet are set out below.

Reduce the total fat intake

Dairy products and meat are the principal sources of saturated fat in the diet. Intake of these products should therefore be reduced, and fish and poultry should be substituted. Visible fat and skin should be removed before cooking and preparing meat dishes. Meat products including sausages and reconstituted meats (such as 'luncheon meat') should be avoided since the concentration of fat is unknown and often high. Baking and grilling of meats reduces the fat content and is preferred to frying. Low-fat or cottage cheese and skimmed or semi-skimmed milk should be substituted for the standard full-fat varieties. Pastries and cakes contain large quantities of fat and should be avoided. The overall aim should be to decrease fat intake such that it is providing approximately 30% of the total energy intake in the diet. Further reduction in fat intake is unacceptable to many patients.

Substitution with monounsaturates and polyunsaturates

Monounsaturated oils, particularly olive oil, and polyunsaturated oils such as sunflower, safflower, corn and soya oil, should be used in cooking instead of saturated fat-rich alternatives.

Reduce the dietary cholesterol intake

Liver, offal and fish roes should be avoided. Although eggs and prawns are rich in cholesterol their total contribution to the body's cholesterol pool is small and they can still be part of a balanced lipid-lowering diet.

Increase the intake of fibre (non-starch polysaccharides, NSPs)

Foods high in soluble fibre, such as pulses, legumes, root vegetables, leafy vegetables and unprocessed cereals, help reduce circulating lipid concentrations. These should be substituted in the diet in the place of higher-fat alternatives.

Reduce alcohol consumption

Excess alcohol is a cause of secondary hyperlipidaemia, and may worsen primary lipid disorders.

Achieve an ideal bodyweight

Treatment of obesity is part of the management of hyperlipidaemia, both because it will help the lipid disorder itself, and because obesity is an independent cardiovascular risk factor.

Foods containing stanol esters

Plant stanols reduce the absorption of cholesterol from the intestine by competing for space in the micelles that deliver lipid to the mucosal cells of the gut. They are largely unabsorbed and excreted in the stool. Increasing the amount of plant stanol in the diet 10-fold by using a margarine (e.g. Benecol) containing added stanol esters lowers LDL cholesterol by approximately 0.35–0.5 mmol/L. A reduction in the risk of heart disease of about 25% would be expected if this reduction in LDL cholesterol was applied to a population.

Drugs

The classes of drugs used to treat hyperlipidaemia are listed in Table 19.17. The statins are the most widely used drugs. Anacetrapib, an inhibitor of cholesteryl ester transfer protein, has recently been introduced. It lowers LDL-C levels similar to statins with increase in HDL-C. Torcetrapib also lowers LDL-C and raises HDL-C but does not reduce atherogenesis, as judged by vascular intraluminal ultrasound.

MANAGEMENT OF SPECIFIC HYPERLIPIDAEMIAS

Screening

Most patients with hyperlipidaemia are asymptomatic and have no clinical signs. Many are discovered during the screening of high-risk individuals. Whose lipids should be measured?

There are great doubts as to whether blanket screening of plasma lipids is warranted. Selective screening of people at high risk of cardiovascular disease should be undertaken:

- a family history of coronary heart disease (especially below 50 years of age)
- a family history of lipid disorders
- the presence of a xanthoma
- the presence of xanthelasma or corneal arcus before the age of 40 years

Table 19.17 Drugs used in the management of hyperlipidaemia

Drug	Mechanism of action	Contraindications and adverse reactions	Expected therapeutic effect	Long-term safety
Statins e.g. Simvastatin Pravastatin Fluvastatin Atorvastatin Rosuvastatin	Inhibit the rate-limiting step in cholesterol synthesis (HMG 6A reductase)	***Contraindications***: Active liver disease, pregnancy, lactation ***Adverse effects:*** Derangement of liver biochemistry, diarrhoea, myositis. Raise ciclosporin level in blood	30–60% reduction in LDL cholesterol Modest triglyceride lowering Tiny effect on HDL cholesterol Atorvastatin and particularly rosuvastatin have the most potent cholesterol-lowering effects (HMG CoA reductase)	Simvastatin, Atorvastatin and Pravastatin have good long-term safety in large-scale trials and in clinical practice. Avoid if possible in women of childbearing age
Cholesterol absorption inhibitors e.g. Ezetimibe	Inhibition of gut absorption of cholesterol from food and also from bile. Mechanism of this inhibition is unclear	***Contraindications***: Lactation ***Adverse effects***: Occasional diarrhoea, abdominal discomfort	Reduce LDL cholesterol by additional 10–15% if given with a statin. Triglyceride concentrations reduced by 10%. Increase HDL cholesterol by 5%	Mostly act in gut and little is absorbed Short term safety good. Long term safety unknown
Cholesterol-binding resins e.g. Colestyramine Colestipol Colesevalam	Bind bile acids in the gut preventing enterohepatic circulation. Liver makes more bile acids from cholesterol, depleting the cholesterol pool	***Adverse effects***: Gastrointestinal side-effects predominate. Palatability is a problem ***Counselling***: Other drugs bind to resins and should be taken 1 h before or 4 h afterwards	8–15% reduction in LDL Little or no effect on HDL cholesterol 5–15% rise in triglyceride concentration	Not systemically absorbed. Safety profile is good. Appear safe in women of childbearing age. Fat-soluble vitamin supplements may be required in children, pregnancy and breast-feeding
Fibric acid derivatives e.g. Gemfibrozil Bezafibrate Ciprofibrate Fenofibrate	Activate peroxisome proliferator-activated nuclear receptors (esp. PPAR-alpha). This has protean effects on lipid metabolism	***Contraindications***: Severe hepatic or renal impairment, gallbladder disease, pregnancy ***Adverse effects***: Reversible myositis, nausea, predispose to gallstones, non-specific malaise, impotence	Reduction of LDL cholesterol by 10–15% and triglycerides by 25–35% HDL cholesterol concentrations increase by 10–50% (newer agents often have greater beneficial effect on HDL)	No knowledge of effect on developing fetus. Avoid in women of child-bearing age. Long-term safety appears good
Nicotinic acid and derivatives e.g. Nicotinic acid Modified-release nicotinic acid Acipimox	Unclear. Probably inhibit lipid synthesis in the liver by reducing free fatty acid concentrations through an inhibitory effect on lipolysis in fat tissue	***Contraindications***: Pregnancy, breast-feeding ***Adverse effects:*** Value limited by frequent side-effects: headache, flushing, dizziness, nausea, malaise, itching, abnormal liver biochemistry. Glucose intolerance, hyperuricaemia, dyspepsia, hyperpigmentation may occur	Reduce LDL cholesterol by 5–10% Reduce triglycerides by 15–20% HDL cholesterol increased by 10–20%	Medium-term safety but marred by the adverse effects listed. Modified-release preparation reduces side-effect incidence
Fatty acid compounds Omega-3 acid ethyl esters Omega-3 acid marine triglycerides	Reduce hepatic VLDL secretion	Occasional nausea and belching	Reduce triglycerides in severe hypertriglyceridaemia No favourable change in other lipids, and may aggravate hypercholesterolaemia in a few patients	Long-term safety is not yet known but seems unlikely to be poor
Cholesteryl ester transfer protein inhibitor e.g. Anacetrapib Torcetrapib	Under development	Unknown	LDL-C similar to statins and increase in HDL-C	Unknown

www.studentconsult.com
Ask the authors

- obesity
- diabetes mellitus
- hypertension
- acute pancreatitis
- those undergoing renal replacement therapy.

Where one family member is known to have a monogenic disorder such as familial hypercholesterolaemia (1 in 500 of the population), siblings and children must have their plasma lipid concentrations measured. It is also worth screening the prospective partners of any patients with this heterozygous monogenic lipid disorder because of the small risk of producing children homozygous for the condition.

Acute severe illnesses such as myocardial infarction can derange plasma lipid concentrations for up to 3 months. Plasma lipid concentrations should be measured either within 48 hours of an acute myocardial infarction (before derangement has had time to occur) or 3 months later.

Serum cholesterol concentration does not change significantly after a meal and as a screening test a random blood sample is sufficient. If the total cholesterol concentration is raised, HDL cholesterol, triglyceride and LDL cholesterol concentrations should be quantitated on a fasting sample. If a test for hypertriglyceridaemia is needed, a fasting blood sample is mandatory.

Hypertriglyceridaemia

A serum triglyceride concentration below 2.0 mmol/L is normal. In the range 2.0–6.0 mmol/L no specific intervention will be needed unless there are coincident cardiovascular risk factors, and in particular a strong family history of early cardiovascular death. In general, patients should be advised that they have a minor lipid problem, offered advice on weight reduction if obese, and advice on correcting other cardiovascular risk factors.

If the triglyceride concentration is above 6.0 mmol/L there is a risk of pancreatitis and retinal vein thrombosis. Patients should be advised to reduce their weight if overweight and start a formal lipid-lowering diet (see below). A proportion of individuals with hypertriglyceridaemia have livers which respond to even moderate degrees of alcohol intake by allowing accumulation or excess production of VLDL particles. If hypertriglyceridaemia persists, lipid measurements should be repeated before and after a 3-week interval of complete abstinence from alcohol. If a considerable improvement results, lifelong abstinence may prove necessary. Other drugs, including thiazides, oestrogens and glucocorticoids, can have a similar effect to alcohol in susceptible patients.

If the triglyceride concentration remains elevated above 6.0 mmol/L, despite the above measures, drug therapy is warranted. A fibric acid derivative is the agent of first choice. Nicotinic acid may be used in addition but its side-effects are often a problem. Fish oil capsules which contain ω-3 long-chain fatty acids are also effective in lowering triglyceride concentrations.

The severe hypertriglyceridaemia associated with the rare disorders of lipoprotein lipase deficiency and apoprotein C-II deficiency may require restriction of dietary fat to 10–20% of total energy intake and the use of special preparations of medium-chain triglycerides in cooking in place of oil or fat. Medium-chain triglycerides are not absorbed via chylomicrons (see p. 273).

Hypercholesterolaemia (without hypertriglyceridaemia)

Familial hypercholesterolaemia

Individuals often require treatment with diet and more than one cholesterol-lowering drug. The cholesterol absorption inhibitor ezetimibe is a logical addition to a statin and has a low side-effect profile (Table 19.17). Bile acid sequestrants are an alternative to ezetimibe, but have greater problems with tolerability. Concurrent therapy with statins and fibrates is usually avoided, in view of their overlapping side-effects, but in very severe cases such mixed therapy has been undertaken under close supervision.

Primary prevention for people with risk factors

Lipid-lowering therapy using a statin, or alternatives as above, should be considered in asymptomatic individuals irrespective of the total or LDL cholesterol level in type 2 diabetes alone, or with two or more of: positive family history of cardiovascular disease, albuminuria, hypertension, smoking.

Primary prevention for people without risk factors

In the absence of risk factors, lipid-lowering therapy should be used in asymptomatic men with LDL cholesterol levels persistently above 6.5 mmol/L despite dietary change. The situation for women is less clear.

Secondary prevention

As a generality, statin treatment is warranted for any patient with known macrovascular disease (coronary artery disease, TIA or stroke, peripheral vascular disease), irrespective of the total or LDL cholesterol level (treatment target is total cholesterol under 4.0, LDL-CK 1.8 mmol/L). If a statin is not tolerated, a fibrate, ezetimibe or fibrate/cholesterol absorption inhibitor combination are alternatives.

Risk prevention tables (see p. 748)

An array of risk prediction tables (see p. 748) are available to allow quantification of the risk of a patient having a cardiovascular event within the next 10 years. Some advocate the use of these risk tables as a guide to treatment – for example if the 10-year risk reaches the 15% or the 30% level. We side with those who have reservations over their use. Such risk analyses are a useful approach in helping to decide whether to use treatments such as aspirin. Aspirin probably has no effect until the day an atherosclerotic plaque ruptures, when it may then prevent thrombosis leading to a heart attack or stroke. Furthermore it has a significant associated morbidity and mortality (from bleeding). At a 10-year cardiovascular risk level of 15% the benefit/risk ratio for aspirin becomes favourable. At the 30% level the benefits are clear.

By contrast, the use of lipid-lowering agents, if initially tolerated, has a low associated morbidity and mortality. These agents probably reduce the rate of atheroma accumulation over a period of decades. When a patient is young, the chance that atheroma will be bad enough to cause a heart attack or stroke within the next 10 years will be small, even if atheroma is accumulating at a swift rate. The level of cardiovascular risk will only rise to the 15% or 30% level when the patient gets older. Yet it seems bizarre not to treat the gradual accumulation of atheroma when the patient is young

with a low 10-year risk, and then to start treatment when age causes the 10-year risk levels to rise to a particular threshold, if all other factors are the same. In choosing whether or not to prescribe, we prefer to consider 'will he/she live long enough to collect some pension and see his/her grandchildren?' rather than 'will he/she have a heart attack or stroke within the next 10 years?'. The answers to these two questions are very different.

Combined hyperlipidaemia (hypercholesterolaemia and hypertriglyceridaemia)

Treatment is the same for all varieties of combined hyperlipidaemia. For any given cholesterol concentration the hypertriglyceridaemia found in the combined hyperlipidaemias increases the cardiovascular risk considerably. Treatment is aimed at reducing serum cholesterol below 4.0 mmol/L and triglycerides below 2.0 mmol/L. Therapy is with diet in the first instance and with drugs if an adequate response has not occurred. Fibrates are the treatment of choice since these reduce both cholesterol and triglyceride concentrations, and also have the benefit of raising cardioprotective HDL concentrations. The combination of a fibrate and ezetimibe or a bile acid-binding resin is of considerable use when a fibrate alone produces an insufficient reduction in LDL cholesterol. Nicotinic acid can be used in addition, although its unwanted effects render it a third-line agent.

OTHER LIPID DISORDERS

Hypolipidaemia

Low lipid levels can be found in severe protein–energy malnutrition. They are also seen occasionally with severe malabsorption and in intestinal lymphangiectasia.

Hypobetalipoproteinaemia (Table 19.16) is a benign familial condition which is being increasingly recognized. The cholesterol levels are in the range 1–3.5 mmol/L.

Abetalipoproteinaemia

This is described on page 284.

FURTHER READING

Brewer HB. Increasing HDL cholesterol levels. *New England Journal of Medicine* 2004; **350**: 1491–1560.

Ford I, Murray H, Packard CJ et al. Long-term follow-up of the West of Scotland Coronary Prevention Study. *New England Journal of Medicine* 2007; **357**: 1477–1486.

Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20 536 high risk individuals: a randomized placebo-controlled trial. *Lancet* 2004; **363**: 757–767.

Smith CR. Lipid lowering therapy – new and established agents reduce risk of cardiovascular events. *Postgraduate Medicine* 2004; **115**: 29–30.

Walsh JME, Pignone M. Drug treatment of hyperlipidaemia in women. *Journal of the American Medical Association* 2004; **290**: 2243–2252.

INBORN ERRORS OF CARBOHYDRATE METABOLISM

Glycogen storage disease

All mammalian cells can manufacture glycogen, but the main sites of its production are the liver and muscle. Glycogen is a high-molecular-weight glucose polymer made up of 1–4 linked glucose units, with a 1–6 branch point every 4–10 residues. In glycogen storage disease there is either an abnormality in the molecular structure or an increase in glycogen concentration owing to a specific enzyme defect. Almost all these conditions are autosomal recessive in inheritance and present in infancy, except for McArdle's disease, which presents in adults.

Table 19.18 shows the classification and clinical features of some of these diseases.

Galactosaemia

Galactose is normally converted to glucose. However, a deficiency of the enzyme *galactose-1-phosphate uridyl-transferase* or, less commonly, *uridine diphosphate galactase-4-epimerase*, results in accumulation of galactose-1-phosphate in the blood. The transferase deficiency, inherited as an autosomal recessive, is due in 70% of patients to a glutamine to arginine missense mutation in Q188R. Galactose ingestion (i.e. milk) leads to inanition, failure to thrive, vomiting, hepatomegaly and jaundice, diabetes, cataracts and developmental delay. A lactose-free diet stops the acute toxicity but poor growth and problems with speech and mental development still occur with the transferase deficiency. A newborn screening programme to detect galactosaemia is used in parts of the USA and other countries.

Prenatal diagnosis and diagnosis of the carrier state are possible by measurement of the level of galactose-1-phosphate in the blood.

Galactokinase deficiency also results in galactosaemia and early cataract formation.

Defects of fructose metabolism

Absorbed fructose is chiefly metabolized in the liver to lactic acid or glucose. Three defects of metabolism in the liver and intestine occur; all are inherited as autosomal recessive traits:

- *Fructosuria* is due to fructokinase deficiency. It is a benign asymptomatic condition.
- *Hereditary fructose intolerance* is due to fructose-1-phosphate aldolase deficiency. Fructose-1-phosphate accumulates after fructose ingestion, inhibiting both glycogenolysis and gluconeogenesis, resulting in symptoms of severe hypoglycaemia. Hepatomegaly and renal tubular defects occur but are reversible on a fructose- and sucrose-free diet. Intelligence is normal and there is an absence of dental caries.
- *Hereditary fructose-1,6-diphosphatase deficiency* leads to a failure of gluconeogenesis. Infants present with hypoglycaemia, ketosis and lactic acidosis. Dietary control can lead to normal growth.

Table 19.18 Some glycogen storage diseases

Type	Affected tissue	Enzyme defect	Clinical features	Tissue needed for diagnosis*	Outcome
Liver glycogenoses					
I (Von Gierke) (25%)	Liver, intestine, kidney	Glycogen synthase deficiency Glucose-6-phosphatase	Hepatomegaly, ketotic hypoglycaemia, short stature, obesity, hypotonia	Liver DNA testing	If patients survive initial hypoglycaemia, prognosis is good; hyperuricaemia is a late complication
III (Forbes) (24%)	Liver, muscle (abnormal glycogen structure)	Glycogen debranching enzyme	Like type I	Leucocytes, liver, muscle	Good prognosis but progressive neuropathy and cardiomyopathy
IV (Anderson) (3%)	Liver (abnormal glycogen structure)	Branching enzyme	Failure to thrive, hepatomegaly, cirrhosis and its complications	Leucocytes, liver, muscle	Death in first 5 years Liver transplantation
VI (Hers) (VI and VIII = 30%)	Liver	Liver phosphorylase or phosphorylase kinase	Hepatomegaly with hypoglycaemia in childhood	Liver	Good
VIII	Liver	Phosphorylase b kinase deficiency	Hepatomegaly, hypoglycaemic fatiguability	Liver, muscle	No treatment
Muscle glycogenoses					
II (Pompé) (15%)	Liver, muscle, heart	Lysosomal acid, α-glucosidase	Respiratory muscle hypotonia, heart failure, cardiomyopathy	Fibroblasts, muscles	Alglucocidase alfa treatment now available Juvenile and adult variants seen
IIb (Dagon)	Muscle	Lysosome-associated maltase	Cardiomyopathy	Membrane protein 2	Muscle hypotonia
V (McArdle)	Muscle only	Phosphorylase	Muscle cramps and myoglobinuria after exercise (in adults)	Muscle	Normal life-span; give sucrose prior to exercise
VII (Tarui)	Muscle	Phosphofructokinase	Like type V	Muscle	Like type V

*Tissue obtained is used for the biochemical assay of the enzyme.
% = percentage of total number of cases in USA and Europe.

Pentosuria

Pentosuria is due to reduced activity of NADP-linked sylitol dehydrogenase. It has no clinical significance.

INBORN ERRORS OF AMINO ACID METABOLISM

Inborn errors of amino acid metabolism are chiefly inherited as autosomal recessive conditions. The major ones are shown in Table 19.19.

Amino acid transport defects

Amino acids are filtered by the glomerulus, but 95% of the filtered load is reabsorbed in the proximal convoluted tubule by an active transport mechanism. Aminoaciduria results from:

- abnormally high plasma amino acid levels (e.g. phenylketonuria)
- any inherited disorder that damages the tubules secondarily (e.g. galactosaemia)
- tubular reabsorptive defects, either generalized (e.g. Fanconi syndrome) or specific (e.g. cystinuria).

Amino acid transport defects can be congenital or acquired.

GENERALIZED AMINOACIDURIAS

Fanconi syndrome

This occurs in a juvenile form (De Toni–Fanconi–Debré syndrome); in adult life it is often acquired through, for example, heavy metal poisoning, drugs or some renal diseases. There is a generalized defective proximal tubular reabsorption of:

- most amino acids
- glucose
- urate
- phosphate, resulting in hypophosphataemic rickets
- bicarbonate, with failure to transport hydrogen ions, causing a renal tubular acidosis that then produces a hyperchloraemic acidosis (see p. 676).

Other abnormalities include:

Table 19.19 The major inborn errors of amino acid metabolism

Disease	Enzyme defect	Incidence	Biochemical and clinical features	Treatment	Prognosis
Albinism	Tyrosinase	1 in 13 000	Amelanosis: whitish hair, pink-white skin, grey-blue eyes Nystagmus, photophobia, strabismus	Symptomatic	Good
Alkaptonuria	Homogentisic acid oxidase	1 in 100 000	Homogentisic acid polymerizes to produce a brown-black product that is deposited in cartilage and other tissues (ochronosis)	None	Good
Homocystinuria				–	–
Type I	Cystathionine synthase		Homocystine is excreted in urine	–	–
			Learning difficulties	–	–
			Marfan-like syndrome	–	–
			Thrombotic episodes	–	–
Type II	Methylene tetrahydrofolate reductase		Survivors have learning difficulties	–	Many die as neonates
Phenylketonuria	Phenylalanine hydroxylase	1 in 20 000	Brain damage with learning difficulties and epilepsy Phenylpyruvate and its derivatives excreted in urine	Diet low in phenylalanine in first few months of life prevents damage	Good but some intellectual impairment
Histidinaemia	Histidase	Very rare	Learning difficulties	–	–
'Maple syrup' disease	Branched-chain ketoacid dehydrogenase	Very rare	Failure to thrive Fits, neonatal acidosis and severe cerebral degeneration Valine, isoleucine and their derivatives are excreted in urine A milder form is seen	–	Early death
Oxalosis (hyperoxaluria)	Alanine: glyoxylate aminotransferase	Very rare	Nephrocalcinosis, renal stones, renal failure due to deposition of calcium oxalate Prenatal and early diagnosis now possible	–	Liver transplantation ? Gene therapy

There are many other enzyme defects producing, for example, alaninaemia, ammonaemia, argininaemia, citrullinaemia, isovaleric acidaemia, lysinaemia, ornithinaemia or tyrosinaemia.

- potassium depletion, primary or secondary to the acidosis
- polyuria
- increased excretion of immunoglobulins and other low-molecular-weight proteins.

Various combinations of the above abnormalities have been described.

The juvenile form begins at the age of 6–9 months, with failure to thrive, vomiting and thirst. The clinical features are as a result of fluid and electrolyte loss and the characteristic vitamin D-resistant rickets.

In the adult, the disease is similar to the juvenile form, but osteomalacia is a major feature.

Treatment of the bone disease is with large doses of vitamin D (e.g. 1–2 mg of 1α-hydroxycholecalciferol with regular blood calcium monitoring). Fluid and electrolyte loss need to be corrected.

Lowe's syndrome (oculocerebrorenal dystrophy)

In this syndrome there is generalized aminoaciduria combined with mental retardation, hypotonia, congenital cataracts and an abnormal skull shape.

SPECIFIC AMINOACIDURIAS

Cystinuria

There is a defective tubular reabsorption and jejunal absorption of cystine and the dibasic amino acids, lysine, ornithine and arginine. Inheritance is either completely or incompletely recessive, so that only heterozygotes who have increased excretion of lysine and cystine can occur. Cystine absorption from the jejunum is impaired but, nevertheless, cystine in peptide form can be absorbed. Cystinuria leads to urinary

stones and is responsible for approximately 1–2% of all urinary calculi. The disease often starts in childhood, although most cases present in adult life.

Treatment is with a high fluid intake in order to keep the urinary cystine concentration low. Patients are encouraged to drink up to 3 L over 24 hours and to drink even at night. Penicillamine should be used for patients who cannot keep the cystine concentration of their urine low.

The condition cystinosis must not be confused with cystinuria.

Hartnup's disease

There is defective tubular reabsorption and jejunal absorption of most neutral amino acids but not their peptides. The resulting tryptophan malabsorption produces nicotinamide deficiency (see p. 222). Patients can be asymptomatic, but others develop evidence of pellagra, with cerebellar ataxia, psychiatric disorders and skin lesions. Treatment is with nicotinamide, which often brings about considerable improvement.

Tryptophan malabsorption syndrome (blue diaper syndrome)

This is due to an isolated transport defect for tryptophan: the tryptophan excreted oxidizes to a blue colour on the baby's diaper.

Familial iminoglycinuria

This occurs when there is defective tubular reabsorption of glycine, proline and hydroxyproline. It seems to have few clinical effects.

Methionine malabsorption syndrome

This is due to failure to absorb and excrete methionine, and results in diarrhoea, vomiting and mental retardation. Patients characteristically have an oast-house smell.

LYSOSOMAL STORAGE DISEASES

Lysosomal storage diseases are due to inborn errors of metabolism which are mainly inherited in an autosomal recessive manner. See Table 19.18.

Glucosylceramide lipidoses: Gaucher's disease

This is the most prevalent lysosomal storage disease and is due to a deficiency in glucocerebrosidase, a specialized lysosomal acid β-glucosidase. This results in accumulation of glucosylceramide in the lysosomes of the reticuloendothelial system, particularly the liver, bone marrow and spleen. Over 300 mutations have been characterized in the glucocerebrosidase gene, the most common being a single base change (N370S) causing the substitution of arginine for serine; this is seen in 70% of Jewish patients. The typical Gaucher cell, a glucocerebroside-containing reticuloendothelial histiocyte, is found in the bone marrow, producing many cytokines such as CD14.

There are three clinical types, the most common presenting in childhood or adult life with an insidious onset of hepatosplenomegaly (type 1). There is a high incidence in Ashkenazi Jews (1 in 3000 births), and patients have a characteristic pigmentation on exposed parts, particularly the forehead and hands. The clinical spectrum is variable, with patients developing anaemia, evidence of hypersplenism and pathological fractures that are due to bone involvement. Nevertheless, many have a normal life-span.

The diagnosis is made on finding a deficiency of lysosomal B glucocerebrosidase in leucocytes. Plasma chitotriosidase (an enzyme secreted by activated macrophages) is grossly elevated in Gaucher's disease and other lysosomal disorders: it is used to monitor enzyme replacement therapy. Acute Gaucher's disease (type 2) presents in infancy with rapid onset of hepatosplenomegaly, with neurological involvement owing to the presence of Gaucher cells in the brain. The outlook is very poor. Type 3 presents in childhood or adolescence with a variable progression of hepatosplenomegaly, neurodegeneration and bone disease. Again, the outlook is poor.

Some patients with non-neuropathic Gaucher's disease show considerable improvement with infusion of human recombinant glucocerebrosidase (imiglucerase – a human recombinant enzyme). Oral miglustat (an inhibitor of glucosylceramide synthase) is used for mild to moderate type 1 Gaucher's disease.

Sphingomyelin cholesterol lipidoses: Niemann–Pick disease

The disease is due to a deficiency of lysosomal sphingomyelinase which results in the accumulation of sphingomyelin cholesterol and glycosphingolipids in the reticuloendothelial macrophages of many organs, particularly the liver, spleen, bone marrow and lymph nodes. The disease usually presents within the first 6 months of life with mental retardation and hepatosplenomegaly; a particular type (11c) presents in adults with dementia. The gene frequency is 1:100 in Ashkenazi Jews. Typical foam cells are found in the marrow, lymph nodes, liver and spleen.

The mucopolysaccharidoses (MPSs)

This is a group of disorders caused by the deficiency of lysosomal enzymes (e.g. α-L-iduronidase) required for the catabolism of glycosaminoglycans (mucopolysaccharides).

The catabolism of dermatan sulphate, heparan sulphate, keratin sulphate or chondroitin sulphate may be affected either singularly or together.

Accumulation of glycosaminoglycans in the lysosomes of various tissues results in the disease. Ten forms of MPS have been described; all are chronic but progressive, and a wide spectrum of clinical severity can be seen within a single enzyme defect. The MPS types show many clinical features though in variable amounts, with dysostosis, abnormal facies, poor vision and hearing and joint dysmobility (either stiff or hypermobile) being frequently seen. Mental retardation is present in, for example, Hurler (MPS IH) and San Filippo A (MPS IIIA) types, but normal intelligence and life-span are seen in Scheie (MPS IS). L-aronidase

infusion reduces lysosomal storage, resulting in clinical improvement.

The GM2 gangliosidoses

In these conditions there is accumulation of GM2 gangliosides in the central nervous system and peripheral nerves. It is particularly common (1 in 2000) in Ashkenazi Jews. Tay–Sachs disease is the severest form, where there is a progressive degeneration of all cerebral function, with fits, epilepsy, dementia and blindness, and death usually occurs before 2 years of age. The macula has a characteristic cherry spot appearance.

Fabry's disease

This X-linked recessive condition involves the glycosphingolipid pathway. There is a deficiency of lysosomal hydrolase (α-galactosidase A) causing an accumulation of globotriaosylceramide with terminal α-galactosyl moieties in the lysosomes of various tissues including the liver, kidney, blood vessels and the ganglion cells of the nervous system. The patients present with peripheral nerve involvement, gastrointestinal symptoms/abdominal pain, diarrhoea and early saiety, but eventually most patients have cardiac strokes and kidney disease in adult life. An absent or very low level of alpha-gal A in leucocytes confirms the diagnosis. Genetic testing is available. Treatment is with agalsidase alpha beta infusions.

Diagnosis

Many of the sphingolipidoses can be diagnosed by demonstrating the enzyme deficiency, usually in peripheral blood leucocytes.

Prenatal diagnosis is possible in a number of the conditions by obtaining specimens of amniotic cells. Carrier states can also be identified, so that sensible genetic counselling can be given.

FURTHER READING

Grabouski GA. Phenotype, diagnosis and treatment of Gaucher's disease. *Lancet* 2008; **372:** 1263–1271.

Van der Ploeg AT, Reuser AJ. Pompe's disease. *Lancet* 2008; **372:** 1342–1353.

Shiffman R. Enzyme replacement in Fabry's disease. *American International Medicine* 2007; **146**: 142–151.

Zarate YA, Hopkin RJ. Fabry's disease. *Lancet* 2008; **372:** 1427–1435.

AMYLOIDOSIS

Amyloidosis is a disorder of protein metabolism in which there is an extracellular deposition of pathological insoluble fibrillar proteins in organs and tissues. Characteristically, the amyloid protein consists of β-pleated sheets that are responsible for its insolubility and resistance to proteolysis.

Amyloidosis can be acquired or inherited. Classification is based on the nature of the precursor plasma proteins (at least 20) that form the fibrillar deposits. The process for the production of these fibrils appears to be multifactorial and differs amongst the various types of amyloid.

AL amyloidosis (immunoglobulin light chain-associated)

This is a plasma cell dyscrasia, related to multiple myeloma, in which clonal plasma cells in the bone marrow produce immunoglobulins that are amyloidogenic. This may be the outcome of destabilization of light chains owing to substitution of particular amino acids into the light chain variable region. There is a clonal dominance of amyloid light (AL) chains – either the dominant κ or λ isotype – which are excreted in the urine (Bence Jones proteins). This type of amyloid is often associated with lymphoproliferative disorders, such as myeloma, Waldenström's macroglobulinaemia or non-Hodgkin's lymphoma. It rarely occurs before the age of 40 years.

The clinical features are related to the organs involved. These include the kidneys (presenting with proteinuria and the nephrotic syndrome) and the heart (presenting with heart failure). Autonomic and sensory neuropathies are relatively common, and carpal tunnel syndrome with weakness and paraesthesia of the hands may be an early feature. Sensory neuropathy is common. There is an absence of central nervous system involvement.

On examination, hepatomegaly and rarely splenomegaly, cardiomyopathy, polyneuropathy and bruising may be seen. Macroglossia occurs in about 10% of cases and periorbital purpura in 15%.

Familial amyloidoses (transthyretin-associated (ATTR))

These are autosomally dominant transmitted diseases where the mutant protein forms amyloid fibrils, starting usually in middle age. The most common form is due to a mutant – transthyretin – which is a tetrameric protein with four identical subunits. It is a transport protein for thyroxine and retinol-binding protein and mainly synthesized in the liver. Over 80 amino acid substitutions have been described; for example, a common substitution is that of methionine for valine at position 30 (Met 30) in all racial groups, and alanine for threonine (Ala 60) in the English and Irish. These substitutions destabilize the protein, which precipitates following stimulation, and can cause disorders such as familial amyloidotic polyneuropathy (FAP), cardiomyopathy or the nephrotic syndrome. Major foci of FAP occur in Portugal, Japan and Sweden.

Other less common variants include mutations of apoprotein A-I, gelsolin, fibrinogen Aα and lysozyme.

Clinically, peripheral sensorimotor and autonomic neuropathy are common, with symptoms of autonomic dysfunction, diarrhoea and weight loss. Renal disease is less prevalent than with AL amyloidosis. Macroglossia does not occur. Cardiac problems are usually those of conduction. There may be a family history of unidentified neurological disease.

Other hereditary systemic amyloidoses include other familial amyloid polyneuropathies (e.g. Portuguese, Icelandic, Dutch). There is a familial Creutzfeldt–Jakob disease. In familial Mediterranean fever, renal amyloidosis is a common serious complication.

Reactive systemic (secondary AA) amyloidoses

These are due to amyloid formed from serum amyloid A (SAA), which is an acute phase protein. It is, therefore, related to chronic inflammatory disorders and chronic infection.

Clinical features depend on the nature of the underlying disorder. Chronic inflammatory disorders include rheumatoid arthritis, inflammatory bowel disease and untreated familial Mediterranean fever. In developing countries it is still associated with infectious diseases such as tuberculosis, bronchiectasis and osteomyelitis. AA amyloidosis often presents with chronic kidney disease, with hepatomegaly and splenomegaly. Macroglossia is not a feature and cardiac involvement is rare. The degree of renal failure correlates with the

SAA level in a more favourable outcome in patients with low normal levels.

Other amyloides

Cerebral amyloidosis, Alzheimer's disease and transmissible spongiform encephalopathy

The brain is a common site of amyloid deposition, although it is not directly affected in any form of acquired systemic amyloidosis. Intracerebral and cerebrovascular amyloid deposits are seen in Alzheimer's disease. Most cases are sporadic, but hereditary forms caused by mutations have been reported. In hereditary spongiform encephalopathies several amyloid plaques have been seen.

Amyloid deposits are frequently found in the elderly, particularly cerebral deposits of A4 protein. This is also seen in Down's syndrome. Apoprotein E (involved in LDL transport, see p. 1062) interacts directly with β-A4 protein in senile plaques and neurofibrillary tangles in the brain. The gene for apoprotein E is on chromosome 19 and may be a susceptibility factor in the aetiology of Alzheimer's disease.

Local amyloidosis

Deposits of amyloid fibrils of various types can be localized to various organs or tissues (e.g. skin, heart and brain).

Dialysis-related amyloidosis

This is due to the β_2-microglobulin producing amyloid fibrils in chronic dialysis patients (see p. 638). It frequently presents with the carpal tunnel syndrome.

Diagnosis

This is based on clinical suspicion and, if possible, on tissue histology. Amyloid in tissues appears as an amorphous, homogeneous substance that stains pink with haematoxylin and eosin and stains red with Congo red. It also has a green fluorescence in polarized light. Tissue can be obtained from the rectum, gums or abdominal fat. The bone marrow may show plasma cells in amyloidosis or a lymphoproliferative disorder. A paraproteinaemia and proteinuria with light chains in the urine may also be seen in AL amyloidosis. In secondary or reactive amyloidosis there will be an underlying disorder. Scintigraphy using ^{123}I-labelled serum amyloid P component is useful for the assessment of AL, ATTR and AA amyloidosis, but it is not widely available and is expensive.

Treatment

This is symptomatic or the treatment of the associated disorder. The nephrotic syndrome and congestive cardiac failure require the relevant therapies. Treatment of any inflammatory source or infection should be instituted. Colchicine may help familial Mediterranean fever. Eprodisate, which interferes with interactions between amyloid proteins and glycosaminoglycans inhibits polymerization of amyloid fibrids; it slows the fall in renal function in AA amyloidosis. Chemotherapy with nephalan plus dexamethasone is showing some efficacy in AL amyloidosis. In ATTR amyloidosis where transthyretin is predominantly synthesized in the liver, liver transplantation (when there would be a disappearance of the mutant protein from the blood) is considered as the definitive therapy.

THE PORPHYRIAS

This heterogeneous group of rare inborn errors of metabolism is caused by abnormalities of enzymes involved in the biosynthesis of haem, resulting in overproduction of the intermediate compounds called 'porphyrins' (Fig. 19.22). The porphyrias show extreme genetic heterogeneity. For example, in acute intermittent porphyria more than 90 mutations have been identified in the porphobilinogen deaminase gene. One mutation has a high prevalence in patients in northern Sweden, suggesting a common ancestor.

Structurally, porphyrins consist of four pyrrole rings. These pyrrole rings are formed from the precursors glycine and succinyl-CoA, which are converted to δ-aminolaevulinic acid (δ-ALA) in a reaction catalysed by the enzyme δ-ALA synthase. Two molecules of δ-ALA condense to form a pyrrole ring.

Porphyrins can be divided into uroporphyrins, coproporphyrins or protoporphyrins depending on the structure of the side-chain. They are termed type I if the structure is symmetrical and type III if it is asymmetrical. Both uroporphyrins and coproporphyrins can be excreted in the urine.

The sequence of enzymatic changes in the production of haem is shown in Figure 19.22. The chief rate-limiting step is the enzyme δ-ALA synthase. This has two isoforms, ALA-N (non-erythroid) and ALA-E (erythroid). ALA-N is under negative feedback by haem but is upregulated by drugs and chemicals; there is no known inherited deficiency and the gene is on 3p21. Conversely ALA-E, encoded by Xp11.21, is unaffected by drugs or haem, and an inherited deficiency causes X-linked sideroblastic anaemia (Fig. 19.22).

Consequently:

- Haem (endogenous or exogenous) produces remission of hepatic porphyria.
- Chemicals and drugs can produce disease.
- Erythopoietic porphyria gives constant symptoms and is affected by sunlight.

Clinical features

All of the haem intermediates shown in Figure 19.22 are potentially toxic. Three patterns of symptoms occur in the various types of porphyria:

- neurovisceral (Table 19.20)
- photosensitive
- haemolytic anaemia.

The most common types of porphyria are acute intermittent (AIP), porphyria cutanea tarda (PCT) and erythropoietic porphyria (EPP).

The diagnosis of various types is based on urinary excretion levels of aminolaevulinic acid (ALA), porphobilinogen (PBG) and porphyrin in aliquots from a 24-hour collection. Erythrocyte PBG deaminase activity is measured for neurovisceral symptoms. Urinary and plasma porphyrin concentrations are helpful for those with photocutaneous manifestations. Faecal porphyrin analysis is helpful in confirmatory testing.

Neurovisceral

Acute intermittent porphyria (AIP) (Fig. 19.22)

This is an autosomal dominant disorder. Presentation is in early adult life, usually around the age of 30 years, and

FURTHER READING

Merlini G, Bellotti V. Molecular mechanisms of amyloidosis. *New England Journal of Medicine* 2003; **349**: 583–596.

Rajkumar SV, Gertz MA. Advances in the treatment of amyloidosis. *New England Journal of Medicine* 2007; **356**: 2413–2415.

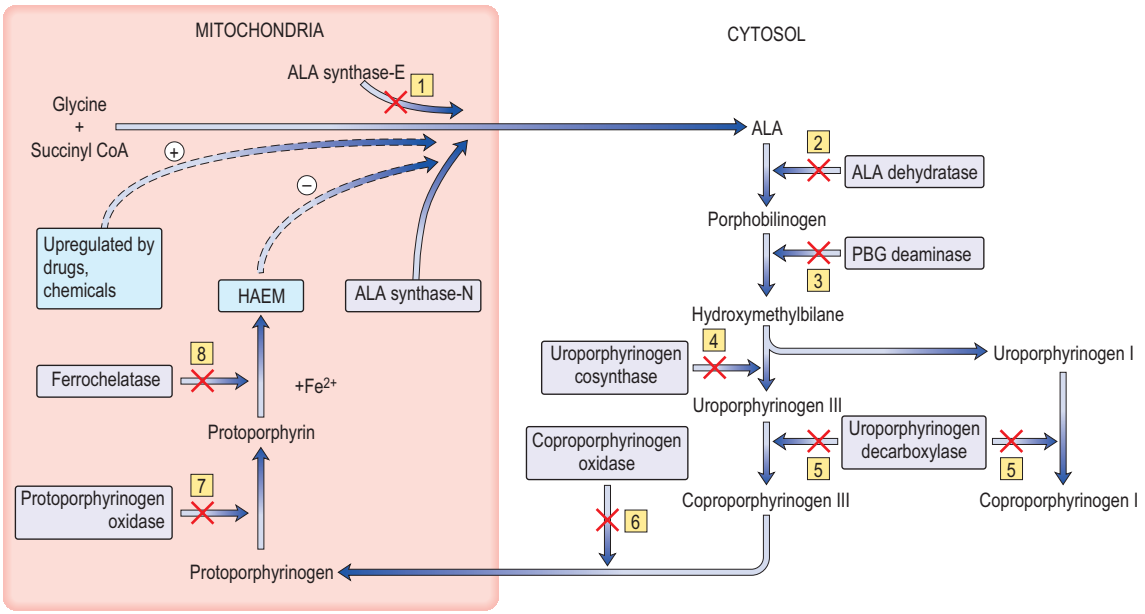


Fig. 19.22 **Porphyrin metabolism showing the various porphyrias.** ALA synthase is the rate-limiting enzyme. Deficiencies of the other seven enzymes (crosses) cause the different porphyrias:

1 X-linked sideroblastic anaemia
2 ALA dehydrase porphyria (ADP)
3 Acute intermittent porphyria (AIP)
4 Congenital erythropoietic porphyria (CEP)
5 Porphyria cutanea tarda (PCT) and hepatoerythropoietic porphyria (HEP)
6 Hereditary coproporphyria (HCP)
7 Variegate porphyria (VP)
8 Erythropoietic protoporphyria (EPP).

Table 19.20 Porphyria: neurovisceral symptoms

Neuropsychiatric	Visceral
Neuropathy:	Abdominal pain
Motor (70%)	Vomiting up to 90%
Sensory	Constipation
Epilepsy (15%)	Diarrhoea (occasional)
Psychiatric disorders (50%):	Fever (~ 30%)
Depression	Hypertension (up to 50%)
Anxiety	Tachycardia (up to 80%)
Psychosis	Muscular pain (~ 50%)

women are affected more than men. It may be precipitated by alcohol and drugs such as barbiturates and oral contraceptives, but a wide range of lipid-soluble drugs have also been incriminated. Acute attacks present with neurovisceral symptoms (Table 19.20). Symptoms of the rare, autosomal recessive aminolaevulinic acid dehydrogenase porphyria (ADP) are similar.

Investigations

The urine turns red-brown or red on standing.

- Blood count. This is usually normal, with occasional neutrophil leucocytosis.
- Liver biochemical tests. There is elevated bilirubin and aminotransferase.
- Serum urea is often raised.
- ALA and PBG are raised.
- Erythrocyte PBG deaminase is decreased.

Screening

Family members should be screened to detect latent cases. Urinalysis is not adequate but measurement of erythrocyte porphobilinogen deaminase and ALA synthase is extremely sensitive.

Mixed neurovisceral and photocutaneous

Variegate porphyria (VP) (Fig. 19.22)

This combines neurovisceral symptoms with those of a cutaneous photosensitive porphyria. A bullous eruption develops on exposure to sunlight owing to the activation of porphyrins deposited in the skin.

Investigation shows an elevated urinary ALA and PBG. Fluorescence emission spectroscopy of plasma differentiates this from other cutaneous porphyrias.

Hereditary coproporphyria (HCP) (Fig. 19.22)

This is extremely rare and broadly similar in presentation to variegate porphyria.

Photocutaneous

Porphyria cutanea tarda (cutaneous hepatic porphyria) (PCT) (Fig. 19.22)

This condition, which has a genetic predisposition, presents with a bullous eruption on exposure to sunlight; the eruption heals with scarring. Alcohol is the most common aetiological agent but hepatitis C, iron overload or HIV can also precipi-

tate the disease. Evidence of biochemical or clinical liver disease may also be present. Polychlorinated hydrocarbons have been implicated and porphyria cutanea tarda has been seen in association with benign or malignant tumours of the liver.

Hepatoerythropoietic porphyria (HEP, Fig. 19.22) is a rare disease clinically very similar to congenital erythropoietic porphyria presenting in childhood; haemolytic anaemia occurs. The defect of HEP is similar to that of PCT.

The diagnosis depends on demonstration of increased levels of urinary uroporphyrin. Histology of the skin shows subepidermal blisters with perivascular deposition of periodic acid–Schiff-staining material. The serum iron and transferrin saturation are often raised. Liver biopsy shows mild iron overload as well as features of alcoholic liver disease.

Congenital erythropoietic porphyria (CEP) (Fig. 19.22)

This is extremely rare and is transmitted as an autosomal recessive trait. Its victims show extreme sensitivity to sunlight and develop disfiguring scars. Dystrophy of the nails, blindness due to lenticular scarring, and brownish discoloration of the teeth also occur.

Erythropoietic protoporphyria (EPP) (Fig. 19.22)

This is more common than congenital erythropoietic porphyria and is inherited as an autosomal dominant trait. It presents with irritation and a burning pain in the skin on exposure to sunlight. The liver is usually normal but protoporphyrin deposition can occur. Diagnosis is made by fluorescence of the peripheral red blood cells and by increased protoporphyrin in the red cells and stools.

Management of porphyrias

Neurovisceral

Acute. The management of acute episodes is largely supportive. Precipitating factors, e.g. drugs, should be stopped. Analgesics should be given (avoiding drugs that may aggravate an attack). Intravenous carbohydrates, e.g. glucose, inhibit ALA synthase activity. Intravenous haem arginate (human haemin) infusion reduces ALA and PBG excretion by having a negative effect on ALA synthase-N activity (Fig. 19.22) and decreases the duration of an attack; this is useful in a severe attack. Calorie and fluid intake should be maintained.

Prevention in remission period. This is by avoidance of possible precipitating factors, e.g. drugs and alcohol. Stopping smoking, treatment of infections and stress avoidance are helpful. Surgery can precipitate attacks. A high-carbohydrate diet should be maintained and haemin infusions may also help.

Photocutaneous episodes

Acute attacks following exposure to UV light can only be treated symptomatically. However, venesection, which reduces urinary porphyria, can be used for PCT in both acute and remission phases. Chloroquine can also aid excretion by forming a water-soluble complex with uroporphyrins.

Liver transplantation is used for severe cases.

Prevention is with avoidance of sunlight, and use of sunscreens and protective clothing.

Oral β-carotene, which quenches free radicals, provides effective protection against solar sensitivity in EPP.

FURTHER READING

American Porphyria Foundation: http://www.porphyriafoundation.com.

Kauppinen R. Porphyrias. *Lancet* 2005; **365**: 241–252.

CHAPTER BIBLIOGRAPHY

Scriver CR, Beaudet AL, Sly WS et al. The Metabolic Basis of Inherited Disease, 9th edn. New York: McGraw-Hill, 2008.

National Collaborating Centre for Chronic Conditions. Type 1 diabetes in adults – national clinical guidelines for diagnosis and management in primary and secondary care. London: Royal College of Physicians, 2004.

SIGNIFICANT WEBSITES

http://medweb.bham.ac.uk/easdec/ *Diabetic retinopathy*

http://www.diabetes.org *American Diabetes Association – heavyweight and authoritative, with an American flavour*

http://www.diabetes.ca *Canadian Diabetes Association site – well designed practical site with many links to other diabetes-related sites; a good jumping off point*

http://www.diabetes.org.uk *Diabetes UK charity – information for patients, researchers and health professionals*

http://www.doh.gov.uk/nsf *UK Government forthcoming National Service Framework draft information*

http://www.dtu.ox.ac.uk *Diabetes Trials Unit (University of Oxford) – research information, particularly the UK Prospective Diabetes Study results*

http://www.eatlas.idf.org *International Diabetes Federation*

http://www.sign.ac.uk/guidelines/published/index.html*Scottish Intercollegiate Guidelines Network – guidelines on a range of subjects including diabetes*

This page intentionally left blank

20 The special senses

DISORDERS OF THE EAR, NOSE AND THROAT

The ear 1077
Anatomy and physiology 1077
Physiology of hearing 1077
Examination 1078
Common disorders 1078

The nose 1080
Anatomy and physiology 1080
Examination 1080
Common disorders 1080

The throat 1082
Anatomy and physiology 1082
Examination 1082
Common disorders 1082

DISORDERS OF THE EYE

Applied anatomy and physiology 1085
Refractive errors 1086
Disorders of the lids 1087
Conjunctivitis 1087
Corneal disorders 1089
Cataracts 1090
Glaucoma 1091
Uveitis 1092
Disorders of the retina 1093
Causes of visual loss 1094

DISORDERS OF THE EAR, NOSE AND THROAT

THE EAR

ANATOMY AND PHYSIOLOGY

The ear can be divided into three parts: outer, middle and inner (Fig. 20.1).

The *outer ear* has a skin-lined tube 2.5 cm long leading down to the tympanic membrane (the ear drum). Its outer third is cartilaginous and contains hair, sebaceous and ceruminous glands, but the walls of the inner two-thirds are bony. The outer ear is self-cleaning as the skin is migratory and there are no indications to use cotton wool buds. Wax should only be seen in the outer third.

The *middle ear* is an air-containing cavity derived from the branchial clefts. It communicates with the mastoid air cells superiorly, and the Eustachian tube connects it to the nasopharynx medially. The Eustachian tube ventilates the middle ear and maintains equal air pressure across the tympanic membrane. It is normally closed but opens via the action of the palatal muscles to allow air entry when swallowing or yawning. A defect in this mechanism, such as with a cleft palate, will prevent air entering the middle ear cleft which may then fill with fluid. Lying within the middle ear cavity are the three ossicles (malleus, incus and stapes) that transmit sound from the tympanic membrane to the inner ear. On the medial wall of the cavity is the horizontal segment of the facial nerve, which may be damaged during surgery or by direct extension of infection in the middle ear.

The *inner ear* contains the cochlea for hearing and the vestibule and semicircular canals for balance. There is a semicircular canal arranged in each body plane and these are stimulated by rotatory movement. The facial, cochlear and vestibular nerves emerge from the inner ear and run through the internal acoustic meatus to the brainstem (see Fig. 21.7, p. 1106).

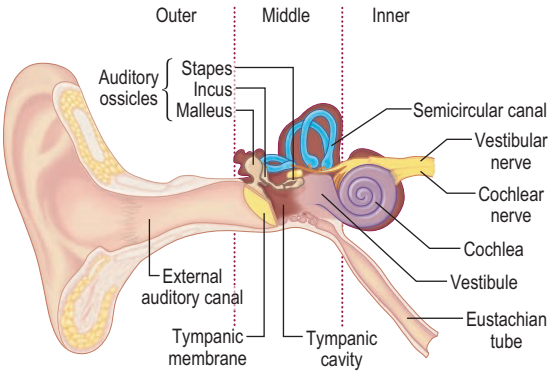


Fig. 20.1 **The anatomy of the ear.**

PHYSIOLOGY OF HEARING

The ossicles, in the middle ear, transmit sound waves from the tympanic membrane to the cochlea. They amplify the waves by about 18-fold to compensate for the loss of sound waves moving from the air-filled middle ear to the fluid-filled cochlea. Hair cells in the basilar membrane of the cochlea detect the vibrations and transduce these into nerve impulses which pass to the cochlear nucleus and then eventually to the superior olivary nuclei of both sides; thus lesions central to the cochlear nucleus do not cause unilateral hearing loss.

If the ossicles are diseased, sound can also reach the cochlea by vibration of the temporal bone (bone conduction).

EXAMINATION

The pinna and postauricular region should first be examined for scars or swellings. An auroscope is used to examine the external ear canal whilst the pinna is retracted backwards and upwards to straighten the canal. Look for wax, discharge or foreign bodies. The tympanic membrane should always be seen with a light reflex anteroinferiorly. Previous repeated infections may cause a thickened, whitish drum but fluid in the middle ear may show as dullness of the drum. Perforations are marginal or central.

COMMON DISORDERS

The discharging ear (otorrhoea)

Discharge from the ear is usually due to infection of the outer or middle ear.

Otitis externa is a diffuse inflammation of the skin of the ear canal. The organism may be bacterial, viral or fungal and the patient usually complains of severe pain. Gentle pulling of the pinna is tender and there may be lymphadenopathy of the preauricular nodes.

Examination may reveal debris in the canal which needs to be removed either by gentle mopping or preferably by suction viewed directly under a microscope. In severe cases the canal may be swollen and a view of the tympanic membrane impossible. Any foreign body seen should be removed with great care by trained personnel.

Treatment is with regular cleansing and topical antibiotics combined with corticosteroids; it resolves in 3–4 days.

Otitis media can also present with discharge from the middle ear through a perforation of the tympanic membrane. There are no mucous glands in the external ear canal, however, and if the discharge is serous then middle ear pathology is unlikely.

Treatment is with systemic antibiotics.

Cholesteatoma

Cholesteatoma is defined as keratinizing squamous epithelium within the middle ear cleft and can present with foul-smelling otorrhoea. Examination may show a defect in the tympanic membrane full of white cheesy material. Mastoid surgery is required to remove this sac of squamous debris as it can erode local structures such as the facial nerve or even extend intracranially.

Hearing loss

Deafness can be conductive or sensorineural and these can be differentiated at the bedside by the Rinne and the Weber tests (Box 20.1) or with pure-tone audiometry.

Rinne test

Normally a tuning fork, 512 Hz, will be heard as louder if held next to the ear (air conduction) compared to being placed on the mastoid bone (*Rinne positive*). If the tuning fork is perceived louder when placed on the mastoid (bone conduction), then a defect in the conducting mechanism of the external or middle ear is present (true *Rinne negative*).

Box 20.1 Hearing loss

Type	Defect	Rinne	Weber
Conductive	Outer or middle ear	Negative	Sound heard louder on the affected side
Sensorineural	Inner ear or more centrally	Positive	Sound heard louder in the normal ear

Table 20.1 Deafness

Conductive	Sensorineural
Congenital – atresia Pendred's syndrome (see p. 985) Long QT syndrome (see p. 727) Björnstad syndrome (pili torti)	**End organ** Advancing age Occupational acoustic trauma Ménière's disease Drugs (e.g. gentamicin, furosemide)
External meatus Wax Foreign body Otitis externa Chronic suppuration	**Eighth nerve lesions** Acoustic neuroma Cranial trauma Inflammatory lesions: Tuberculous meningitis Sarcoidosis Neurosyphilis Carcinomatous meningitis
Drum Perforation/trauma	**Brainstem lesions (rare)** Multiple sclerosis Infarction
Middle ear Otosclerosis Ossicular bone problems Suppuration (otitis media)	

Weber test

A tuning fork placed on the forehead or vertex of a patient with normal hearing (or with symmetrical hearing loss) should be perceived centrally.

Conductive hearing loss may be due to many causes (Table 20.1) but wax is the commonest.

Pure tone audiometry

The patient is asked to respond to a series of pure tones presented to each ear, in turn, in a soundproof room. An audiogram is produced (Fig. 20.3).

Perforated tympanic membrane

This may arise from trauma or chronic middle ear disease where recurrent infection results in a permanent defect. Surgical repair is only indicated if the patient is symptomatic with hearing loss or recurrent discharge.

Otitis media

This is an acute inflammation of the middle ear, causing severe pain (otalgia) and conductive hearing loss. This occurs because fluid accumulation in the middle ear impairs sound conduction to the cochlea. It is often viral in origin, e.g. following a cold, and will settle within 72 hours without antibacterial treatment. In patients with systemic features or after 72 hours, a systemic antibiotic, e.g. amoxicillin, should be given, particularly in the under 2 year olds. Topical therapy is of no value. Complications include infection of the mastoid bone.

Acute otitis media may spread to the mastoid area and if there is tenderness and swelling over the mastoid then an urgent ENT opinion should be obtained.

Secretory otitis media with effusion (also called serous otitis media or glue ear) (Fig. 20.2)

This is common in children because of Eustachian tube dysfunction. The effusion resolves naturally in the majority of cases but can persist giving hearing loss, and it predisposes to recurrent attacks of acute otitis media. A grommet (tympanostomy tube) is a tube that is inserted into the tympanic membrane and ventilates the middle ear cavity, i.e. it takes over the Eustachian tube's function. Grommets are extruded from the tympanic membrane as it heals (lasting from 6 months to 2 years). However, developmental outcomes are not improved by grommet insertions. In most children the middle third of the face grows around the age of 7–14 years and Eustachian tube dysfunction is rare after this.

Otosclerosis

This is usually a hereditary disorder where new bony deposits occur within the derivatives of the otic capsule, specifically the stapes footplate and in the cochlea. Characteristically seen in the second and third decades, it is commoner in females and can become worse during pregnancy. The hearing loss may be mixed, and treatment includes a hearing aid or replacement of the fixed stapes with a prosthesis (stapedectomy).

Presbycusis

This is the commonest cause of deafness. It is a degenerative disorder of the cochlea and is typically seen in old age. It can be due to the loss of outer hair cells (sensory), loss of the ganglion cells (neural), strial atrophy (metabolic) or it could be a mixed picture. Ageing itself does not cause outer hair cell loss but environmental noise toxicity over the years is a major factor. The onset is gradual and the higher frequencies are affected most (Fig. 20.3). Speech has two components: low frequencies (vowels) and high frequencies (consonants). When the consonants are lost, speech loses its intelligibility. Increasing the volume merely increases the low frequencies and the characteristic response of 'Don't shout. I'm not deaf!' A high-frequency-specific hearing aid will do much to ease the frustrations of both the patients and their close contacts.

Noise trauma

Cochlear damage can occur, for example, when shooting without ear protectors or from industrial noise (see p. 959) and characteristically has a loss at 4 kHz.

Fig. 20.2
Secretory otitis media (glue ear). Auroscopic view of tympanic membrane which is dull with loss of light reflex.

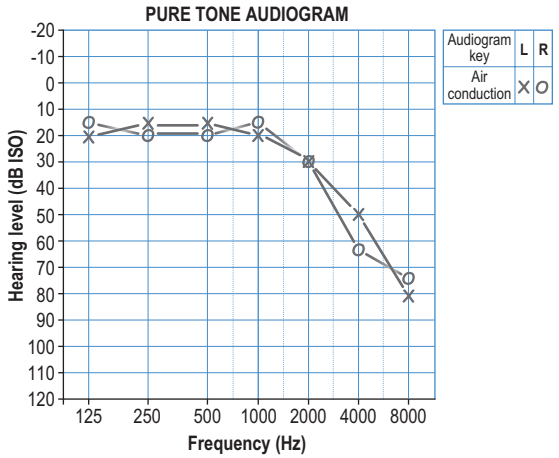


Fig. 20.3 **Audiogram showing presbycusis (high frequency loss).**

Acoustic neuroma

This is a slow-growing benign schwannoma of the vestibular nerve (see p. 1160) which can present with progressive sensorineural hearing loss. Any patient with an asymmetric sensorineural hearing loss or sudden sensorineural hearing loss should be investigated, e.g. with an MRI scan.

Vertigo

Vertigo is usually rotatory when it arises from the ear. The presence of otalgia, otorrhoea, tinnitus or hearing loss suggests an otologic aetiology. Vestibular causes can be classified according to the duration of the vertigo:

- seconds to minutes – benign paroxysmal positional vertigo
- minutes to hours – Ménière's disease
- hours to days – labyrinthine or central pathology.

Benign paroxysmal positional vertigo (BPPV)

BPPV is thought to be due to loose otoliths in the semicircular canals, commonly the posterior canal. Positional vertigo is precipitated by head movements, usually to a particular position, and may occur when turning in bed or on sitting up. The onset is typically sudden and distressing. The vertigo lasts seconds or minutes and the phenomenon becomes less severe on repeated movements (fatigue). There is no serious underlying cause but it sometimes follows vestibular neuronitis (see p. 1109), head injury or ear infection.

Diagnosis

This is made not only on the history but by precipitating an attack by carrying out a positional test in the outpatient clinic. The *Hallpike manoeuvre* involves sitting the patient on a couch and then turning the patient's head towards the affected ear. The patient's head is supported by the examiner and the patient then lies down so that the head is just below the horizontal. Nystagmus (following a latent interval of a few seconds) is commonly noted, as is any subjective sensation of vertigo. A positive Hallpike test confirms BPPV, which can be cured in over 90% of cases by the Epley manoeuvre. This involves gentle but specific manipulation and rotation of the patient's head to shift the loose otoliths from the semicircular canals.

A differential diagnosis is a cerebellar mass, but here positional nystagmus (and vertigo) is immediately apparent (no latent interval) and does not fatigue.

Ménière's disease

This condition is characterized by recurring episodic rotatory vertigo lasting 30 minutes to a few hours; attacks are recurrent over months or years. Classically it is associated with a low frequency sensorineural hearing loss, feeling of fullness in the affected ear, loss of balance, tinnitus and vomiting. There is a build-up of endolymphatic fluid in the inner ear, although its precise aetiology is still unclear.

Treatment involves the use of vestibular sedatives, e.g. cinnarizine in the acute phase. Preventative measures, such as a low-salt diet, betahistine and avoidance of caffeine are useful. If the disease cannot be controlled in this way, then a chemical labyrinthectomy, which involves perfusing the round window orifice with ototoxic drugs such as gentamicin, is possible. Gentamicin destroys the vestibular epithelium; therefore the patient has severe vertigo for around 2 weeks until the body compensates for the lack of vestibular input on that side. The patient will happily trade occasional mild vertigo when the balance system is challenged to the unpredictable, severe and disabling attacks of vertigo of Ménière's disease.

Labyrinthine or central causes of vertigo

(see Table 21.10)

These are managed with vestibular sedatives in the acute phase. Most patients will settle over a few days but continuous true vertigo with nystagmus suggests a central lesion. A patient with a deficit of vestibular function due to viral labyrinthitis or neuronitis should be able to come off the vestibular sedatives within 2 weeks, as long-term use can give parkinsonian side-effects, delay central compensation and thus prolong the vertigo. Vestibular rehabilitation by a physiotherapist or audiological scientist can speed up the compensation process, although most patients will be able to do this themselves with time.

Tinnitus

This is a sensation of a sound when there is no auditory stimulus. It can occur without hearing loss. Patients describe a hissing or ringing in their ears and this can cause much distress. It usually does not have a serious cause but vascular malformation, e.g. aneurysms, or vascular tumours may be associated.

Tinnitus associated with deafness is described above.

Treatment

This is difficult. A tinnitus masker (a mechanically produced continuous soft sound) can help. Specialist audiological services are of use and may rehabilitate patients well.

FURTHER READING

Halmagyi GM. Diagnosis and management of vertigo. *Clinical Medicine* 2005; **5**: 159–165.

Hilton M, Pinder D. Benign paroxysmal positional vertigo. *British Medical Journal* 2003; **326**: 673.

Lockwood AH. Tinnitus. *New England Journal of Medicine* 2002; **348**: 1027–1032.

Luxon LM, Davies RA (eds). *Handbook of Vestibular Rehabilitation.* London: Whurr, 1997.

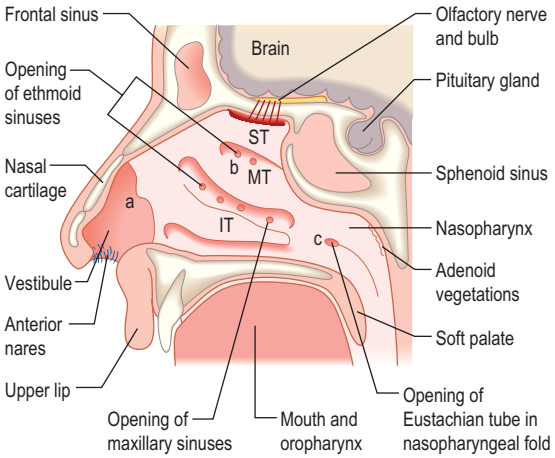


Fig. 20.4 **The anatomy of the nose in longitudinal section.** IT, inferior turbinate; MT, middle turbinate; ST, superior turbinate; a, internal ostium; b, respiratory region; c, choanae.

THE NOSE

ANATOMY AND PHYSIOLOGY (Fig. 20.4)

The function of the nose is to facilitate smell and respiration. Smell is a sensation conveyed by the olfactory epithelium in the roof of the nose. The olfactory epithelium is supplied by the first cranial nerve (see p. 1100). The nose also filters, moistens and warms inspired air and in doing so assists the normal process of respiration.

The external portion of the nose consists of two nasal bones attached to the rest of the facial skeleton and to the upper and lower lateral cartilages. The internal nose is divided by a midline septum that comprises both cartilage and bone. This separates the internal nose from the external nostril to the posterior choanae. The posterior choanae are in continuity with the nasopharynx posteriorly. The paranasal sinuses open into the lateral wall of the nose and are a system of aerated chambers within the facial skeleton.

The blood supply of the nose is derived from branches of both the internal and external carotid arteries. The internal carotid artery supplies the upper nose via the anterior and posterior ethmoidal arteries. The external carotid artery supplies the posterior and inferior portion of the nose via the superior labial artery, greater palatine artery and sphenopalatine artery. On the anterior nasal septum is an area of confluence of these vessels (Little's area) (Fig. 20.5a).

EXAMINATION

The anterior part of the nose can be examined using a nasal speculum and light source. Endoscopes are required to examine the nasal cavity and postnasal space.

COMMON DISORDERS

Epistaxis

Nose bleeds vary in severity from minor to life-threatening. Little's area (Fig. 20.5a) is a frequent site of nasal haemorrhage. First aid measures should be administered immediately, including external compression of the anterior lower portion of the external nose, ice packs and leaning forward. The patient should be asked to avoid swallowing any blood

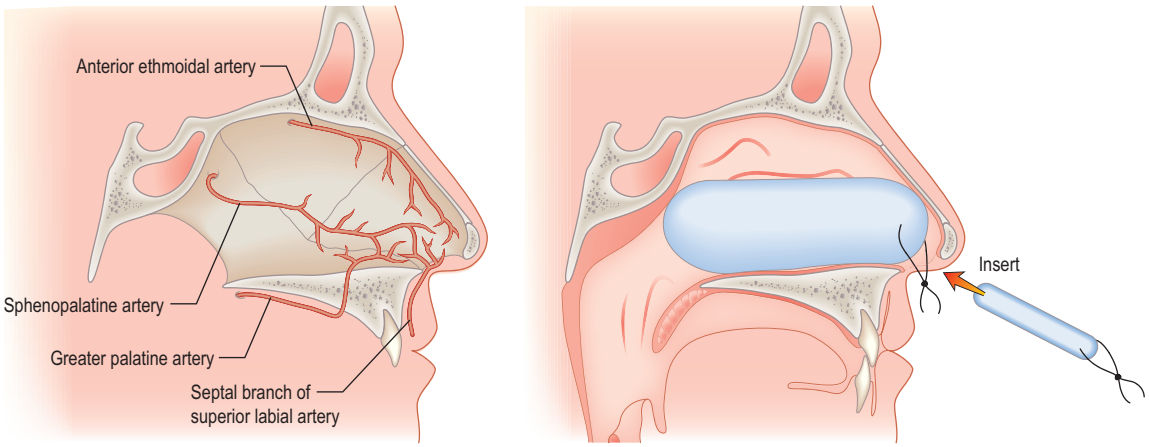


Fig. 20.5 **(a) Blood supply of Little's area on the septum of the nose. (b) Nasal pack.**

Table 20.2	Aetiology of epistaxis
Local	Idiopathic
	Trauma – foreign bodies, nose-picking and nasal fractures
	Iatrogenic – surgery, intranasal steroids
	Neoplasm – nasal, paranasal sinus and nasopharyngeal tumours
General	Anticoagulants
	Coagulation disorders (see p. 433)
	Hypertension
	Osler–Weber–Rendu syndrome (familial haemorrhagic telangiectasia)

running posteriorly. If the bleeding continues profusely then resuscitation in the form of intravenous access, fluid replacement or blood, and oxygen can be administered. If further intervention is necessary, consideration should be given to intranasal cautery of the bleeding vessel, or intranasal packing may be undertaken using a variety of commercially available nasal packs (Fig. 20.5b). In addition to direct treatment of the epistaxis, a cause and appropriate treatment of a cause should be sought (Table 20.2).

Rhinitis

See page 831.

Nasal obstruction

Nasal obstruction is a symptom and not a diagnosis. It can significantly affect a patient's quality of life. Causes include:

- *Rhinitis* (see p. 831). If an allergen is identified, then allergen avoidance is the mainstay of treatment. Topical steroids and/or antihistamines can be tried. If severe, then oral antihistamines or referral to an allergy clinic for immunotherapy may be warranted.
- *Septal deviation*. Correction of this deviation can be undertaken surgically.
- *Nasal polyps*. This condition occurs with inflammation and oedema of the sinus nasal mucosa. This oedematous mucosa prolapses into the nasal cavity and can cause significant nasal obstruction. In allergic rhinitis (see p. 831) the mucosa lining the nasal septum and inferior turbinates are swollen and a dark red or plum colour. Nasal polyps can be identified as glistening swellings which are not tender. Treatment with intranasal steroids may help but if polyps are large or unresponsive to medical treatment then surgery is necessary.
- *Foreign bodies*. These are usually seen in children who present with unilateral nasal discharge. Clinical examination of the nose with a light source often reveals the foreign body, which requires removal either in clinic or in theatre with a general anaesthetic.
- *Sinonasal malignancy*. This is extremely rare. The diagnosis must be considered if unusual unilateral symptoms are seen, including nasal obstruction, epistaxis, pain, epiphoria, cheek swelling, paraesthesia of the cheek and proptosis of the orbit.

Sinusitis

Sinusitis is an infection of the paranasal sinuses that may be bacterial (mainly *Streptococcus pneumoniae* and *Haemophilus influenzae*) or occasionally fungal. It is most commonly associated with an upper respiratory tract infection and can occur with asthma. Symptoms include frontal headache, purulent rhinorrhoea, facial pain with tenderness and fever. It can be confused with a variety of other conditions such as migraine, trigeminal neuralgia and cranial arteritis.

Treatment

Treatment for a bacterial sinusitis includes nasal decongestants, e.g. xylometazoline, broad-spectrum antibiotics, e.g. co-amoxiclav because *H. influenzae* can be resistant to amoxicillin, anti-inflammatory therapy with topical corticosteroids such as fluticasone propionate nasal spray to reduce mucosal swelling, and steam inhalations.

If the symptoms of sinusitis are recurrent (Box 20.2) or complications such as orbital cellulitis arise, then an ENT opinion is appropriate and a CT scan of the paranasal sinuses is undertaken. Plain sinus X-rays are now rarely used to image the sinuses.

CT scan of the sinuses or an MRI scan can demonstrate bony landmarks and soft tissue planes (Fig. 20.6).

Functional endoscopic sinus surgery (FESS) is used for ventilation and drainage of the sinuses.

Ask the authors
www.studentconsult.com

Box 20.2 Types of sinusitis

Acute sinusitis	Symptoms of sinusitis lasting between 1 week and 1 month
Recurrent acute sinusitis	More than four episodes of acute sinusitis per year
Subacute sinusitis	Symptoms of sinusitis lasting between 1 and 3 months
Chronic sinusitis	Symptoms of sinusitis lasting longer than 3 months

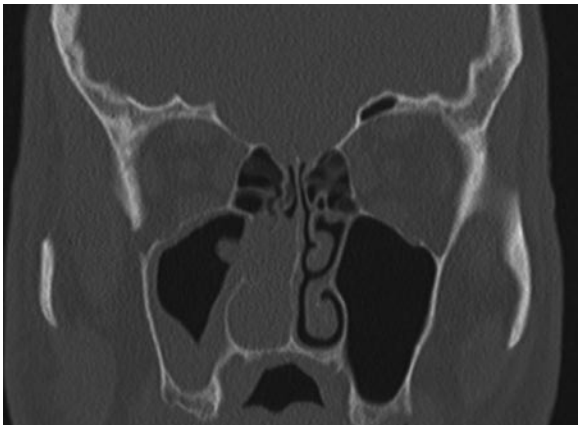


Fig. 20.6 **CT of the sinuses showing an inverting papilloma obstructing the right nostril and mucosal thickening of the right maxillary sinus.**

Anosmia

Olfaction is mainly under the control of cranial nerve I, although irritant, unpleasant nasal sensations are carried by cranial nerves V, IX and X. Anosmia is a complete loss of the sense of smell and *hyposmia* is a decreased sense of smell. If odorant molecules do not reach the olfactory epithelium high in the nose, a *conductive deficit* of smell occurs. If the neural transmission of smell is affected then a *sensorineural loss* of smell is incurred. Some conditions may predispose to a mixed (conductive and sensorineural) loss of smell. The main cause of a loss of smell is nasal obstruction due to upper respiratory infection or nasal polyps. Other causes include sinonasal disease, old age, drug therapy and head injury/trauma. It is difficult to predict the speed and extent of recovery in the latter causes.

Idiopathic cases will account for many patients but before this diagnosis is accepted an assessment of the patient for the possibility of an intranasal tumour or intracranial mass should be undertaken.

Fractured nose

Patients with a fractured nose present with epistaxis, bruising of the eyes and nasal bridge swelling. Initially, it is often difficult to assess if the bones are deviated, particularly if there is significant swelling. Reduction of the fracture should be undertaken in the first 2 weeks after injury and can be achieved by manipulation. However, if the fracture sets, a more formal rhinoplasty may have to be undertaken at a later stage. The patient should be examined for a head injury and the nose should also be checked for a septal haematoma. This is painful, can cause nasal obstruction, is fluctuant to touch on the nasal septum and requires immediate drainage.

THE THROAT

ANATOMY AND PHYSIOLOGY

The throat can be considered as the oral cavity, the pharynx and the larynx. The oral cavity extends from the lips to the tonsils. The pharynx can be divided into three areas:

- nasopharynx – extending from the posterior nasal openings to the soft palate
- oropharynx – extending from the soft palate to the tip of the epiglottis
- hypopharynx – extending from the tip of the epiglottis to just below the level of the cricoid cartilage where it is continuous with the oesophagus.

Lying within the hypopharynx is the larynx. This consists of cartilaginous, ligamentous and muscular tissue with the primary function of protecting the distal airway. The pharynx is innervated from the pharyngeal plexus. In the larynx there are two vocal cords which abduct (open) during inspiration and adduct (close) to protect the airway and for voice production (phonation). The main nerve supply of the vocal cords comes from the recurrent laryngeal nerves (branches of the vagus nerve) which arise in the neck, but on the left side passes down around the aortic arch and then ascends in the tracheo-oesophageal groove to the larynx.

Normal vocal cords in phonation vibrate between 90 (male) and 180 (female) times per second, giving the voice its pitch or frequency. A healthy voice requires full closure of the vocal cords with a smooth, regular pattern of vibration, and any pathology that prevents full closure will result in air escaping between the vocal cords during phonation and a 'breathy' voice.

EXAMINATION

Good illumination is essential. Look at teeth, gums, tongue, floor of mouth and oral cavity. Tonsils, soft palate and uvula are easily seen, and a gag reflex (see p. 1110) is present. The remainder of the pharynx and larynx can be inspected with a laryngeal mirror or endoscope.

Examination of the neck for lymph nodes and other masses is also performed.

COMMON DISORDERS

Hoarseness (dysphonia)

There are three essential components for voice production: an air source (the lungs); a vibratory source (the vocal cords); and a resonating chamber (the pharynx, the nasal and oral cavities). Although chest and nasal disorders can affect the voice, the majority of hoarseness is due to laryngeal pathology.

Inflammation which increases the 'mass' of the vocal cords will cause the vocal cord frequency to fall, giving a much deeper voice. Thus listening to a patient's voice can often give a diagnosis before the vocal cords are examined.

Nodules

Nodules (always bilateral and commoner in females) and polyps (Fig. 20.7) are found on the free edge of the vocal cord preventing full closure and giving a 'breathy, harsh' voice. They are commonly found in professions that rely on their voice for their livelihood, such as teachers, singers and lawyers. They are usually related to poor technique of voice production and can usually be cured with speech therapy. If surgery is needed, great care must be taken to remain in the superficial layers of the vocal cord in order to prevent deep scarring which may leave the voice permanently hoarse.

Reinke's oedema (Fig. 20.8)

This is due to a collection of tissue fluid in the subepithelial layer of the vocal cord. The vocal cord has poor lymphatic drainage, predisposing it to oedema. Reinke's oedema is associated with irritation of the vocal cords: smoking, voice abuse, acid reflux and very rarely hypothyroidism. Treatment is to remove the irritation in most cases but surgery to thin the cords will also allow the voice to return to its normal pitch.

Acute-onset hoarseness

Hoarseness, in a smoker, is a danger sign. Any patient with a hoarse voice for over 6 weeks should be seen by an ENT surgeon. The voice may be deep, harsh and breathy indicating a mass on the vocal cord or it can be weak suggesting a paralysed left vocal cord secondary to mediastinal disease, e.g. bronchial carcinoma.

Early squamous cell carcinoma of the larynx (Fig. 20.9) has a good prognosis. Treatment is with carbon dioxide laser resection or radiotherapy. Spread of the tumour can lead to referred otalgia which may then require a laryngectomy with possible neck dissection. A patient with a paralysed left vocal cord must have a CT of the neck and chest. Medialization of the paralysed cord to allow contact with the opposite cord can return the voice and give a competent larynx. This can be done under local anaesthesia, giving an immediate result whatever the long-term prognosis of the chest pathology.

Fig. 20.7 **Vocal cord polyp.**

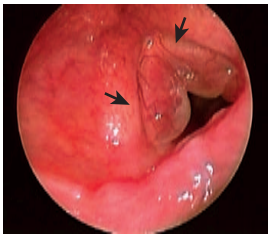


Fig. 20.8 **Reinke's oedema of the vocal cords.**

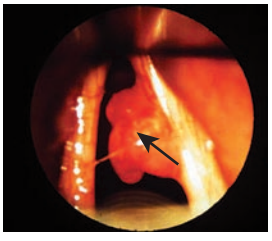


Fig. 20.9 **Carcinoma of the right cord.**

Stridor

Stridor or noisy breathing can be divided into inspiratory (source is glottic or above), expiratory (intrathoracic trachea or below) or mixed (subglottis or extrathoracic trachea). All patients with stridor, both paediatric and adult, are potentially at risk of asphyxiation and should be investigated fully. Severe stridor may be an indication for either intubation or a tracheostomy (Table 20.3).

Tracheostomy tubes may be:

- *Cuffed or uncuffed*. A high-volume, low-pressure cuff is used to prevent aspiration and to allow positive-pressure ventilation.
- *Fenestrated or unfenestrated*. This is a small hole on the greater curvature of the tube (both outer and inner) allowing air to escape upwards to the vocal cords and therefore the patient can speak. This tube often has a valve which allows air to enter from the stoma but closes on expiration, directing the air through the fenestration.

Most long-term tracheostomy tubes have an inner and outer tube. The inner tube fits inside the outer tube and projects beyond its lower end. A major problem with a tracheostomy tube is crusting of its distal end with dried secretions, and this arrangement allows the inner tube to be removed, cleaned and replaced as frequently as required, without disrupting the outer tube.

When to decannulate a patient is often a difficult issue if laryngeal competence is unclear. Movement of the vocal cords requires an ENT examination but owing to the risk of aspiration, a speech therapist's opinion can also be useful. The tracheostomy tube itself can also give problems due to compression of the oesophagus with a cuffed tube and by preventing the larynx from rising during normal swallowing.

Tonsillitis and pharyngitis

Viral infections of the throat are common and, although many practitioners are under pressure from the patient to give antibiotics, the vast majority are usually self-limiting, settling with bed rest, analgesia and encouraging fluid intake. Fungal infections, usually candidiasis, are uncommon and may indicate an immunocompromised patient or undiagnosed diabetes.

Tonsillitis

Tonsillitis, with a good history of pyrexia, dysphagia, lymphadenopathy and severe malaise is usually bacterial with β-haemolytic streptococcus the commonest organism.

Glandular fever (see p. 107)

This can also present with tonsillitis, although clinically the tonsils have a white exudate, there is often a petechial rash on the soft palate and an accompanying lymphadenopathy.

Table 20.3 Indications for tracheostomy

Upper airway obstruction (real or anticipated)
Long-term ventilation
Bronchial lavage
Incompetent larynx with aspiration

Quinsy

Quinsy is a collection of pus outside the capsule of the tonsil usually located adjacent to its superior pole. The patient often has trismus making examination difficult but the pus pushes the uvula across the midline to the opposite side. The area is usually hyperaemic and smooth but unilateral tonsil ulceration is more likely to be a malignancy. In either case urgent referral to an ENT specialist is essential.

Indications for a tonsillectomy are shown in Table 20.4. This is carried out under a general anaesthetic and current surgical techniques include diathermy dissection, laser excision and coblation (using an ultrasonic dissecting probe). There are strong advocates for each technique and much will depend on the individual surgeon's preference. Some departments now carry out tonsillectomy as a day case procedure as most reactionary bleeding will occur within the first 8 hours postoperatively.

Snoring

Snoring is due to vibration of soft tissue above the level of the larynx. It is a common condition (50% of 50-year-old males will snore to some extent) and can be considered to be related to obstruction of three potential areas: the nose, the palate or/and the hypopharynx (see Fig. 14.28). An Epworth questionnaire (Table 20.5) can also assist in the discrimination of sleep apnoea from simple snoring. Patients with a history of habitual, non-positional, heroic (can be heard through a wall) snoring require a full ENT examination and can be investigated by sleep nasendoscopy in which a sedated, snoring patient has a flexible nasendoscope inserted to identify the source of vibration. Nasal pathology such as polyps can be removed surgically with good results and most patients will benefit from lifestyle changes such as weight loss. Stiffening or shortening the soft palate via surgery, often using a laser, can help for palatal snorers but hypopharyngeal snorers require either a dental prosthesis at night to hold the mandible forward or continuous positive airway pressure (CPAP) via a mask (see p. 915).

Table 20.4 Indications for tonsillectomy

Suspected malignancy
Obstructive sleep apnoea due to tonsillar hypertrophy
Recurrent tonsillitis: five attacks a year for at least 2 years
Quinsy in a patient with a history of recurrent tonsillitis

Table 20.5 Epworth sleepiness scale

How likely are you to doze off or fall asleep in the following situations, in contrast to just feeling tired? This refers to your usual way of life in recent time. Even if you have not done some of these things recently, try to work out how they would have affected you. Use the following scale to choose the most appropriate number for each situation.

0 = would never doze
1 = slight chance of dozing
2 = moderate chance of dozing
3 = high chance of dozing

Situation	Chance of dozing
Sitting and reading	_____
Watching TV	_____
Sitting and inactive in a public place (theatre or meeting)	_____
As a passenger in a car for an hour without a break	_____
Lying down to rest in the afternoon when circumstances permit	_____
Sitting and talking to someone	_____
Sitting quietly after lunch (without alcohol)	_____
In a car, while stopped for a few minutes in the traffic	_____
TOTAL	_____

Normal 5 ± 4
Severe obstructive sleep apnoea 16 (±4)
Narcolepsy 17

Dysphagia (see p. 249)

Difficulty in swallowing is a common symptom but can be the presenting feature of carcinoma of the pharynx and therefore requires investigation.

A pharyngeal pouch is a herniation of mucosa through the fibres of the inferior pharyngeal constrictor muscle (cricopharyngeus) (Fig. 20.10). An area of weakness known as Killian's dehiscence allows a pulsion diverticulum to form. This will collect food which may regurgitate into the mouth or even down to the lungs at night with secondary pneumonia. Diagnosis is made radiologically and treatment is surgical, either via an external approach through the neck where the pouch is excised or more commonly endoscopically with stapling of the party wall (Fig. 20.10c).

Foreign bodies in the pharynx can be divided into three general categories: soft food bolus, coins (smooth), bones (sharp). Soft food bolus can be initially treated conservatively with muscle relaxants for 24 hours. Impacted coins should be removed at the earliest opportunity but sharp objects require emergency removal to avoid perforation of the muscle wall. If the patient perceives the foreign body to be to one side, then it should be above the cricopharyngeus and an ENT examination will locate it; common areas are the tonsillar fossae, base of tongue, posterior pharyngeal wall and valleculae. Radiology will identify coins, and it can be a clinical decision to see whether a coin will pass down to the stomach, in which case no further treatment is required as it will exit naturally. Some departments advocate the use of a metal detector to monitor the position of the coin in the patient, who is usually a child or has a mental disorder. Fish can be divided into those with a bony skeleton (teleosts) and those with a cartilaginous skeleton (elasmobranchs), and therefore radiology may only be useful in some cases. Radiology can also identify air in the cervical oesophagus indicating a radiolucent foreign body lying distally.

Globus pharyngeus

This is not a true dysphagia. It is a condition with classic symptoms of an intermittent sensation of a lump in the throat. This is perceived to be in the midline at the level of the cricoid cartilage and is worse when swallowing saliva; indeed it often disappears when ingesting food or liquids. ENT examination is clear and normal laryngeal mobility can be felt when gently rocking the larynx across the postcricoid tissues. A contrast swallow will show not only the structures below the pharynx but also assess the swallowing dynamically. Any suspicious area will require an endoscopy with biopsy.

FURTHER READING

Burton M. *Diseases of the Ear, Nose and Throat*, 15th edn. Edinburgh: Churchill Livingstone, 2000.

Corbridge R. *Essential ENT Practice. A Clinical Text*. London: Arnold, 1998.

Roland NJ, McRae RDR, McCombe AW. *Key Topics in Otolaryngology*, 2nd edn. London: Bios Scientific Publishers, 2000.

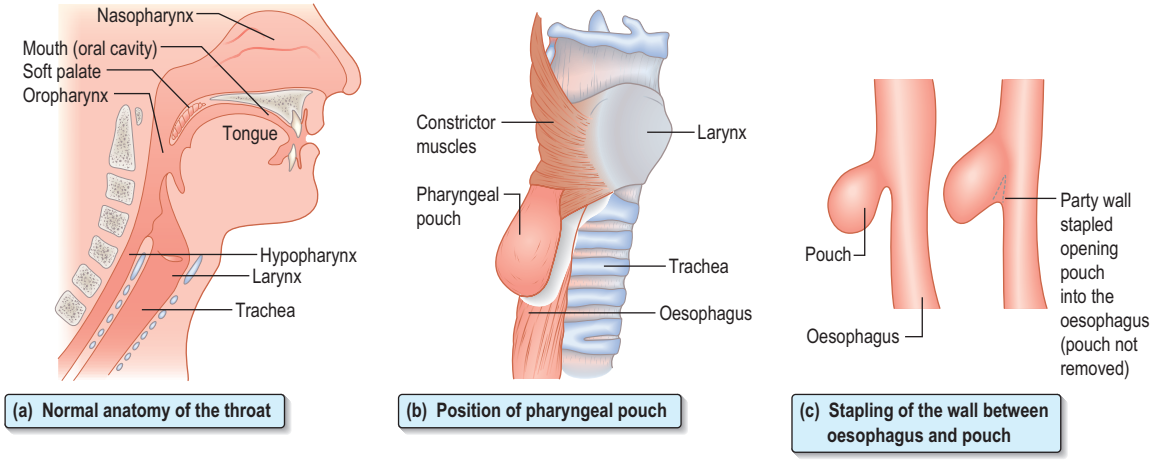


Fig. 20.10 **Pharyngeal pouch. (a)** Normal anatomy of the throat. **(b)** Position of pharyngeal pouch. **(c)** Stapling of the party wall between oesophagus and pouch.

DISORDERS OF THE EYE

A careful and detailed history gives most of the facts needed to make a working diagnosis. The eye has only a few mechanisms by which it can convey a diseased state. Common symptoms include alteration in visual acuity, redness, pain, discharge and photophobia. It is vital that an accurate Snellen visual acuity is recorded in all patients with an eye problem.

Diabetic eye disease (see p. 1050) and hypertensive eye disease (see p. 801) are discussed elsewhere.

APPLIED ANATOMY AND PHYSIOLOGY

The average diameter of the human eye is 24 mm. The *cornea* occupies the central aspect of the globe and is one of the most richly innervated tissues in the body. This clear, transparent and avascular structure, measuring 12 mm horizontally and 11 mm vertically, provides 78% of the focusing power of the eye. The endothelial cells lining the inner surface of the cornea are responsible for maintaining the clarity of the cornea by continuously pumping fluid out of the tissue. Any factor which alters the function of these cells will result in corneal oedema and cause blurred vision. The anterior surface of the cornea is lined by epithelial cells which, together with the tear film, provide a smooth and regular surface for the refraction of light. The eyelids prevent the cornea from drying and becoming an irregular surface by distributing the tear film over the surface of the globe with each blink. In addition, the lids protect the globe and provide an outflow channel for excess tears via the punctae to the nose (Fig. 20.11).

The cornea and the sclera give mechanical strength and shape to the exposed surface of the globe. The *sclera* is a white opaque structure covering four-fifths of the globe and is continuous with the cornea at the limbus. The six extraocular muscles responsible for eye movements are attached to the sclera and the optic nerve perforates it posteriorly.

The *conjunctiva* covers the anterior surface of the sclera. This richly vascularized and innervated mucous membrane stretches from the limbus over the anterior sclera and is then reflected onto the undersurface of the upper and lower lids. The area of conjunctival reflection under the lids makes up the upper and lower fornix. The conjunctiva over the sclera is referred to as the bulbar conjunctiva whilst that over the lids makes up the tarsal conjunctiva.

The *anterior chamber* is the space between the cornea and the iris, and is filled with aqueous humour (Fig. 20.12). This fluid is produced by the ciliary body (2 µL/min) and provides nutrients and oxygen to the avascular cornea. The outflow of aqueous humour is through the trabecular

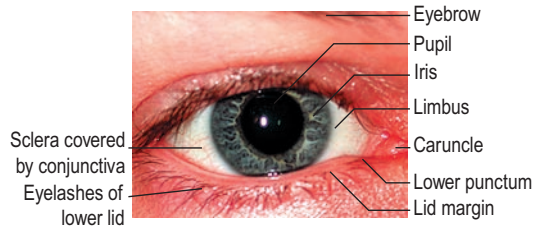


Fig. 20.11 **Photograph of the right eye.**

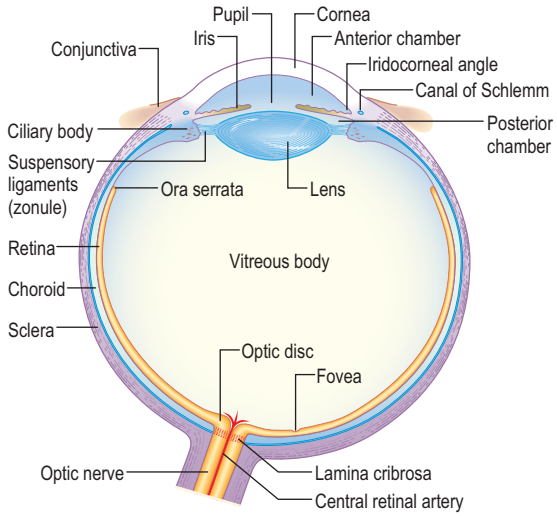


Fig. 20.12 **Section of the eyeball.**

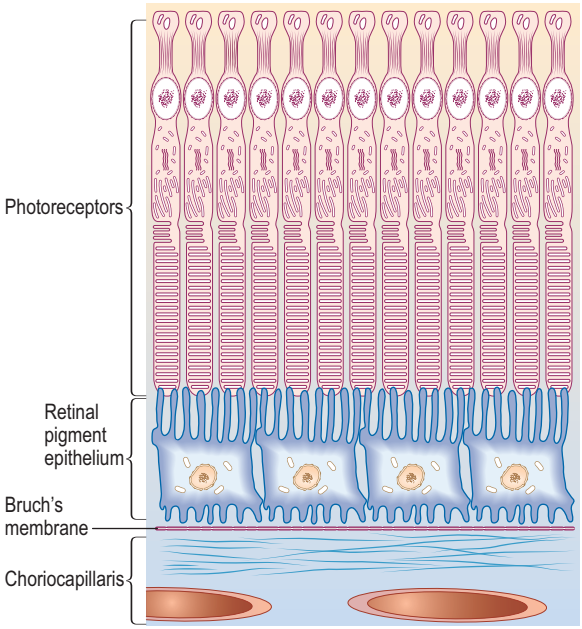


Fig. 20.13 **Schematic diagram of the retina.**

meshwork and canal of Schlemm adjacent to the limbus. Any factor which impedes its outflow will increase the intraocular pressure. The upper range of normal for intraocular pressure is 21 mmHg.

The *iris* is the coloured part of the eye under the transparent cornea. The muscles of the iris diaphragm regulate the size of the pupil, thereby controlling the amount of light entering the eye. Immediately posterior to the pupil and anterior to the vitreous humour lies the *lens* of the eye. This is a transparent biconvex structure and is responsible for 22% of the refractive power of the eye. By changing its shape it can alter its refractive power and help to focus objects at different distances from the eye. In the fourth decade of life this ability to change shape starts to decline (presbyopia) and the need for reading glasses becomes inevitable. With age the lens starts to become less transparent and cataracts begin to develop.

The inner aspect of the *sclera* is lined by a highly vascular tissue, the choroid, and then the retina. The vitreous humour fills the cavity between the retina and the lens.

The *retina* is a multi-layered structure. The metabolically active region of the retina is represented in Figure 20.13. There are two types of photoreceptors in the retina, rods and cones. There are approximately 6 million cones mainly confined to the *macula* and are responsible for detailed central vision and colour vision. The peripheral retina has around 125 million rods that are responsible for peripheral vision. The axons of the ganglion cells form the optic nerve (or disc) of the eye (Fig. 20.14).

The *blood supply* to the eye is via the ophthalmic artery and, in particular, the central retinal artery is responsible for supplying the inner retinal layers. Venous return is through the central retinal and ophthalmic veins. Local lymphatic drainage is to the pre-auricular and submental nodes.

The sensory *innervation* of the eye (see p. 1106) is through the trigeminal (V) nerve. Of the six extraocular muscles the abducens (VI) nerve supplies the lateral rectus, the trochlear (IV) nerve supplies the superior oblique, whilst the rest are supplied by the oculomotor (III) nerve. The III nerve also supplies the upper lid and indirectly the pupil (parasympathetic fibres are attached to it). The facial (VI) nerve supplies the obicularis and other muscles of facial expression.

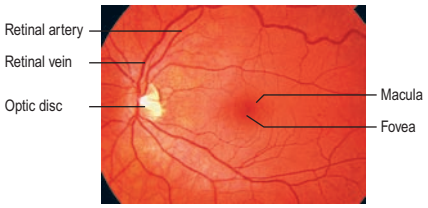


Fig. 20.14 **Fundus of the left eye.**

REFRACTIVE ERRORS

The eye projects a sharp and focused image onto the retina. Refractive errors refer to any abnormality in the focusing mechanism of the eye and not to any opacity in the system such as a corneal or retinal scar.

The refraction of light in emmetropic (normal), myopic and hypermetropic eyes is shown in Figure 20.15.

Astigmatism occurs when the eye is not of the same radius of curvature for refraction. It may be myopic in one plane and hypermetropic or emmetropic in the other plane. In this situation the front surface of the eye is more rugby ball shaped than football shaped.

Presbyopia is the term used to describe the normal ageing of the lens and leads to a change in the refractive state of the eye. As the lens ages it becomes less able to alter its curvature and this causes difficulty with near vision, especially reading.

Treatment

Errors of refraction can be corrected by using spectacles or contact lenses. The latter result in better quality, but carry the risk of infection. They may be the only option in some refractive states such as keratoconus. There are a number

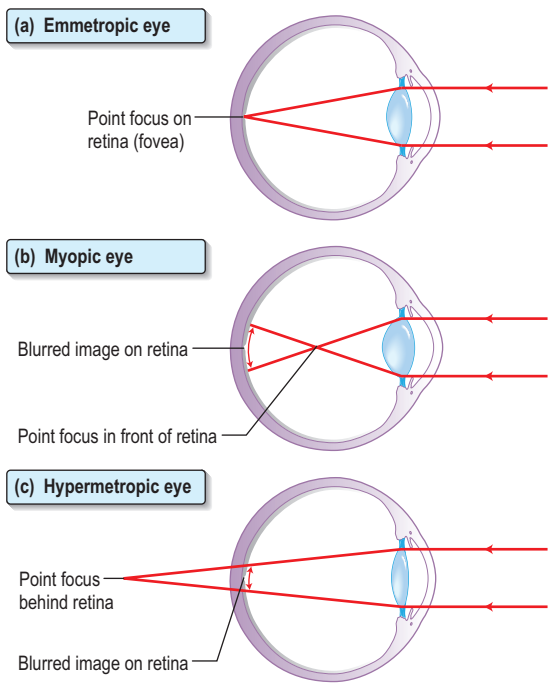
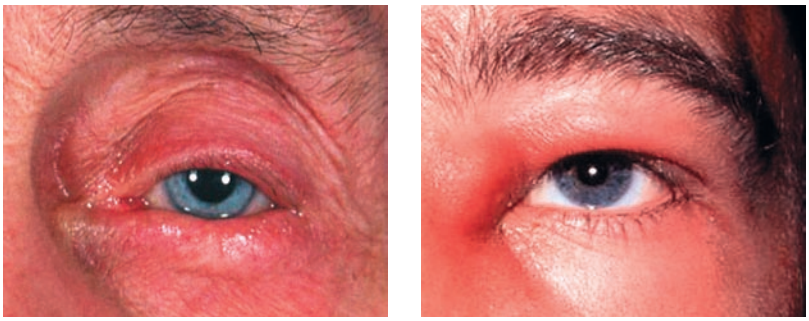
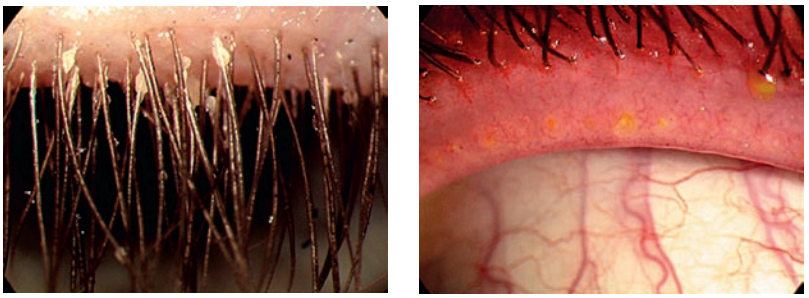


Fig. 20.15 **Refraction. (a)** Emmetropic (normal). **(b)** Myopic (short-sighted). **(c)** Hypermetropic (long-sighted).



(a) (b)

Fig. 20.16 **Disorders of the eyelid. (a)** Lid entropion. The lower lid appears inverted. **(b)** Acute dacryocystitis showing a lump on the side of the nose.



(a) (b)

Fig. 20.17 **Blepharitis. (a)** Crusty and scaly deposits on the lashes and lash bases. **(b)** Upper lid showing meibomian glands plugged with oily secretions.

of surgical techniques available with varying degrees of accuracy. The most popular method is to use an excimer laser to re-profile the corneal curvature (PRK, LASIK, LASEK). Here the laser either removes corneal tissue centrally to flatten the cornea in myopia or it removes tissue from the peripheral cornea to steepen it in hypermetropia.

DISORDERS OF THE LIDS

The lids afford protection to the eyes and help to distribute the tear film over the front surface of the globe. Excess tears are drained via the punctae and lacrimal system. Malposition of the lids, factors which affect blinking or lacrimal drainage, can all cause problems.

Entropion. The lid margin rolls inwards so that the lashes are against the globe (Fig. 20.16a). The lashes act as a foreign body and cause irritation, leading to a red eye which can mimic conjunctivitis. Occasionally the constant rubbing of lashes against the cornea causes an abrasion. The commonest cause is ageing and surgery is usually required.

Ectropion. The lid margin is not apposed to the globe. As a result the lacrimal puncta is not in the correct anatomical position to drain tears and patients usually complain of a watery eye. Underlying factors include age, VII nerve palsy and cicatricial skin conditions. Surgery is usually required.

Dacryocystitis. This is inflammation of the lacrimal sac. Patients usually present with a painful lump at the side of the nose adjacent to the lower lid (Fig. 20.16b). This should be treated with oral broad-spectrum antibiotics such as cefalexin, and patients should be watched carefully for signs of cellulitis. These patients should be referred to the ophthalmologist as some may have an underlying mucocele or dilated sac, and will require surgery.

Blepharitis. This is inflammation of the lid margins. Inflammation may involve the lashes and lash follicles (Fig. 20.17a) resulting in styes, or inflammation and blockage of meibomian glands (Fig. 20.17b) leading to chalazion (Fig. 20.18). Treatment involves lid toilet and topical antibiotics such as chloramphenicol or fusidic acid. If there is associated cellulitis then broad-spectrum oral antibiotics are required. Some patients are left with a lump once the acute inflammatory phase of the chalazion has subsided. Most of these patients find the lump cosmetically unacceptable and require incision and curettage.

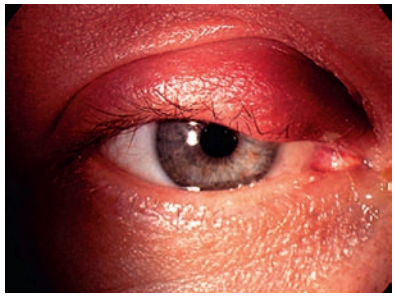


Fig. 20.18 **Blepharitis.** Blockage of the meibomian glands leads to swelling of the lid (chalazion).

CONJUNCTIVITIS

This is the commonest cause of a red eye. Inflammation of the conjunctiva leading to a red eye can arise from a number of causes, with viral, bacterial and allergic being the commonest. Common features in all types of conjunctivitis include soreness, redness and discharge, and in general the visual acuity is good. History should include the speed of onset of the inflammation, the colour and consistency of the discharge, whether the eye is itchy, and if there has been a recent history of a cold or sore throat. In the neonate it is vital to ensure there has been no associated maternal sexually transmitted infection. The differential diagnosis of conjunctivitis is shown in Table 20.6.

Table 20.6 Conjunctivitis

	Discharge	Preauricular node	Corneal involvement	Comment
Bacterial	Mucopurulent	−ve except gonococci	+ve Gonococcus	Rapid onset
Viral	Watery	+ve	+ve Adenovirus	Cold and/or sore throat
Chlamydial	Watery	+ve	+ve	GU discharge
Allergic	Stringy	−ve	+ve	Itchy

Bacterial conjunctivitis

Bacterial conjunctivitis is uncommon, making up 5% of all cases of conjunctivitis. In the vast majority of patients it causes a sore or gritty eye in the presence of good vision. Bacterial conjunctivitis is invariably bilateral and should be suspected when conjunctival inflammation is associated with a purulent discharge. The rapidity of onset together with the severity of discharge is helpful in determining the underlying organism.

Clinical features

Gonococcal conjunctivitis should be suspected when the onset of symptoms is rapid, the discharge is copious, and ocular inflammation includes chemosis (conjunctival oedema) and lid oedema. Gonococci are a cause of conjunctivitis giving rise to a pre-auricular node.

Less acute or subacute purulent conjunctivitis with moderate discharge can be attributed to organisms such as *Haemophilus influenzae* and *Streptococcus pneumoniae*. Chronic conjunctivitis is usually associated with mild conjunctival injection and scant purulent discharge. Common organisms include *Staphylococcus aureus* and *Moraxella lacunata*.

Treatment

Prompt treatment with oral and topical penicillin is given in gonococcal conjunctivitis to ensure a reduced rate of corneal perforation. A Gram stain of the conjunctival swab can quickly confirm the presence of diplococci. Gonococcal conjunctivitis is a notifiable disease in the UK.

Empirical treatment for both subacute and chronic conjunctivitis involves a topical broad-spectrum antibiotic such as chloramphenicol. Swabs should be taken if these cases do not respond to this initial treatment.

Chlamydial conjunctivitis

This is a sexually transmitted infection and is most prevalent in sexually active adolescents and young adults. Direct or indirect contact with genital secretions is the usual route of infections but shared eye cosmetics can also be involved. It is caused by *Chlamydia trachomatis* (see p. 176). Neonatal chlamydial conjunctivitis is a notifiable disease in the UK and should be suspected in newborns with a red eye. Mothers should be asked about sexually transmitted infections. Trachoma caused by the same organism is found mainly in the tropics and the Middle East and is a very common cause of blindness in the world (see p. 143).

Clinical features

The onset of symptoms is slow, and patients may complain of mild discomfort for weeks. In these cases the red eye is associated with a scanty mucopurulent discharge and a palpable preauricular lymph node. In chronic cases it is not unusual to see superior corneal vascularization. In neonates the onset of the red eye is typically around 2 weeks after birth, whereas gonococcal conjunctivitis occurs within days of birth. Conjunctival swabs should be taken prior to commencement of treatment.

Treatment

Topical erythromycin twice daily is commenced and patients referred to the genitourinary physician. Neonates should be started on topical erythromycin and referred to the paediatrician as there may be associated otitis media or pneumonitis.

Viral conjunctivitis

Adenoviral conjunctivitis

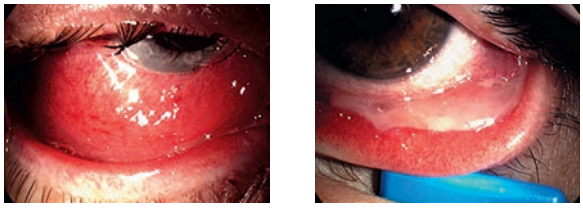
This is highly contagious and can cause epidemics in communities. Transmission is through direct or indirect contact with infected individuals. The onset of symptoms may be preceded by a cold or flu-like symptoms. The eye becomes inflamed and this is commonly associated with chemosis, lid oedema and a palpable preauricular lymph node. Some patients may develop a membrane on the tarsal conjunctiva (Fig. 20.19a,b) and haemorrhage on the bulbar conjunctiva. Viral conjunctivitis can cause deterioration in visual acuity owing to corneal involvement (focal areas of inflammation). In 50% of the patients the conjunctivitis is unilateral.

Treatment

The condition is largely self-limiting in the majority of cases. Lubricants together with a cold compress can be soothing for patients. Strict hygiene and keeping towels separate from the rest of the household goes a long way towards reducing the spread of the infection. Clinicians also need to ensure good hygiene practice to reduce cross-infection and infecting themselves. In patients with corneal involvement or intense conjunctival inflammation, topical steroids are indicated.

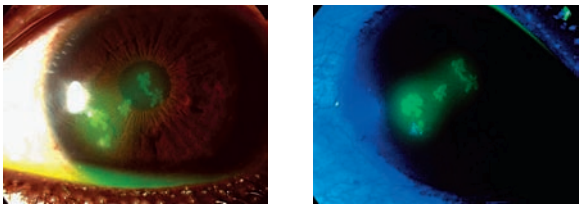
Herpes simplex conjunctivitis

Primary ocular herpes simplex conjunctivitis is typically unilateral. It usually causes a palpable preauricular lymph node



(a) (b)

Fig. 20.19 **(a and b) Viral conjunctivitis; (b) with pseudomembrane on lower lid.**



(a) (b)

Fig. 20.20 **Dendritic ulcers on the cornea. (a)** Stained with fluorescein. **(b)** Stained with fluorescein and viewed with blue light.

and cutaneous vesicles develop on the eyelids and the skin around the eyes in the majority of patients. Over 50% of these patients may develop a dendritic corneal ulcer (Fig. 20.20). The organism responsible for this condition is the herpes simplex virus (HSV), which is usually HSV-1 but HSV-2 can give rise to ocular infection.

Treatment

Primary ocular HSV infection is a self-limiting condition but most clinicians choose to treat it with topical aciclovir in order to limit the risk of corneal epithelial involvement.

Molluscum contagiosum conjunctivitis

This is typically unilateral and produces a red eye that generally goes unrecognized and comes to the forefront because patients fail to improve and the cornea starts to become involved. A closer look at the eyelids and the margin will reveal pearly umbilicated nodules and these are filled with the DNA poxvirus. A high index of suspicion is needed to make an early diagnosis.

Treatment

This includes curetting the central portion of the lesion, freezing the centre or completely excising the lesion. If the corneal involvement is severe or the eye is very inflamed, a short course of topical steroids such as prednisolone 0.5% or dexamethasone 0.1% is helpful.

Allergic conjunctivitis

There are five main types of allergic conjunctivitis: seasonal, perennial, vernal, atopic and giant papillary. The last three are difficult to treat, chronic and can be sight-threatening. They should be referred to an ophthalmologist.

Seasonal/perennial conjunctivitis

Seasonal allergic conjunctivitis and perennial conjunctivitis affect 20% of the general population in the UK. Seasonal allergic conjunctivitis involves an allergic reaction to grass, pollen and fungal spores and occurs mainly in spring and summer. Perennial allergic conjunctivitis occurs all year round but peaks in the autumn and includes allergens such as house-dust mites.

The main symptoms include itching, redness, soreness, watering and a stringy discharge. Occasionally the conjunctiva may become so hyperaemic that chemosis results. This is usually associated with swollen lids.

Treatment

Reducing the allergen load (reducing dust, p. 832) is helpful. Medical treatment includes the use of antihistamine drops such as azelastine and emedastine together with topical mast cell-stabilizing agents such as sodium cromoglicate and nedocromil. Olopatadine (bd) has dual action and is very effective. Corticosteroid drops should be avoided. Oral antihistamines help the itching.

CORNEAL DISORDERS

Trauma

Corneal abrasions

Trauma resulting in the removal of a focal area of epithelium on the cornea is very common. Abrasions usually occur when the eye is accidentally poked with a finger, a foreign body (FB) flies into the eye or something brushes against the eye.

Clinical features

Symptoms include severe pain, due to exposure of the corneal nerve endings, lacrimation and inability to open the eye (blepharospasm). Blinking and eye movement can aggravate the pain and foreign body sensation. The visual acuity is usually reduced. Most cases will need topical anaesthetic drops such as oxybuprocaine or tetracaine to be administered before it is possible to examine the eye. The cornea should be inspected with a blue light after instillation of fluorescein drops. The orange dye will stain the area of the abrasion. Under blue light the abrasion lights up as green. Occasionally FBs can lodge on the under surface of the upper lid and give rise to linear vertical abrasions. Eversion of the upper lid is necessary in all cases of abrasions (Fig. 20.21).

Treatment

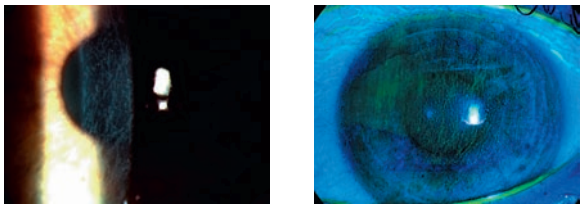
This involves instillation of a broad-spectrum antibiotic such as chloramphenicol drops or ointment four times a day for 5 days. The role of padding is controversial but common practice is to pad the affected eye for 24 hours once chloramphenicol ointment has been applied to the eye.

Corneal foreign body

Occasionally when something flies into the eye it gets stuck on the cornea (Fig. 20.22a). It may be associated with lacrimation and photophobia. Examination is best attempted following instillation of a topical anaesthetic and should include everting the upper lid (Fig. 20.22b). Corneal foreign bodies can usually be seen directly with a white light.

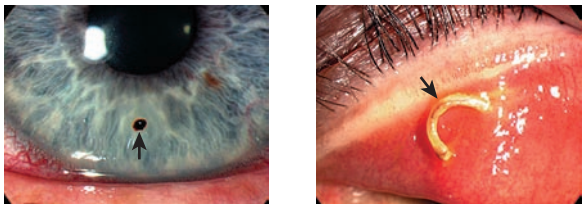
Treatment

The corneal FB should be removed and the patient given a topical antibiotic such as chloramphenicol four times a day for 5 days or fusidic acid twice a day for 5 days.



(a) (b)

Fig. 20.21 **Linear corneal abrasion. (a)** Stained with fluorescein. **(b)** Stained with fluorescein and viewed with blue light.



(a) (b)

Fig. 20.22 **Foreign bodies. (a)** Corneal. **(b)** Subtarsal.

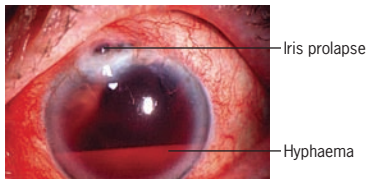


Fig. 20.23 **Blunt trauma.** Hyphaema and globe rupture with iris prolapse.

Trauma

In cases of high-velocity trauma corneal perforation or intraocular FB should be suspected. Examination may show a corneal laceration with or without the FB embedded in the cornea. The FB may be present on the iris or the lens. Other clues of a penetrating injury include a flat anterior chamber or the presence of blood in the anterior chamber (hyphaema) or a large subconjunctival haemorrhage. Urgent referral to the ophthalmologist is mandatory, ensuring that no drops are instilled into the eye and that a plastic shield has been placed over the eye to minimize further risk of trauma.

Blunt trauma usually results in periorbital bruising and gross lid oedema, which can make examination to exclude perforating injury difficult. These patients should be referred to the ophthalmologist for a detailed ocular examination to exclude a perforation retinal detachment or a traumatic hyphaema (Fig. 20.23).

Keratitis

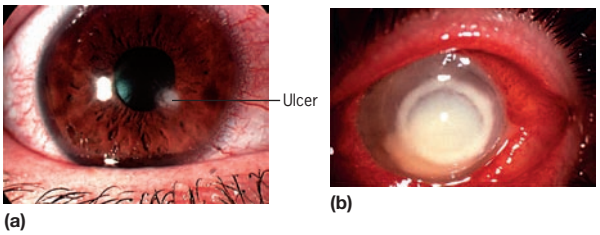
This is a general term to describe corneal inflammation. Common causes include herpes simplex virus, contact lens-associated infection and blepharitis. Symptoms include the sensation of a foreign body or pain (depending on the size and depth of the ulcer), photophobia and lacrimation. Vision is reduced if the ulcer is in the visual axis.

Herpes simplex keratitis

Corneal epithelial cells infected with the virus eventually undergo lysis and form an ulcer which is typically dendritic in shape (Fig. 20.20). The ulcer stains with fluorescein and can be observed easily with a blue light. Topical immunosuppression, e.g. steroid drops, or systemic immunosuppression, e.g. AIDS, can lead to the centrifugal spread of the virus such that the ulcer increases in area and is referred to as a geographic ulcer. Recurrent attacks of HSV keratitis can be triggered by ultraviolet light, stress and menstruation. All these factors are responsible for activating the virus, which normally lies dormant in the ganglion of the V nerve.

Treatment

Aciclovir ointment five times a day for 2 weeks is usually very effective.



(a) (b)

Fig. 20.24 **Contact lens-related keratitis. (a)** Corneal ulcer. **(b)** Severe keratitis with a corneal abscess and a hypopyon.

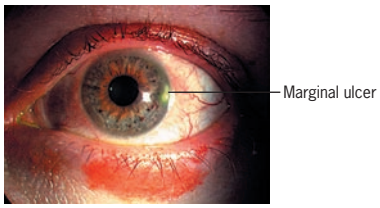


Fig. 20.25 **Marginal keratitis.**

Contact lens-related keratitis

A small number of contact lens wearers develop infective corneal ulcers which are potentially sight-threatening (Fig. 20.24a,b). The organisms usually responsible include Gram-positive and Gram-negative bacteria. Patients should be referred to an ophthalmologist for scraping of the ulcer and commencement of antibiotic treatment.

Blepharitis

This is an extremely common condition where the lid margins are inflamed. Common underlying causes of blepharitis include meibomian gland dysfunction, seborrhoea and *Staphylococcus aureus* infection. Patients can be asymptomatic or complain of itchy, burning eyes because of tear film instability resulting from meibomian gland dysfunction. *Staphylococcus aureus* is frequently responsible for chronic blepharo-conjunctivitis and some patients may develop keratitis (Fig. 20.25).

Treatment

Lid hygiene is the mainstay of treatment as it helps to reduce the bacterial load and unblock meibomian glands. A short course of topical chloramphenicol is useful in chronic cases but in severe cases or cases where acne rosacea is suspected, oral doxycycline may be required. Patients with keratitis should be referred to the ophthalmologist as they will require topical steroids.

CATARACTS

Cataract (Fig. 20.26a,b) is by far the commonest cause of preventable blindness in the world with an effective surgical treatment. In the UK approximately 250 000 cataract operations are performed each year, making it the commonest surgical procedure.

Aetiology

Age-related opacification of the lens (cataract) is the commonest cause of visual impairment with 30% of people over 65 years having visual acuities below that required for driving (Snellen acuity less than 6/12). The common causes of cataracts are summarized in Table 20.7.

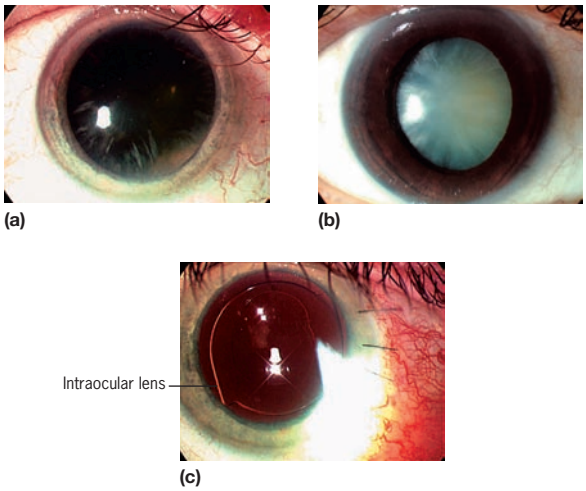


Fig. 20.26 **Cataracts. (a)** Early cataract. **(b)** White mature cataract. **(c)** Artificial intraocular lens following phacoemulsification.

Table 20.7 Cataracts: aetiology

Congenital	Maternal infection
	Familial
Age	Elderly
Metabolic	Diabetes, galactosaemia, hypocalcaemia, Wilson's disease
Drug-induced	Corticosteroids, phenothiazines, miotics, amiodarone
Traumatic	Post-intraocular surgery
Inflammatory	Uveitis
Disease associated	Down's syndrome (see p. 40)
	Dystrophia myotonica (see p. 1181)
	Lowe's syndrome (see p. 1070)

In young patients, familial or congenital causes should be excluded. Any history of ocular inflammation is noted. Cataracts diagnosed in infants demand urgent referral to the ophthalmologist in order to minimize the subsequent development of amblyopia.

Clinical features

Gradual painless deterioration of vision is the commonest symptom. Other symptoms are dependent upon the type of cataract, for example a posterior capsular type would lead to glare and problems with night driving. Early changes in the lens are correctable by spectacles but eventually the opacification needs surgical intervention.

Investigations

Blood glucose, serum calcium and liver biochemistry should be measured to diagnose metabolic disorders.

Treatment

Cataract extraction with the insertion of an intraocular lens is the treatment of choice (Fig. 20.26c). The aetiology and density of the cataracts usually determines the technique employed. Currently, small incision surgery called phacoemulsification is the most popular technique used for routine cases.

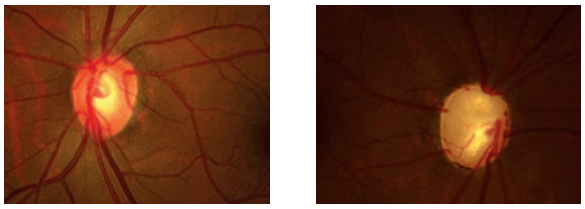


Fig. 20.27 **(a) Normal optic disc and (b) glaucomatous optic disc.** The central cup is enlarged and deepened.

GLAUCOMA

This refers to a group of diseases where the pressure inside the eye is sufficiently elevated to cause optic nerve damage and result in visual field defects (Fig. 20.27). Normal intraocular pressure (IOP) is between 10 and 21 mmHg. Some types of glaucoma can result in an IOP exceeding 70 mmHg. Glaucoma as a disease entity is the second commonest cause of blindness world-wide and the third commonest cause of blind registration in the UK.

Primary open-angle glaucoma (POAG)

This is the commonest form of glaucoma. High intraocular pressures result from reduced outflow of aqueous humour through the trabecular meshwork. The cause of this increased resistance to aqueous outflow at the level of the meshwork is not fully understood. Common risk factors include age (0.02% of 40-year-olds vs 10% of 80-year-olds), race (black Africans are at five times greater risk than whites), positive family history and myopia.

POAG causes a gradual, insidious, painless loss of peripheral visual field. The central vision remains good until the end-stage of the disease. Diagnosis is only made if the IOP is measured. The optic disc is inspected and shows an enlarged cup with a thin neuroretinal rim. Visual fields are performed and show a normal blind spot with scotomas. Most patients are identified as having glaucoma whilst undergoing a routine ophthalmic examination.

Treatment

Treatment aims to reduce the IOP and this is achieved either by reducing aqueous production or increasing aqueous drainage. Beta-blockers such as timolol, carteolol and levobunolol reduce aqueous production and are the commonest prescribed topical agents. These drugs are contraindicated in patients with COPD, asthma or heart block. Prostaglandin analogues such as latanoprost and travoprost, increase aqueous outflow and are also used (alone or in combination with beta-blockers) for POAG as they can reduce IOP by 30%. Patients should be monitored as the irises can become brown. Carbonic anhydrase inhibitors such as acetazolamide reduce aqueous production and are available in both topical and oral form. In its oral form acetazolamide is the most potent drug for reducing ocular pressure. It should not be used in patients with sulphonamide allergy.

Acute angle-closure glaucoma (AACG)

This is an ophthalmic emergency. In this type of glaucoma there is a sudden rise in intraocular pressure to levels greater

Table 20.8 Differential diagnosis of the acute red eye

Cause	Conjunctival injection	Unilateral or bilateral	Pain	Photophobia	Vision	Pupil	Intraocular pressure
Conjunctivitis	Diffuse	Bilateral	Gritty	Occasionally with adenovirus	Normal	Normal	Normal
Anterior uveitis	Circum-corneal	Unilateral	Painful	Yes	Reduced	Constricted	Normal or raised
Acute glaucoma	Diffuse	Unilateral	Severe pain	Mild	Reduced	Mid-dilated	Raised

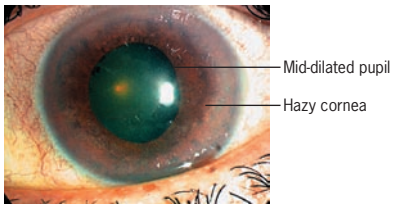


Fig. 20.28 **Acute angle-closure glaucoma.**

than 50 mmHg. This occurs due to reduced aqueous drainage as a result of the ageing lens pushing the iris forward against the trabecular meshwork. People most at risk of developing AACG are those with shallow anterior chambers such as hypermetropes and women. The attack is more likely to occur under reduced light conditions when the pupil is dilated.

AACG causes sudden onset of a red painful eye and blurred vision. Patients become unwell with nausea and vomiting and complain of headache and severe ocular pain. The eye is injected, tender and feels hard. The cornea is hazy and the pupil is semi-dilated (Fig. 20.28). Table 20.8 shows the differential diagnosis of the acute red eye. Emergency Box 20.1 shows features that require urgent referral to an ophthalmologist.

Prompt treatment is required to preserve sight and includes i.v. acetazolamide 500 mg (provided there are no contraindications) to reduce IOP, and instillation of pilocarpine 4% drops to constrict the pupil to improve aqueous outflow and prevent iris adhesion to the trabecular meshwork. Other topical drops such as beta-blockers and prostaglandin analogues can also be instilled if available, provided there are no contraindications. Analgesia and antiemetics are given as required.

These patients must be referred to an ophthalmologist immediately so that reduction in IOP can be monitored and other agents such as oral glycerol or i.v. mannitol can be administered to non-responding patients. Definitive treatment involves making a hole in the periphery of the iris of both eyes either by laser or surgically.

Emergency Box 20.1

Red Flags for a red eye

The following symptoms require urgent referral:
- severe pain
- photophobia
- reduced vision
- coloured halos around point of light in a patient's vision
- proptosis
- smaller pupil in affected eye

+

On medical assessment:
- high intraocular pressure
- corneal epithelial disruption
- shallow anterior chamber depth
- ciliary flush

UVEITIS

Uveitis is inflammation of the uveal tract, which includes the iris, ciliary body and choroid. Inflammation confined to the anterior segment of the eye (in front of the iris) is referred to as iritis or anterior uveitis, that involving the ciliary body is referred to as intermediate uveitis whilst inflammation of the choroid is termed posterior uveitis. If all three regions are involved then the term pan-uveitis is used.

The most common symptoms of uveitis are blurred vision, pain, redness, photophobia and floaters. Each symptom is determined by the location of the inflammation such that photophobia and pain are common features of iritis whilst floaters are commonly seen with posterior uveitis.

Uveitis is commonly encountered with ankylosing spondylitis and positive HLA-B27 (see p. 533), arthritis, inflammatory bowel disease, sarcoid, tuberculosis, syphilis, toxoplasmosis, Behçet's syndrome, lymphoma and viruses such as herpes, cytomegalovirus and HIV. In a number of patients no cause is found (idiopathic uveitis).

Anterior uveitis (iritis)

The classic presentation entails a triad of eye symptoms: redness, pain and photophobia. Vision can be normal or blurred depending on the degree of inflammation. The eye can be generally red or the injection can be localized to the limbus. The anterior chamber shows features consistent with inflammation including cells with keratic precipitates (KP) on the corneal endothelium, fibrin or hypopyon (pus), and the pupil may have adhered to the lens (posterior synechiae) (Fig. 20.29). The IOP may be normal or raised either due to cells

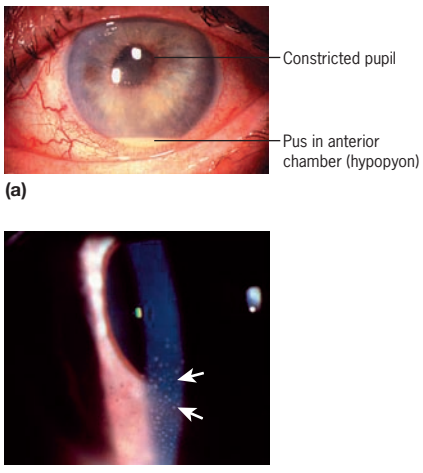


Fig. 20.29 **Anterior uveitis. (a)** Acute with hypopyon. **(b)** Showing keratic precipitates on the corneal endothelium.

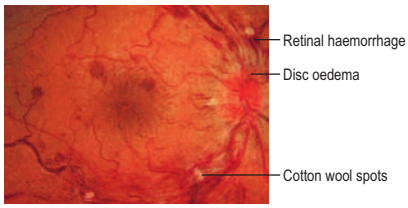


Fig. 20.30 **Central retinal vein occlusion.**

clogging up the trabecular meshwork or posterior synechiae causing aqueous to build up behind the iris and force the iris against the trabecular meshwork and so reduce aqueous drainage.

Treatment

This consists of reducing inflammation with the use of topical steroids such as dexamethasone 0.1% and dilating the pupil with cyclopentolate 1% to prevent formation of posterior synechiae. Dilatation also allows fundoscopy to exclude posterior segment involvement. If the IOP is raised, this is treated with either topical beta-blockers, prostaglandin analogues, or oral or i.v. acetazolamide. These patients should be referred to the ophthalmologist.

DISORDERS OF THE RETINA

Central retinal vein occlusion (CRVO)

This usually leads to profound sudden painless loss of vision with thrombosis of the central retinal vein at or posterior to the lamina cribrosa where the optic nerve exits the globe. The thrombus causes obstruction to the outflow of blood leading to a rise in intravascular pressure. This results in dilated veins, retinal haemorrhage, cotton wool spots, and abnormal leakage of fluid from vessels resulting in retinal oedema (Fig. 20.30). In severe cases an afferent papillary defect is usually present (see p. 1103).

Predisposing factors include increasing age, hypertension and cardiovascular disease, diabetes, glaucoma, blood dyscrasias and vasculitis.

Treatment

Treatment of any underlying medical condition is mandatory. Referral to an ophthalmologist is essential to monitor the eye, as some patients can develop retinal ischaemia with resulting neovascularization of the retina and iris. Patients who develop iris neovascularization, rubeosis, are at risk of developing rubeotic glaucoma.

Central retinal artery occlusion (CRAO)

This results in sudden painless severe loss of vision. Retinal arterial occlusion results in infarction of the inner two-thirds of the retina. The arteries become narrow and the retina becomes opaque and oedematous. A cherry red spot is seen at the fovea because the choroid shows up through the thinnest part of the retina (Fig. 20.31). An afferent papillary defect is usually present (see p. 1103).

Arteriosclerosis-related thrombosis is the most common cause of CRAO. Emboli from atheromas and diseased heart valves are other causes. Giant cell arteritis (see p. 1138) must be excluded.

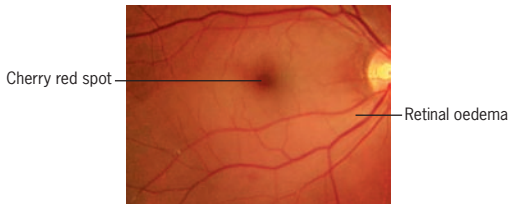


Fig. 20.31 **Central retinal artery occlusion.**

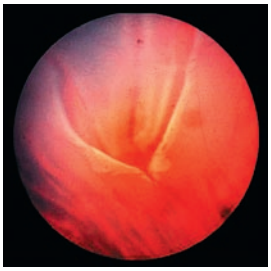


Fig. 20.32 **Retinal tear leading to detachment.**

Treatment

CRAO is an ophthalmic emergency since studies have shown that irreversible retinal damage occurs after 90 minutes of onset. Ocular massage and 500 mg i.v. acetazolamide help to reduce ocular pressure and may help in dislodging the emboli. Breathing into a paper bag allows a build-up of carbon dioxide which acts as a vasodilator and so helps in dislodging the emboli. Other options include making a corneal paracentesis to drain off some aqueous humour, thereby reducing the intraocular pressure.

Patients with CRAO should have a thorough medical evaluation to determine the aetiology of the emboli or thrombus. Some patients may present with transient loss of vision or amaurosis fugax (see p. 1128). All patients with CRAO and amaurosis fugax should be started on oral aspirin if it is not medically contraindicated.

Retinal detachment

This causes a painless progressive visual field loss. The shadow corresponds to the area of detached retina. Following a tear in the retina, fluid collects in the potential space between the sensory retina and the pigment epithelium (Fig. 20.32). Patients usually report a sudden onset of floaters often associated with flashes of light prior to the detachment. These patients should be referred to an ophthalmologist for a detailed fundal examination.

Age-related macular degeneration (AMD)

This is the commonest cause of visual impairment in patients over 50 years in the western world, and blind registration in this age group. It affects 10% of people over 65 years and 30% over 80 years. Mutations in various genes have been reported; fibulin 5, complement factor 8, Arg 80 Gly variant of complement C3.

The cause is unknown but suggested risk factors include increasing age, smoking, hypertension, hypercholesterolaemia and ultraviolet exposure.

There are two types:

- **Non-exudative (dry) macular degeneration** describes a painless and progressive loss of vision. With age,

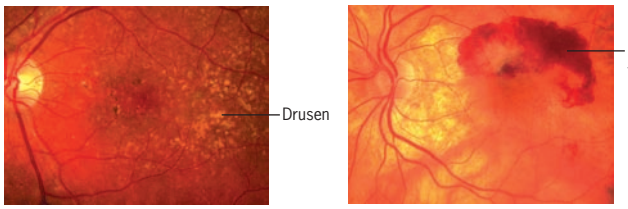


Fig. 20.33 **Age-related macular degeneration. (a)** With deposition of material (drusen) over the macula. **(b)** With choroidal neovascularization.

lipofuscin deposits (drusen) are found between the retinal pigment epithelium (RPE) and Bruch's membrane (see Fig. 20.13 and Fig. 20.33a). Drusen may be hard or soft and there may be focal RPE detachment. Not all patients with these changes will be affected visually but some develop distortion and blurring of their central vision. Extensive atrophy of RPE can occur (geographic atrophy).

- **Exudative (wet) AMD** (10% of cases) occurs with the development of abnormal subfoveal choroidal neovascularization in the region of the macula and causes severe central visual loss (Fig. 20.33b).

Treatment

The Age-Related Eye Disease Study (AREDS) has shown that vitamins C and E, β carotene, zinc and copper have been shown to slow progression of the disease.

Anti-vascular endothelial growth factors (anti-VEGF) are being used in the treatment of wet AMD. Pegaptanib, ranibizumab and bevacizumab are given by intravitreal injections while verteporfin is given intravenously and activated by local eye irradiation using non-thermal red light.

Patients with central distortion or with frank macular pathology should be referred to the ophthalmologist for assessment of treatment. Only a few patients will fulfil the clinical requirement to receive laser treatment or photodynamic therapy (PDT). All patients should be assessed for low-vision aids such as magnifying glasses as this may help to improve their independence.

CAUSES OF VISUAL LOSS

Box 20.3 summarizes the causes of visual loss. In developing countries the common causes of blindness are similar to those in the developed world and include trauma, cataract, diabetic retinopathy, glaucoma and macular degeneration. In addition, trachoma due to *Chlamydia trachomatis* (see p. 143) accounts for 10% of global blindness and onchocerciasis (river blindness) due to *Onchocerca volvulus* (see p. 165) accounts for blindness in about 1 million people, although this is decreasing with treatment. In leprosy, 70% of patients have ocular involvement, and blindness occurs in 5–10% of these. Ocular involvement is common in cerebral malaria (see p. 155), although loss of vision is rare. HIV infection can produce uveitis but the major problem is severe opportunistic infection of the eye when the CD4 count falls (see p. 202).

Vitamin A deficiency and xerophthalmia affects millions each year, and the WHO classification of xerophthalmia by ocular signs is shown in Table 5.10 (see p. 219).

Box 20.3 Loss of vision: summary

Painless loss of vision	Painful loss of vision
Cataract	Acute angle-closure glaucoma
Open-angle glaucoma	Giant cell arteritis
Retinal detachment	Optic neuritis
Central retinal vein occlusion	Uveitis
Central retinal artery occlusion	Scleritis
Diabetic retinopathy	Keratitis
Vitreous haemorrhage	Shingles
Posterior uveitis	Orbital cellulitis
Age-related macular degeneration	Trauma
Optic nerve compression	
Cerebral vascular disease	

FURTHER READING

Asbell PA, Dualan I, Mindel J. Age related cataract. *Lancet* 2005; **365**: 599–609.

Jager RD, Mieler WF. Miller JW. Age-related macular degeneration. *New England Journal of Medicine* 2008; **358**: 2606–2617.

Lynn WA, Lightman S. The eye in systemic infection. *Lancet* 2004; **364**: 1439–1450.

Sakimoto E, Roseblatt MI, Azar DT. Laser eye surgery for refractive errors. *Lancet* 2006; **367**: 1432–1447.

Weinreb RN, Khaw PT. Primary open-angle glaucoma. *Lancet* 2004; **363**: 1711–1720.

21 Neurological disease

The impact of neurological disease 1095
Common symptoms and signs 1095
Examination and formulation 1097
Functional neuroanatomy 1098
Cranial nerves 1100
I: Olfactory nerve 1100
II: Optic nerve and visual system 1100
III, IV, VI: Oculomotor, trochlear and abducens nerves 1104
V: Trigeminal nerve 1106
VII: Facial nerve 1107
VIII: Vestibulo-cochlear nerve 1108
Lower cranial nerves IX, X, XI, XII 1110
Brainstem lesions 1110
Motor control systems 1111
Corticospinal (pyramidal) system 1111
Extrapyramidal system 1112
Cerebellum 1113
Tremor 1114
Lower motor neurone lesions 1115
Spinal reflex arc 1115
Sensory pathways and pain 1115
Lesions of the sensory pathways 1116
Pain 1118
Bladder control and sexual dysfunction 1119
Neurological tests 1119
Neuro-imaging 1119
Unconsciousness and coma 1122
The unconscious patient 1124
Stroke and cerebrovascular disease 1126
Clinical syndromes 1128
Cortical venous thrombosis and dural venous sinus thrombosis 1136
Headache, migraine and facial pain 1136
Epilepsy and loss of consciousness 1139
Epilepsy 1139
Other forms of disturbed consciousness 1144
Sleep and its disturbances 1145
Movement disorders 1145
Akinetic–rigid syndromes 1145
Other akinetic–rigid syndromes 1148
Dyskinesias 1149
Multiple sclerosis 1151
Nervous system infection and inflammation 1153
Meningitis 1153
Encephalitis 1156
Myelitis 1157
Herpes zoster (shingles) 1157
Neurosyphilis 1157
Neurocysticercosis 1157
HIV and neurology 1158
Other inflammatory conditions 1158
Brain and spinal abscesses 1159
Brain tumours 1160
Hydrocephalus 1162
Traumatic brain injury 1162
Spinal cord disease 1163
Spinal cord compression 1163
Other causes of paraparesis 1164
Management of paraplegia 1165
Degenerative neuronal diseases 1166
Motor neurone disease 1166
Congenital and inherited diseases 1169
Cerebral palsy 1169
Dysraphism 1169
Basilar impression of skull (platybasia) 1169
Neuroectodermal syndromes 1169
Spinocerebellar degenerations 1170
Peripheral nerve disease 1170
Mononeuropathies 1171
Polyneuropathies 1172
Plexus and nerve root lesions 1176
Muscle diseases 1178
Inflammatory myopathies 1179
Metabolic and endocrine myopathies 1179
Neuromuscular junction conditions 1180
Muscular dystrophies 1181
Myotonias 1181
Channelopathies 1182
Metabolic myopathies 1182

THE IMPACT OF NEUROLOGICAL DISEASE

Neurology is a large and diverse subject which covers many conditions that require long-term coordinated care and have serious effects on the daily lives of patients and their families. Prevalence rates vary widely between countries, being dependent upon incidence (Table 21.1), prognosis and level of medical care provided.

COMMON SYMPTOMS AND SIGNS

Pattern recognition in neurology – interpretation of history, symptoms and examination – is very reliable. Practical experience is vital. There are three critical questions:

FURTHER READING

Hirtz D, Thurman DJ, Gwinn-Hardy K et al. How common are the 'common' neurologic diseases? *Neurology* 2007; **68**: 326–337.

World Health Organization. *The Global Burden of Disease*. WHO, Geneva, 1996.

Ask the authors
www.studentconsult.com

Table 21.1 UK incidence: common neurological conditions/100 000/year

Headaches (GP consultations)	2200
Cerebrovascular events	210
Shingles (herpes zoster) and post-herpetic neuralgia	150
Diabetic and other neuropathies	105
Epilepsy	46
Parkinson's disease	19
Severe brain injury and subdural haematoma	13
All CNS tumours	9
Trigeminal neuralgia	8
Meningitis	7
Multiple sclerosis	7
Presenile dementia (below 65 years)	4
Myasthenia, all muscle and motor neurone disease	5

- What is/are the site(s) of the lesion(s)?
- What is the likely pathology?
- Does a recognizable disease fit this pattern?

This is the essence of clinical diagnosis. Notes of symptoms should portray the *story* from the patient or relatives.

Headaches

Headache is an almost universal experience and one of the most common symptoms in medical practice. It varies from an infrequent and trivial nuisance to a pointer to serious disease.

Mechanisms

Pain receptors are located at the base of the brain in arteries and veins and throughout meninges, extracranial vessels, scalp, neck and facial muscles, paranasal sinuses, eyes and teeth. Curiously, brain substance is almost devoid of pain receptors.

Head pain is mediated by mechanical and chemical receptors (e.g. stretching of meninges, 5-HT and histamine stimulation). Nerve impulses travel centrally via the Vth and IXth cranial nerves and upper cervical sensory roots.

Most headaches are benign, but the diagnostic issue – and usual concern – is the question of serious disease. Here are some useful clinical pointers.

Chronic (benign) and recurrent headaches

Almost all recurring headaches lasting hours or days – band-like, generalized head pains, with a history for several years or months – are vaguely ascribed to muscle tension and/or migraine (p. 1136). Depression is a common accompaniment.

In localized pain of short duration (minutes to hours), sinusitis, glaucoma and migrainous neuralgia (p. 1137) should be considered. Headaches are not caused by essential hypertension; malignant hypertension, with arterial damage and brain swelling, occasionally causes headache (p. 1096).

Eyestrain from refractive error does not cause headache, though new prescription lenses sometimes provoke pain.

Pressure headaches

Intracranial mass lesions displace and stretch meninges and basal vessels. Pain is provoked when these structures are shifted either by a mass or by changes in cerebrospinal fluid (CSF) pressure, e.g. coughing. Cerebral oedema around brain tumours causes further shift. These 'pressure headaches' typically become worse on lying down.

Any headache present on waking and made worse by coughing, straining or sneezing may be due to a mass lesion. Vomiting often accompanies pressure headaches. Such headaches are caused early, over days or weeks, by posterior fossa masses (see hydrocephalus, p. 1162), but over a longer time scale – months or years – by hemisphere tumours.

A rare cause of prostrating headache with lower limb weakness is an intraventricular tumour causing intermittent hydrocephalus.

Table 21.2 Common gait abnormalities

Spasticity/hemiparesis
Parkinson's disease
Cerebellar ataxia
Position sense loss (stamping)
Distal weakness (slapping)
Proximal weakness (waddling)
Apraxia of gait

Headache of subacute onset

The onset and progression of a headache over days or weeks with or without features of a pressure headache should always raise suspicion of an intracranial mass or serious intracranial disease. Encephalitis (p. 1156), viral meningitis (p. 1154) and chronic meningitis (p. 1155) should also be considered.

Headaches with scalp tenderness

Patches of exquisite tenderness overlying superficial scalp arteries are caused by giant cell arteritis (p. 1138) in patients over 50.

Headache following head injury

The majority of post-trauma headaches lasting days, weeks or months are not caused by any serious intracranial pathology. However, subdural haematoma (p. 1135) must be considered.

A single episode of severe headache

This common emergency is caused by one of the following:

- subarachnoid haemorrhage (SAH) and cervical arterial dissection
- migraine, or other benign headaches
- meningitis (occasionally).

Particular attention should be paid to suddenness of onset (suggestive of SAH). The exact time of onset, time to peak, duration, associated symptoms and previous headache history should be documented. Neck stiffness, vomiting (meningeal irritation) and a rash and/or fever suggest bacterial meningitis.

Difficulty walking and falls

Change in walking pattern is a common complaint (Table 21.2). Arthritis and muscle pain make walking stiff and slow (antalgia). The *pattern* of gait is valuable diagnostically.

Spasticity and hemiparesis

Spasticity (p. 1112), more pronounced in extensor muscles, with or without weakness, causes stiff and jerky walking. Toes of shoes become scuffed, catching level ground. Pace shortens; a narrow base is maintained. Clonus – involuntary extensor rhythmic leg jerking – may occur.

In a hemiparesis when spasticity is unilateral and weakness marked, the stiff, weak leg is circumducted and drags.

Parkinson's disease: shuffling gait

There is muscular rigidity (p. 1146) throughout extensors and flexors. Power is preserved; the pace shortens, and slows to a shuffle; its base remains narrow. A stoop and diminished arm swinging become apparent. Gait becomes festinant (hurried) with short rapid steps. There is difficulty turning quickly and initiating movement, sometimes with falls. *Retropulsion* means small backward steps, taken involuntarily when a patient is halted.

Cerebellar ataxia: broad-based gait

In lateral cerebellar lobe disease (p. 1114) stance becomes broad-based, unstable and tremulous. Ataxia describes this incoordination. When walking, the person tends to veer to the side of the affected cerebellar lobe.

In disease of midline structures (cerebellar vermis), the trunk becomes unsteady without limb ataxia, with a tendency to fall backwards or sideways – truncal ataxia.

Sensory ataxia: stamping gait

Peripheral sensory lesions (e.g. polyneuropathy, p. 1172) cause ataxia because of loss of *proprioception* (position sense). Broad-based, high-stepping, stamping gait develops. This form of ataxia is exacerbated by removal of sensory input (e.g. vision) and worse in the dark. Romberg's test, first described in sensory ataxia of tabes dorsalis (p. 1157), becomes positive.

Lower limb weakness: slapping and waddling gaits

When weakness is distal, each foot must be lifted over obstacles. When ankle dorsiflexors are weak, e.g. in a common peroneal nerve palsy (p. 1172), the sole returns to the ground with an audible *slap*.

Weakness of proximal lower limb muscles (e.g. polymyositis, muscular dystrophy) causes difficulty rising from sitting. Walking becomes a *waddle*, the pelvis being poorly supported by each leg.

Gait apraxia

With frontal lobe disease (e.g. tumour, hydrocephalus, infarction), acquired walking *skills* become disorganized. Leg movement is normal when sitting or lying but initiation and organization of walking fail. Shuffling small steps (*marche à petits pas*), gait ignition failure or undue hesitancy may predominate. Urinary incontinence and dementia are often present.

Falls

Falls in the elderly are a major cause of hospital admission, e.g. following fractures. Often no precise cause can be found. A multidisciplinary approach is essential, e.g. reviewing risk factors such as rugs, stairs, footwear and home circumstances.

Dizziness, vertigo and blackouts

Dizziness covers many complaints, from a vague feeling of unsteadiness to severe, acute vertigo. It is frequently used to describe light-headedness, panic, anxiety, palpitations and chronic ill-health. The real nature of this symptom must be determined.

Vertigo (p. 1108) means the illusion of movement, a sensation of rotation or tipping. The patient feels the surroundings are spinning or moving. This is distressing and often accompanied by nausea or vomiting.

Blackout, like dizziness, is simply descriptive, implying either altered consciousness, visual disturbance or falling. Epilepsy (p. 1139) and syncope are mentioned in detail (p. 1144); hypoglycaemia and anaemia must be considered. Commonly no sinister cause is found. A careful history is essential.

Collapse is a vague term, but often used. Avoid it.

No serious disease is found in many patients (>20%) referred with symptoms suggestive of possible neurological conditions. y.

Fatigue is common: when it is an isolated symptom, neurological disease is rarely discovered. There is a borderland (sometimes contentious) between neurology and psychiatr

EXAMINATION AND FORMULATION (see Practical boxes 21.1 and 21.2 and Table 21.3)

Following a short or detailed examination, relevant findings are summarized in a brief formulation – the basis for investigation, transfer of information and management.

Practical Box 21.1 Five-part short neurological examination

1 **Look at the patient**
- General demeanour
- Speech
- Gait
- Arm swinging

2 **Head**
- Fundi
- Pupils
- Eye movements
- Facial movements
- Tongue

3 **Upper limbs**
- Posture of outstretched arms
- Wasting, fasciculation
- Power, tone
- Coordination
- Reflexes

4 **Lower limbs**
- Power
- Tone
- Reflexes
- Plantar responses

5 **Sensation**
- Ask the patient

FURTHER READING

Clarke C. The language of neurology – symptoms, signs and basic investigations. In: Clarke C, Howard R, Rossor M, Shorvon S (eds) *Neurology: A Queen Square Textbook.* Oxford: Blackwell, 2009.

Practical Box 21.2 Ten-part neurological examination

1 **State of consciousness, arousal, appearance**
2 **Mental state, attitude, insight** (Box 22.5, p. 1189)
3 **Cognitive function**
Orientation, recall, level of intellect, language, other cortical problems, e.g. apraxia
4 **Gait and balance tests**
5 **Skull – contour, bruits**
6 **Neck – stiffness, palpation and auscultation of carotids**
7 **Cranial nerves** (Table 21.5)
8 **Motor system**
Upper limbs:
Wasting and fasciculation
Posture of arms: drift, rebound, tremor
Tone: spasticity, clonus, extrapyramidal rigidity
Power: 0–5 scale (Table 21.3)
Tendon reflexes: + or ++ normal; +++ pathological; 0 = absent with reinforcement
Thorax and abdomen:
Respiration
Thoracic and abdominal muscles
Abdominal reflexes
Cremasteric reflexes
Lower limbs:
Wasting and fasciculation
Tone, power and tendon reflexes
Plantar responses
9 **Coordination and fine movements**
10 **Sensory system**
First: is feeling in limbs, face and trunk normal?
Posterior columns:
Vibration (128 Hz tuning fork)
Joint position
Light touch
2-point discrimination (normal: 0.5 cm finger tips, 2 cm soles)
Spinothalamic tracts:
Pain: use a split orange-stick or a sterile pin
Temperature: warm and cold objects
If sensation is abnormal, chart areas involved

Table 21.3 Six grades of muscle power

Grade	Definition
5	Normal power
4	Active movement against gravity and resistance
3	Active movement against gravity
2	Active movement with gravity eliminated
1	Flicker of contraction
0	No contraction

FUNCTIONAL NEUROANATOMY

The neurone and synapse

The neurone is the functional unit of the entire nervous system (Fig. 21.1). Its cell body and axon terminate in a synapse. Size and type of each group of neurones vary. A thoracic spinal cord α-motor neurone has an axonal length of more than 1 metre and innervates between several hundred and 2000 muscle fibres in one leg – a motor unit. By contrast, some spinal or intracerebral interneurones have axons under 100 μm long, terminating on one neuronal cell body.

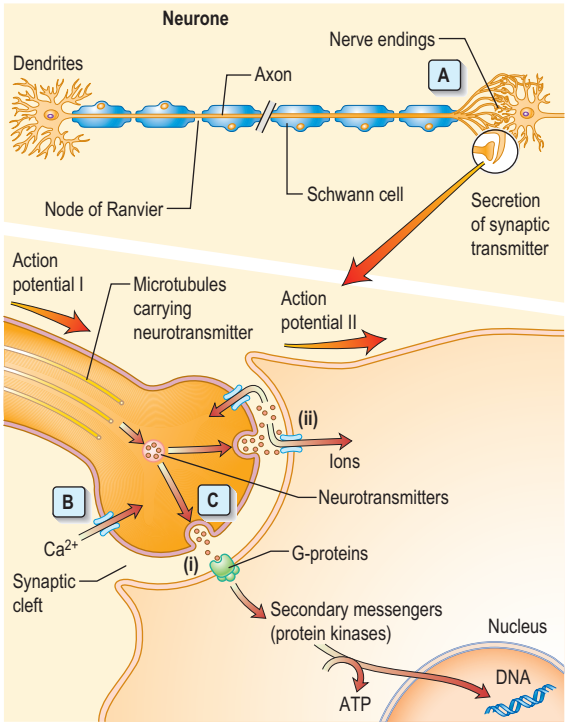


Fig. 21.1 **The functional unit: neurone and neurotransmitters. (A) The action potential**, i.e. nerve impulse, travels down the axon. Microtubules carry neurotransmitters to nerve endings. **(B) Action potential I** depolarizes the synaptic membrane, opening voltage-gated calcium channels. **(C) Influx of calcium ions** cause vesicles to fuse with the membrane allowing neurotransmitter binding to receptors **(i)** and activation of secondary messengers that modulate gene transcription and also open ligand-gated channels **(ii)**. This allows ions to enter, depolarize the membrane and initiate **action potential II**.

Neurotransmitters

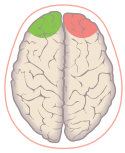
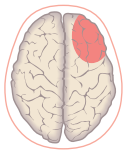
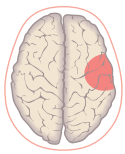
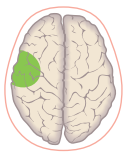
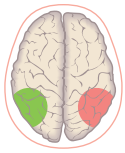
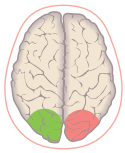
Neurotransmitters are excitatory (acetylcholine, noradrenaline, adrenaline, 5-hydroxytryptamine, dopamine, glutamate and aspartate) or inhibitory (gamma-aminobutyric acid (GABA), histamine and glycine). Neuropeptides, e.g. vasopressin, ACTH, substance P and opioid peptides, as well as the purines (ATP and AMP) are both excitatory and inhibitory. Nitric oxide also has modulatory properties.

Synaptic transmission is mediated by neurotransmitters released by action potentials passing down an axon. Neurotransmitters react with postsynaptic receptors and are removed by transporter proteins. The neurotransmitter–receptor reaction increases ionic permeability and propagates a further action potential. Axonal electrical activity and synaptic chemical release is the basis of neurological function.

The role of neurotransmitters in pathogenesis continues to be evaluated; acute and chronic neuronal injury is mediated by a final common pathway involving excessive stimulation of glutamate receptors.

Clinical features of focal brain lesions: general mechanisms

The symptoms and signs suggest the area of the brain that is malfunctioning (for example, aphasia – the left frontal lobe, hemiparesis – internal capsule, or a Bell's palsy – VIIth cranial (facial) nerve).

Site of lesion	Disorder	R L
Frontal, either	Intellectual impairment Personality change Urinary incontinence Monoparesis or hemiparesis	
Frontal, left	Broca's aphasia	
Temporo-parietal, left	Acalculia Alexia Agraphia Wernicke's aphasia Right-left disorientation Homonymous field defect	
Temporal, right	Confusional states Failure to recognize faces Homonymous field defect	
Parietal, either	Contralateral sensory loss or neglect Agraphaesthesia Homonymous field defect	
Parietal, right	Dressing apraxia Failure to recognize faces	
Parietal, left	Limb apraxia	
Occipital/ occipitoparietal	Visual field defects Visuospatial defects Disturbances of visual recognition	

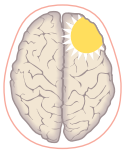
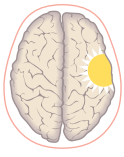
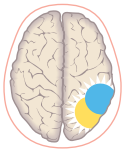
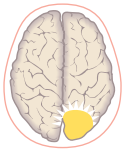
Site of lesion	Effects	R L
Frontal	Partial seizures–focal motor seizures of contralateral limbs Conjugate deviation of head and eyes away from the lesion	
Temporal	Formed visual hallucinations Complex partial seizures Memory disturbances (e.g. *déjà vu*)	
Parietal	Partial seizures–focal sensory seizures of contralateral limbs	
Parieto-occipital	Crude visual hallucinations (e.g. shapes in one part of the field)	
Occipital	Visual disturbances (e.g. flashes)	

Fig. 21.2 **Principal features of destructive cortical lesions in a right-handed individual.**

Fig. 21.3 **Effects of irritative cortical lesions.**

Focal lesions of the cortex, and lesions throughout the nervous system, cause symptoms and signs by two processes:

- *Suppression or destruction of neurones* and surrounding structures (Fig. 21.2). This is the commonest process – part of the system simply fails to work.
- Synchronous discharge of neurones by *irritative* lesions (Fig. 21.3), e.g. cortical lesions, causes epilepsy, either partial or generalized.

Localization within the cerebral cortex

This subject causes unnecessary difficulty. Work on neuronal networks, functional imaging and plasticity questions traditional views of highly specific localization of cortical function. The following paragraphs summarize areas of clinical relevance.

The dominant hemisphere (usually left)

The concept of cerebral dominance arose from a simple observation: right-handed stroke patients with acquired language disorders had destructive lesions within the left hemisphere. Right-handed (and 70% of left-handed) people have language function on the left.

More specifically, destructive lesions within the left fronto-temporo-parietal region cause disorders of communication:

- spoken language – *aphasia*, also called *dysphasia*
- writing – *agraphia*
- reading – *acquired alexia*.

Developmental dyslexia describes delayed, disorganized reading and writing ability in children, usually with normal intelligence.

Aphasia

Aphasia is loss of or defective language from damage to the speech centres within the left hemisphere. Numerous varieties have been described.

Broca's (expressive, anterior) aphasia

Damage in the left frontal lobe causes reduced speech fluency with relatively preserved comprehension. The patient makes great efforts to initiate language, which becomes reduced to a few disjointed words with failure to construct sentences. Patients who recover say they knew what they wanted to say, but could not get the words out.

Wernicke's (receptive, posterior) aphasia

Left temporo-parietal damage leaves fluency of language but words are muddled. This varies from insertion of a few incorrect or non-existent words into speech to a profuse outpouring of jargon (i.e. rubbish with wholly non-existent words). Severe jargon aphasia is bizarre and often mistaken for psychotic behaviour.

Patients who recover from Wernicke's aphasia say that they found speech, both their own and others', like an unintelligible foreign language, i.e. incomprehensible, but they could neither stop speaking nor understand speech.

Nominal (anomic, amnestic) aphasia

This means difficulty naming familiar objects. Naming difficulty is an early feature in all types of aphasia.

Global (central) aphasia

This means the combination of the expressive problems of Broca's aphasia and the loss of comprehension of Wernicke's with loss of both language production and understanding. This is due to widespread damage to speech areas and is the commonest aphasia after a severe left hemisphere infarct. Writing and reading are also affected.

Dysarthria

Dysarthria is disordered articulation – slurred speech. Language is intact. Paralysis, slowing or incoordination of muscles of articulation or local discomfort causes various patterns of dysarthria. Examples are the *gravelly* speech of pseudobulbar palsy (p. 1110), the *jerky*, ataxic speech of cerebellar lesions, the *monotone* of Parkinson's, and speech in myasthenia that *fatigues* and dies away. Many aphasic patients are also dysarthric.

The non-dominant hemisphere

Disorders in right-handed patients with right hemisphere lesions are often difficult to recognize. There are abnormalities of perception of internal and external space. Examples are losing the way in familiar surroundings, failing to put on clothing correctly (dressing apraxia), or failure to draw simple shapes – constructional apraxia.

Memory and its disorders

Disorders of memory follow damage to the medial surfaces of both temporal lobes and their brainstem connections – the hippocampi, fornices and mammillary bodies. Bilateral lesions are necessary to cause *amnesia*. In all organic memory disorders recent events are recalled poorly, in contrast to the relative preservation of distant memories.

Memory loss (the amnestic syndrome) is part of dementia (p. 1167) but also occurs as an isolated entity (Table 21.4).

Neuroanatomy: essential elements

For clinical purposes the complexity of neuroanatomy must be reduced to its core elements:

- cranial nerves
- three systems of motor control:
  - corticospinal or pyramidal system
  - extrapyramidal system
  - cerebellum
- motor unit
- reflex arc
- sensory pathways and pain
- control of the bladder and sexual function.

Table 21.4 Causes of an amnestic syndrome

Dementia
Alcohol (Wernicke–Korsakoff syndrome)
Head injury (severe)
Anoxic brain damage and following carbon monoxide poisoning
Posterior cerebral artery occlusion (bilateral)
Herpes simplex encephalitis
Chronic sedative and solvent abuse
Bilateral invasive tumours
Following hypoglycaemia
Arsenic poisoning (very rare)

CRANIAL NERVES (Table 21.5)

I: OLFACTORY NERVE

This sensory nerve arises from olfactory (smell) receptors within nasal mucosa. Branches pierce the cribriform plate and synapse in the olfactory bulb. The olfactory tract passes to the olfactory cortex (anteromedial surface of frontal and temporal lobes).

Anosmia (loss of sense of smell) is caused by head injury and tumours of the olfactory groove (e.g. meningioma, frontal glioma). Olfaction is temporarily (occasionally permanently) lost or diminished after upper respiratory infections and when the nostrils are blocked.

II: OPTIC NERVE AND VISUAL SYSTEM (Fig. 21.4)

The photic energy of light regulated by the pupillary aperture (p. 1086), is converted into action potentials by retinal rod, cone and ganglion cells. The lens, under control of the ciliary muscle, produces the image (inverted) on the retina (1), i.e. an object in the lower field of vision is projected to the upper retina; one in a temporal field to the nasal retina. Each optic nerve (2), sheathed in meninges, carries axons from retinal ganglion cells.

At the chiasm (3), fibres travelling in the nasal portions of optic nerves cross and join with uncrossed temporal optic nerve fibres to form each optic tract. These fibres synapse at each lateral geniculate body (4). One optic tract thus carries fibres from the temporal ipsilateral retina and the nasal contralateral retina. Some fibres reaching each lateral geniculate body pass to the brainstem to control refraction (lens) and pupillary aperture.

From the lateral geniculate body, fibres pass in the optic radiation through the parietal and temporal lobes

Table 21.5 Cranial nerves

Number	Name	Main clinical action
I	Olfactory	Smell
II	Optic	Vision, fields, afferent light reflex
III	Oculomotor	Eyelid elevation, eye elevation, ADduction, depression in ABduction, efferent (pupil)
IV	Trochlear	Eye intorsion, depression in ADduction
V	Trigeminal	Facial (and corneal) sensation, mastication muscles
VI	Abducens	Eye ABduction
VII	Facial	Facial movement, taste fibres
VIII	Vestibular Cochlear	Balance and hearing
IX	Glossopharyngeal	Sensation – soft palate, taste fibres
X	Vagus	Cough, palatal and vocal cord movements
XI	Accessory	Head turning, shoulder shrugging
XII	Hypoglossal	Tongue movement

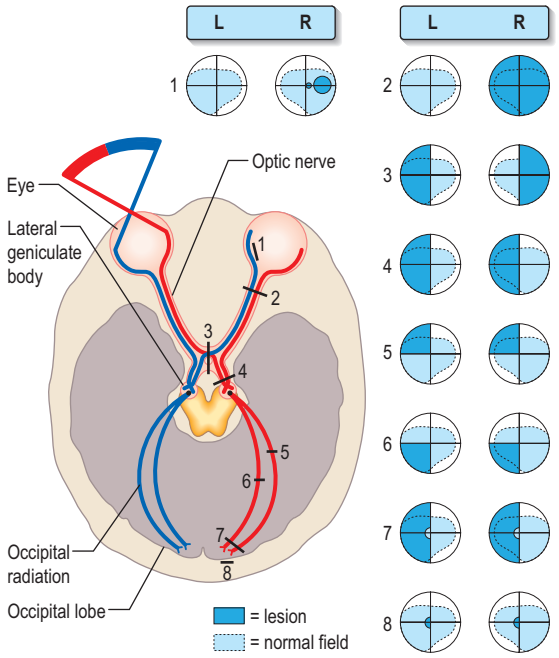


Fig. 21.4 **The visual pathway.**

1 Paracentral scotoma – retinal lesion.
2 Mononuclear field loss – complete optic nerve lesion.
3 Bitemporal hemianopia – chiasmal lesion.
4 Homonymous hemianopia – optic tract lesion.
5 Homonymous quadrantanopia – temporal lesion.
6 Homonymous quadrantanopia – parietal lesion.
7 Homonymous hemianopia with macular sparing – occipital cortex or optic radiation.
8 Homonymous hemianopia (hemiscotoma) – occipital pole lesion.

(5 and 6) to reach the visual, or calcarine, cortex of the occipital lobe (7 and 8). Upper retinae (lower visual fields) project in the optic radiation through the parietal lobes to the upper part of the visual cortex, and the lower retinae (upper fields) through the temporal lobes, to the lower visual cortex.

Impulses reach the cortex in strictly vertical topographical order (i.e. *upper* field to *lower* retina, tract, radiation and cortex, and vice versa.). Within the visual cortex itself there are synaptic connections between groups of cells that detect lines, orientation, shapes, movement, colour and depth. These are processed by neighbouring visual association areas.

Visual field defects caused by lesions of each optic tract, radiation and cortex are called *homonymous* to indicate the different (i.e. bilateral) origins of each unilateral pathway. (A homonym is one word used to denote different things.)

Field defects are hemianopic when half the field is affected and quadrantanopic when a quadrant is affected. Congruous denotes symmetry, and incongruous, lack of symmetry. Bitemporal defects (damage to crossing nasal fibres) are caused by optic chiasm lesions.

Visual acuity

This is assessed with a Snellen chart and/or Near Vision Reading Types, corrected for refractive errors with lenses or a pinhole occluder. Normal acuity is 6/6 to 6/9; if less, an explanation is necessary.

Table 21.6 Principal causes of an optic nerve lesion

Optic and retrobulbar neuritis
Optic nerve compression, e.g. primary or secondary tumour, aneurysm
Toxic optic neuropathy, e.g. tobacco, ethambutol, methyl alcohol, quinine, hydroxy chloroquine
Syphilis
Ischaemic optic neuropathy, e.g. giant cell arteritis
Hereditary optic neuropathies, e.g. Leber's
Severe anaemia
Vitamin B_{12} deficiency
Trauma
Infection (spread of paranasal sinus infection or orbital cellulitis)
Papilloedema and its causes (Table 21.7)
Bone disease affecting optic canal, e.g. Paget's

Visual loss is discussed on page 1094. Causes include:

- ocular causes, e.g. glaucoma, macular degeneration, cataract, retinal detachment, diabetes, trachoma, leprosy, vitamin A deficiency, trauma, onchocerciasis (river blindness)
- lesions of the visual pathway (optic nerve, chiasm, tract, radiation or cortex).

Visual field defects (Fig. 21.4)

These are checked by confrontation with white- and red-headed pins and, if abnormal or in doubt, recorded with a Goldmann (or similar) screen.

Retinal and local eye lesions (1, Fig. 21.4)

These produce either scotomata (holes) or peripheral visual loss (tunnel vision). Common causes are diabetic retinal vascular disease, glaucoma and retinitis pigmentosa. Local lesions (e.g. cataract) also cause visual loss.

Optic nerve lesions (2, Fig. 21.4)

Unilateral visual loss, commencing with a central or paracentral (off-centre) scotoma, is the hallmark of an optic nerve lesion. A total optic nerve lesion causes unilateral blindness with loss of pupillary light reflex (direct and consensual) when the blind eye is illuminated (see afferent pupillary defect, p. 1103). Causes are listed in Table 21.6.

The principal pathological appearances of the nerve (optic disc) are:

- disc swelling and hyperaemia (papilloedema)
- disc pallor (optic atrophy).

Papilloedema and optic neuritis

Papilloedema simply means swelling of the papilla – the optic disc (Table 21.7). In all forms of disc oedema there is axonal swelling within the optic nerve and blockage of axonal transport, with capillary and venous congestion.

Optic neuritis means *inflammation* in the nerve with swelling of the disc.

The earliest ophthalmoscopic signs of swelling are disc pinkness, with blurring and heaping up of disc margins, nasal first. There is loss of normal, visible, spontaneous pulsation

Table 21.7 Causes of papilloedema

Raised intracranial pressure
Brain tumour, abscess, haematoma, intracranial haemorrhage and SAH, idiopathic intracranial hypertension, hydrocephalus, encephalitis
Optic nerve disease
Optic neuritis, e.g. multiple sclerosis
Hereditary optic neuropathy, e.g. Leber's
Ischaemic optic neuropathy, e.g. giant cell arteritis
Toxic optic neuropathy, e.g. methanol
Hypervitaminosis A
Venous occlusion
Cavernous sinus thrombosis
Central retinal vein thrombosis/occlusion
Orbital mass lesions
Retinal vascular disease
Malignant hypertension
Vasculitis, e.g. systemic lupus erythematosis
Metabolic causes
Hypercapnia, chronic hypoxia, hypocalcaemia
Disc infiltration
Leukaemia, sarcoidosis, optic nerve glioma

SAH, subarachnoid haemorrhage.

of retinal veins within the disc. The physiological cup becomes obliterated, the disc engorged with dilated vessels. Small haemorrhages often surround the disc.

Various conditions simulate true disc oedema. Marked hypermetropic (long-sighted) refractive errors make a disc appear pink, distant and ill-defined. Opaque (myelinated) nerve fibres at disc margins and hyaline bodies (drusen, p. 1094) can be mistaken for disc swelling.

Disc infiltration also causes a swollen disc with raised margins (e.g. in leukaemia).

When there is doubt about disc oedema, i.v. fluorescein angiography is diagnostic; retinal leakage is seen with papilloedema.

Early papilloedema from causes other than optic neuritis (e.g. a brain tumour) often produces few if any visual problems – the underlying disease is the source of the patient's symptoms. However, as disc oedema progresses, blind spot enlargement and visual blurring develop. The disc becomes further engorged, its arterial blood flow falls; eventually, optic nerve infarction occurs, with sudden severe and permanent visual loss.

Optic neuritis

Multiple sclerosis is the most common cause of inflammation within the optic nerve. Disc swelling and visual loss due to optic neuritis is usually accompanied by dull ocular pain. *Retrobulbar neuritis* means that the inflammation is within the optic nerve but behind the eye; no abnormality is seen at the disc despite visual impairment.

Leber's hereditary optic neuropathy (LHON)

LHON is a cause of isolated blindness in otherwise healthy young men. Unilateral or bilateral optic nerve neuropathy develops over days or weeks. There is sometimes disc swelling and telangiectasia around the disc in the acute phase, followed by optic atrophy. Severe bilateral visual loss is usual by the age of 40. Genetic analysis points to mitochondrial DNA mutations (p. 37), in many cases at G11778A. Exceptional cases occur in women.

Optic atrophy

Optic atrophy means disc pallor, from loss of axons, glial proliferation and decreased vascularity. This follows many processes, e.g. nerve infarction from thromboembolism or following papilloedema, inflammation (demyelinating optic neuritis in MS, syphilis, LHON), optic nerve compression, previous trauma, and toxic and metabolic causes (vitamin B_{12} deficiency, quinine and methyl alcohol). Optic atrophy is called *consecutive* or secondary when it follows papilloedema of any cause. The degree of visual loss depends upon underlying pathology.

Optic chiasm (3, Fig. 21.4)

Bitemporal hemianopic fields occur when a mass compresses the chiasm. Common causes are:

- pituitary neoplasm (p. 972)
- meningioma
- craniopharyngioma
- secondary neoplasm.

In *any* case of bilateral visual loss, chiasmal compression must always be considered; with unilateral visual loss consider an optic nerve compression.

Optic tract and optic radiation (4, 5 and 6, Fig. 21.4)

Optic tract lesions (rare) cause field defects that are homonymous, hemianopic and often incomplete and incongruous. Optic radiation lesions cause homonymous quadrantanopic defects. Temporal lobe lesions (e.g. tumour, infarction) cause upper quadrantic defects, and parietal lobe, lower.

Occipital cortex (7 and 8, Fig. 21.4)

Homonymous hemianopic defects are produced by unilateral posterior cerebral artery infarction. The macular cortex (at each occipital pole) is spared; it has a separate blood supply via the middle cerebral artery. Infarction of one occipital pole causes a small, congruous, scotomatous, homonymous hemianopia (8).

Widespread bilateral occipital lobe damage by infarction, trauma or tumour causes *cortical blindness* (Anton's syndrome). The patient cannot see but characteristically lacks insight into this; he or she may even deny it. Pupillary responses remain normal (p. 1130).

Pupil dilatation

Sympathetic impulses via fibres in the nasociliary nerve pass to dilator pupillae. Sympathetic preganglionic fibres to the eye (and face) originate in the hypothalamus, pass uncrossed through midbrain and lateral medulla, and emerge from the spinal cord at T1 (close to the lung apex) to form the superior cervical ganglion at C2. Postganglionic fibres form a plexus around the carotid bifurcation. Fibres pass to the pupil from part of this plexus surrounding the *internal* carotid. Those fibres to the face (sweating and piloerection) arise from the part of the plexus surrounding the *external* carotid. See Horner's syndrome (p. 1103).

Pupil constriction

Parasympathetic impulses cause pupillary constriction. Fibres in the short ciliary nerves arise from the ciliary ganglion and pass to *sphincter pupillae*, causing constriction. Parasympathetic and light reflex pathways are shown in Figure 21.5.

Light reflex

Light constricts the pupil being illuminated (direct reflex) and, by the consensual reflex, the contralateral pupil (Fig. 21.5).

Convergence reflex

Fixation (voluntary or reflex) on a near object requires convergence of the ocular axes and is accompanied by pupillary constriction. Afferent fibres in each optic nerve, passing through each lateral geniculate body, also relay to the *convergence centre*. This centre receives 1a spindle afferents from extraocular muscles – principally medial recti, innervated by the IIIrd nerve.

The efferent autonomic route is convergence centre to Edinger–Westphal nucleus to ciliary ganglion and pupils.

Clinical abnormalities of the pupils

Pupillary abnormalities in coma are discussed on page 1124, and those in brainstem death on page 920. It is easier to see pupillary reactions in a darkened room and with a bright torch.

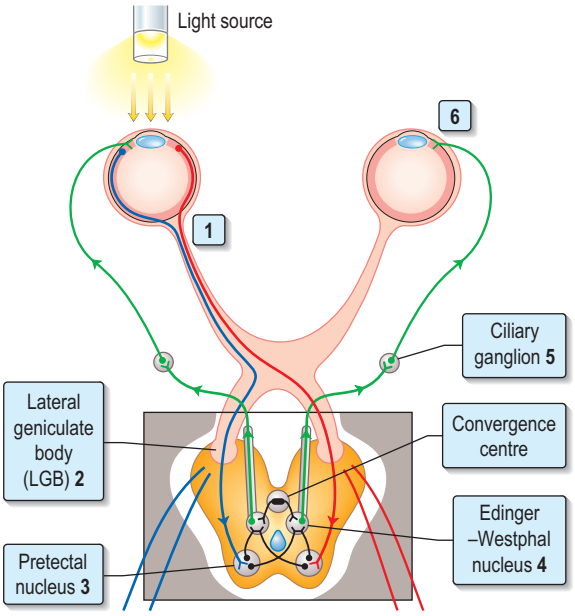


Fig. 21.5 **Pupillary light reflex.**
Afferent pathway

(1) Light generates action potentials in optic nerve axons.
(2) Axons (some decussating at the chiasm) pass through each lateral geniculate body, and
(3) synapse at each pretectal nucleus.

Efferent pathway

(4) Action potentials pass to each Edinger–Westphal nucleus of III, then,
(5) via IIIrd nerve, ciliary ganglion and short ciliary nerves to cause pupil constriction.

Physiological changes and old age

A slight difference between the size of each pupil is common (*physiological anisocoria*). The pupil tends to become small (3–3.5 mm) and irregular in old age (senile miosis); anisocoria is more pronounced. *Convergence* becomes sluggish with ageing.

Afferent pupillary defect. A blind left eye, for example following complete optic nerve section, has a pupil larger than the right. The features of a left *afferent pupillary defect* are:

- The left pupil is unreactive to light (i.e. the direct reflex is absent).
- The consensual reflex (constriction of the right pupil when the left is illuminated) is absent. Conversely, the left pupil constricts when light is shone in the intact right eye, i.e. the consensual reflex of the right eye remains intact.

Relative afferent pupillary defect (RAPD). This occurs with incomplete damage to *one* afferent pupillary pathway (i.e. of one optic nerve relative to the other). RAPD can provide evidence of a left optic nerve lesion, when, for example, left retrobulbar neuritis occurred previously with complete clinical recovery of vision (p. 1152):

- Light shone in the left eye causes both pupils to constrict.
- When light is shone into the intact right eye, both pupils again constrict (i.e. right direct and consensual reflexes are intact).
- When the light is *swung back* to the affected, left eye, its pupil *dilates* slightly, relative to its previous size.

A left RAPD by the swinging light test indicates residual damage in the afferent pupillary fibres of the left optic nerve – the consensual reflex is stronger than the direct.

Horner's syndrome

Unilateral pupillary constriction with slight ptosis and enophthalmos (Horner's) indicates a sympathetic pathway lesion on the same side. The conjunctival vessels become slightly injected (Table 21.8). There is diminution of sweating on the same side, the extent depending upon the level of the lesion:

- Central lesions affect sweating over the entire half of the head, arm and upper trunk.
- Neck lesions proximal to the superior cervical ganglion cause diminished facial sweating.
- Lesions distal to the superior cervical ganglion do not affect sweating at all.

Testing with adrenaline can help indicate the level of the lesion. When drops of 1 : 1000 adrenaline are placed into the conjunctiva, the pupil dilates when the Horner's lesion is distal to the superior cervical ganglion pupil (denervation hypersensitivity). This has little effect on a Horner's pupil from a proximal lesion or on a normal pupil.

Myotonic pupil (Holmes–Adie pupil)

This is a dilated, often irregular, pupil, more frequent in women; it is common and usually unilateral. There is no (or very slow) reaction to bright light and also incomplete constriction to convergence. This is due to denervation in the ciliary ganglion, of unknown cause, and has no

other pathological significance. A myotonic pupil is sometimes associated with diminished or absent tendon reflexes.

Argyll Robertson pupil

This small, irregular (3 mm or less) pupil is fixed to light but constricts on convergence. The lesion is in the brainstem surrounding the aqueduct of Sylvius.

Table 21.8 Causes of Horner's syndrome

Hemisphere and brainstem
Massive cerebral infarction
Pontine glioma
Lateral medullary syndrome
Coning of the temporal lobe

Cervical cord
Syringomyelia
Cord tumours

T1 root
Bronchial neoplasm (apical)
Apical tuberculosis
Cervical rib
Brachial plexus trauma

Sympathetic chain in neck
Following thyroid/laryngeal/carotid surgery
Carotid artery occlusion and dissection
Neoplastic infiltration
Cervical sympathectomy

Miscellaneous
Congenital
Migrainous neuralgia, usually transient
Isolated and of unknown cause

An Argyll Robertson pupil is (almost) diagnostic of neurosyphilis, though it is occasionally seen in diabetes.

III, IV, VI: OCULOMOTOR, TROCHLEAR AND ABDUCENS NERVES

Mechanisms controlling eye movement are:

- central upper motor neurone mechanisms driving *yoked* parallel eye movements (*conjugate gaze*)
- movements generated via each oculomotor, abducens and trochlear nerve to muscles they supply.

Conjugate gaze: yoked eye movements

Fast voluntary and reflex eye movements originate in each frontal lobe. Fibres pass in the anterior limb of the internal capsule and cross in the pons to end in the centre for lateral gaze (paramedian pontine reticular formation – PPRF, Fig. 21.6a), close to each VIth nerve nucleus. Each PPRF also receives fibres from:

- the ipsilateral occipital cortex – pathway concerned with tracking objects
- both vestibular nuclei – pathways linking eye movements with position of the head and neck (doll's head reflexes, p. 920).

Conjugate lateral eye movements are coordinated from each PPRF via the medial longitudinal fasciculus (MLF, Fig. 21.6b). Fibres from the PPRF pass both to the ipsilateral VIth nerve nucleus and, having crossed the midline, to the opposite IIIrd nerve nucleus via the MLF. Each VIth nucleus (lateral rectus) and the opposite IIIrd nerve nucleus (medial rectus and other muscles) are thus linked, enabling the eyes to be driven laterally, with axes parallel at the same velocity.

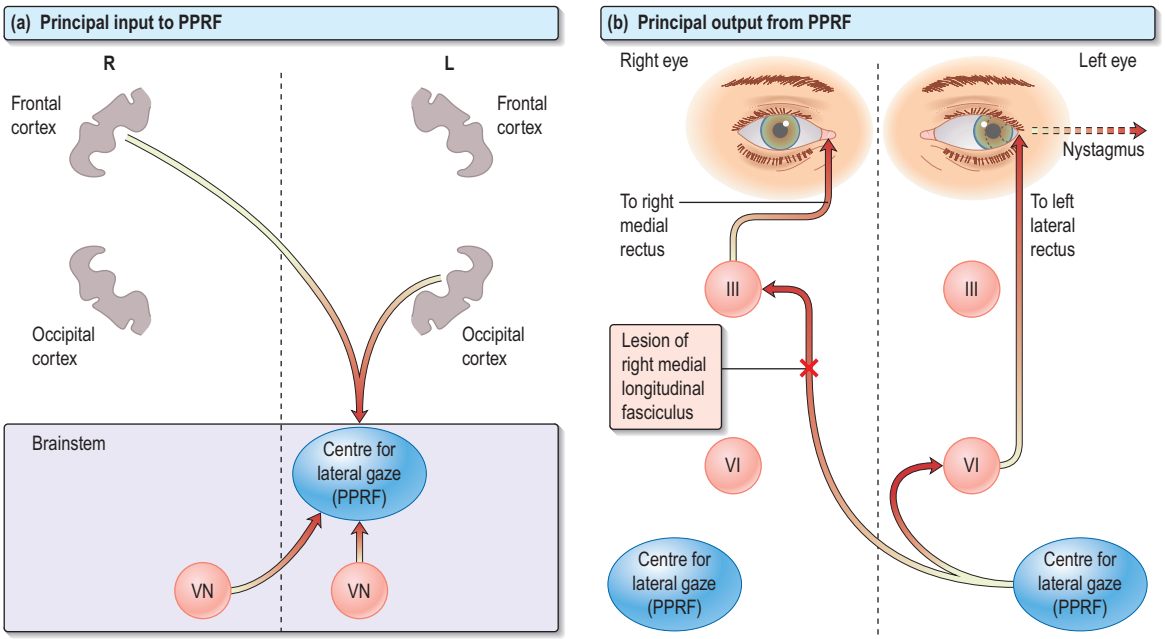


Fig. 21.6 **(a) Paramedian pontine reticular formation (PPRF): principal input.** Impulses from right frontal cortex, left occipital cortex and both vestibular nuclei (VN) drive left PPRF. **(b) PPRF: principal output.** Impulses from PPRF pass via ipsilateral VIth nerve nucleus to lateral rectus muscle (ABduction) and via medial longitudinal fasciculus to right IIIrd nerve nucleus and thus to opposite medial rectus muscle (ADduction) of the right eye. A lesion at X gives limited ADduction to the right eye and nystagmus of the left eye.

Abnormalities of conjugate lateral gaze

A destructive lesion on one side allows yoked lateral gaze to be driven by the intact opposite pathway. A left frontal destructive lesion (e.g. an infarct) leads to failure of conjugate lateral gaze to the right. In an acute lesion the eyes are often deviated to the side of the lesion, past the midline and therefore look towards the left (normal) limbs; there is usually a contralateral (i.e. right) hemiparesis.

An irritative left frontal lobe lesion (e.g. an epileptic focus), stimulates the opposite, right, PPRF and drives lateral gaze *away from* the side of the lesion (i.e. to the right) during an attack.

In the brainstem a unilateral destructive lesion involving the PPRF leads to failure of conjugate lateral gaze *towards* that side. There is usually a contralateral hemiparesis and lateral gaze is deviated towards the side of the paralysed limbs.

Internuclear ophthalmoplegia

Damage to one MLF causes internuclear ophthalmoplegia (INO), a common complex brainstem oculomotor sign seen frequently in MS. In a right INO there is a lesion of the right MLF (Fig. 21.6b). On attempted left lateral gaze the right eye fails to ADduct. The left eye develops coarse nystagmus in ABduction. The side of the lesion is on the side of impaired ADduction, not on the side of the (obvious, unilateral) nystagmus. When present bilaterally, INO is almost pathognomonic of MS. A unilateral INO can also be caused by a small brainstem infarct.

Doll's head reflexes and skew deviation

These are of some diagnostic value in coma (p. 920).

Abnormalities of vertical gaze

Failure of up-gaze is caused by an upper brainstem lesion such as a supratentorial mass or a brainstem tumour (e.g. pinealoma). When the pupillary convergence reflex fails in addition, this is called Parinaud's syndrome. Convergence/retraction nystagmus can occur. Defective up-gaze also develops in certain degenerative disorders (e.g. progressive supranuclear palsy). Some impairment of up-gaze occurs as part of normal ageing.

Weakness of extraocular muscles

Diplopia (double vision) indicates weakness of one or more extraocular muscles. Causes are:

- lesions of the IIIrd, IVth and/or VIth cranial nerves or nuclei
- disorders of the neuromuscular junction (e.g. myasthenia gravis)
- disease of the ocular muscles and trauma
- orbital lesions.

Squint (strabismus)

Squint describes crossed or divergent eyes – the visual axes fail to meet at fixation. Squint is convergent or divergent, non-paralytic or paralytic.

Non-paralytic squint. Non-paralytic or concomitant squint describes a squint dating from infancy in which the angle between the visual axes does not vary when the eyes are moved – the squint remains the same in all directions of gaze. Diplopia is almost never a symptom. The deviating eye (the one that does not fixate) usually has defective (amblyopic) vision; this is called *amblyopia ex anopsia*. Non-paralytic squint may be latent (e.g. only visible when tired).

The cover test. The cover test is used principally to recognize latent squint. The patient is asked to fix on a light. When there is no squint, a pinpoint reflection is seen in the *exact centre* of each pupil. One eye is covered quickly. If a latent squint is present, the other (uncovered) eye moves to take up central fixation. The test is repeated with the opposite eye. The dominant, fixing eye will not move when the other, squinting, amblyopic eye is covered and uncovered.

Paralytic squint. Paralytic or incomitant squint occurs when there is an acquired defect of movement of an eye – the usual situation in neurological disease. There is a squint (and hence diplopia) maximal in the direction of action of the weak muscle.

III: Oculomotor nerve

The nucleus of the IIIrd nerve lies ventral to the aqueduct in the midbrain. Efferent fibres to four external ocular muscles (superior, inferior and medial recti, and inferior oblique), levator palpebrae superioris and sphincter pupillae (parasympathetic) enter the orbit through the superior orbital fissure. Causes of a IIIrd nerve lesion are listed in Table 21.9.

Signs of a complete IIIrd nerve palsy are:

- unilateral complete ptosis
- the eye facing down and out
- a fixed and dilated pupil.

Sparing of the pupil means that parasympathetic fibres remain undamaged; these run in a discrete bundle on the superior surface of the nerve, thus the pupil is of normal size and reacts normally. In diabetes, IIIrd nerve infarction usually spares the pupil.

In a IIIrd nerve palsy the eye can still ABduct (VIth nerve) and rotate inwards (intort, IVth). Preservation of intorsion (inward rotation) means that the IVth (trochlear) nerve is intact. To test this in a patient with a right IIIrd nerve palsy, ask the patient to look down and converge: conjunctival vessels of the right eye are seen to *twist clockwise*, indicating intorsion and an intact IVth nerve.

IV: Trochlear nerve

This supplies the superior oblique muscle. The patient complains of torsional diplopia (two objects at an angle) when attempting to look down (e.g. descending stairs); the head is tilted away from that side.

Table 21.9 Some causes of a IIIrd nerve lesion

Aneurysm of the posterior communicating artery
Infarction of IIIrd nerve: in diabetes and atheroma
Coning of the temporal lobe
Midbrain infarction
Midbrain tumour, primary or secondary

VI: Abducens nerve

This supplies the lateral rectus muscle (ABduction). There is an evident convergent squint with diplopia maximal looking to the side of the lesion. The eye cannot be ABducted beyond the midline. The VIth nerve has a long intracranial course. It can be damaged in the brainstem (e.g. by MS or pontine glioma). In raised intracranial pressure it is compressed against the tip of the petrous temporal bone. The nerve sheath may be infiltrated by tumours, particularly nasopharyngeal carcinoma. An isolated VIth nerve palsy occurs in diabetes (infarction) and is a common sequel of head trauma.

Complete external ophthalmoplegia

Complete external ophthalmoplegia describes an immobile eye when IIIrd, IVth and VIth nerves are paralysed at the orbital apex (e.g. by metastasis) or within the cavernous sinus (e.g. by sinus thrombosis).

FURTHER READING

Balcer LJ. Optic neuritis. *New England Journal of Medicine* 2006; **354**: 1273–1280.

Miller JW. Gene therapy for retinal degeneration. *New England Journal of Medicine* 2008; **358**: 2282–2284.

Miller NR, Newman NJ. The eye in neurological disease. *Lancet* 2004; **364**: 2045–2054.

V: TRIGEMINAL NERVE

This nerve is large and mainly sensory but has motor fibres.

Sensory fibres (Fig. 21.7; see also Figs 21.11 and 21.12) of the three divisions – ophthalmic (V_1), maxillary (V_2) and mandibular (V_3) – pass to the trigeminal (Gasserian) ganglion within the cavernous sinus at the apex of the petrous temporal bone. Central fibres enter the brainstem. Ascending fibres transmitting light touch enter the Vth nucleus in the pons. Descending central fibres carrying pain and temperature form the spinal tract of V, to end in the spinal Vth nucleus that extends from the medulla into the cervical cord.

Motor fibres arise in the upper pons and join the mandibular branch (V_3) to supply muscles of mastication.

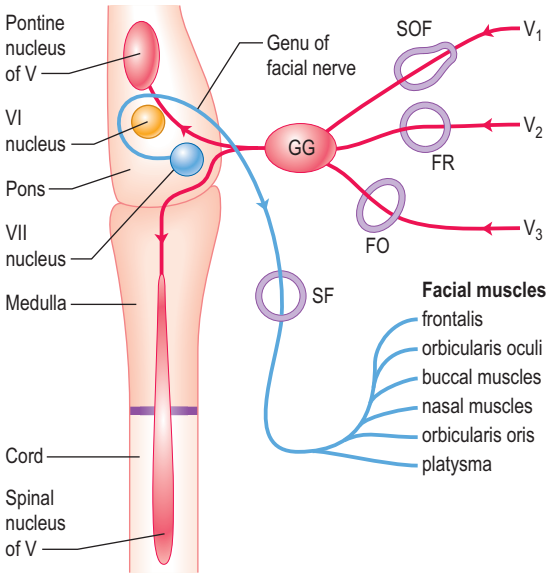


Fig. 21.7 **Sensory input of Vth nerve (red) and motor output of VIIth nerve (blue).** GG, Gasserian ganglion; SOF, superior orbital fissure; FR, foramen rotundum; FO, foramen ovale; SF, stylomastoid foramen.

Signs of a Vth nerve lesion

A complete Vth nerve lesion causes unilateral sensory loss on the face, tongue and buccal mucosa, and the jaw deviates to that side as the mouth opens (motor fibres). Diminution of the corneal reflex is an early and sometimes isolated sign of a Vth nerve lesion.

Central (brainstem) lesions of the lower trigeminal nuclei (e.g. in syringobulbia, p. 1164) produce a characteristic circumoral sensory loss.

When the spinal nucleus alone is involved, sensory loss is restricted to pain and temperature sensation, i.e. dissociated (p. 1117).

Causes

Within the *brainstem*, lesions involve Vth nuclei and central connections, e.g.:

- brainstem glioma
- multiple sclerosis
- infarction
- syringobulbia.

At the *cerebellopontine angle* (CPA), the nerve is compressed by:

- acoustic neuroma
- meningioma
- secondary neoplasm.

As a CPA lesion enlarges, the neighbouring VIIth and VIIIth nerves become involved, producing facial weakness, deafness and vertigo.

At the *apex of the petrous temporal bone*, spreading middle ear infection or a secondary tumour damages the nerve. The combination of a painful Vth with a VIth nerve lesion is called *Gradenigo's syndrome*.

Within the cavernous sinus, the trigeminal (Gasserian) ganglion is compressed by:

- aneurysm of the internal carotid artery
- lateral extension of a pituitary neoplasm or metastasis
- thrombosis of the cavernous sinus.

The trigeminal ganglion itself becomes infected in ophthalmic herpes zoster (p. 1157), the most common lesion of the ganglion. Postherpetic neuralgia commonly follows (p. 1157).

Peripheral branches of V are infiltrated by skull base neoplasms, for example the 'numb chin' seen with a breast cancer metastasis.

Trigeminal neuralgia

Trigeminal neuralgia (*tic douloureux*) develops most commonly in the over 50 years age group and is almost always unilateral. Compression, most commonly from aberrant branches of the superior cerebellar artery in the cerebellopontine angle, causes demyelination. It is thought that this sends ephaptic (short circuited) signals to cause pain.

Features

Paroxysms of knife-like or electric shock-like pain, lasting seconds, occur in the distribution of the Vth nerve. Pain tends to commence in the mandibular division (V_3) and spreads upwards to maxillary (V_2) and ophthalmic divisions (V_1). Spasms occur many times a day, but rarely in sleep. Each paroxysm is stereotyped, brought on by stimulation of one

or more trigger zones in the face. Washing, shaving, a cold wind or eating are examples of trivial stimuli that provoke pain. The face may be screwed up in agony (hence *tic*). Spontaneous remissions last months or years before (almost invariable) recurrence. There are no signs of Vth nerve dysfunction; the corneal reflex is preserved.

Treatment

Carbamazepine 600–1200 mg daily reduces severity of attacks in the majority. Phenytoin, gabapentin and clonazepam are used, but are less effective. If drugs fail, surgery (radiofrequency extirpation of the ganglion, neurovascular decompression or sectioning of the sensory root) is useful. Alcohol injection into the trigeminal ganglion or peripheral Vth nerve branches can also be carried out.

Secondary trigeminal neuralgia

Trigeminal neuralgia occurs in MS (p. 1152), with tumours of the Vth nerve (e.g. neuroma) and with cerebellopontine angle lesions. There are often physical signs – initially a depressed corneal reflex, progressing to trigeminal sensory loss.

Idiopathic trigeminal neuropathy

Chronic isolated Vth nerve lesions sometimes occur without apparent cause. When sensory loss is severe, trophic changes (facial scarring and corneal ulceration) develop.

VII: FACIAL NERVE

The VIIth nerve is largely motor, supplying muscles of facial expression. VII carries sensory taste fibres from the anterior two-thirds of the tongue via the chorda tympani and supplies motor fibres to stapedius. The VIIth nerve (Fig. 21.7) arises from its nucleus in the pons and leaves the skull through the stylomastoid foramen. Neurones in each VIIth nucleus supplying the upper face (principally frontalis) receive supranuclear fibres from each hemisphere.

Unilateral facial weakness

Upper motor neurone (UMN lesions) cause weakness of the lower part of the face on the opposite side. Frontalis is spared: normal furrowing of the brow is preserved; eye closure and blinking are largely unaffected. The earliest sign is slowing of one side of the face, e.g. on baring teeth. There is sometimes relative preservation of spontaneous emotional movement (e.g. smiling) compared with voluntary movement.

Lower motor neurone (LMN) lesions. A complete unilateral LMN VIIth lesion causes weakness (ipsilateral) of all facial expression muscles. The angle of the mouth falls; unilateral dribbling develops. Frowning (frontalis) and eye closure are weak. Corneal exposure and ulceration occur if the eye does not close during sleep. Platysma is weak.

Causes of facial weakness

A common cause of a UMN lesion is a stroke, e.g. internal capsule infarction with hemiparesis.

At lower levels, lesion sites are recognized by LMN weakness with additional signs.

Pons. Here the VIth (abducens) nucleus is encircled by the VIIth nerve (Fig. 21.7), leading to a lateral rectus palsy (convergent squint) with unilateral LMN facial weakness. When the neighbouring PPRF and corticospinal tract are involved, there is the triple combination of:

- LMN facial weakness
- failure of conjugate lateral gaze (towards lesion)
- contralateral hemiparesis.

Causes include pontine tumours (e.g. glioma), MS and infarction.

The facial nucleus itself can be affected in *bulbar* poliomyelitis (p. 110) and in motor neurone disease (p. 1166).

Cerebellopontine angle (CPA). The neighbouring Vth, VIth and VIIIth nerves are compressed with VII in the CPA, e.g. by acoustic neuroma, meningioma or metastasis.

Petrous temporal bone. The geniculate ganglion (sensory for taste) lies at the genu (knee) of VII (Fig. 21.7). Fibres join VII from the chorda tympani – taste from the tongue (anterior two-thirds). The (motor) nerve to stapedius leaves VII distal to the genu. Lesions in this region cause:

- loss of taste on the anterior two-thirds of the tongue
- hyperacusis (loud noise distortion – paralysis of stapedius).

Causes include:

- Bell's palsy
- trauma
- middle ear infection
- herpes zoster (Ramsay Hunt syndrome, p. 1108)
- tumours (e.g. glomus tumour).

Skull base, parotid gland and within the face. Swelling of VII within the stylomastoid foramen develops in Bell's palsy. VII can be compressed by skull base tumours and in Paget's disease of bone. Branches of VII that pierce the parotid are damaged by parotid gland tumours, mumps (p. 118), sarcoidosis (p. 869) and trauma. Each VIIth nerve can be involved in polyneuritis (e.g. Guillain–Barré syndrome, p. 1172), usually symmetrically.

Facial weakness (usually symmetrical) is also seen in muscle and neuromuscular junction disease, e.g.:

- myotonic dystrophy (p. 1181)
- facioscapulohumeral dystrophy (p. 1181)
- myasthenia gravis (p. 1180).

Bell's palsy

This common, acute, isolated facial palsy is usually due to viral infection (often herpes simplex) that causes swelling of VII within the petrous temporal bone or as it traverses the stylomastoid foramen. There is sudden unilateral facial weakness, sometimes with loss of taste on the tongue, and hyperacusis. Pain behind the ear is common at onset. Diagnosis is made on clinical grounds. Other cranial nerves are not involved. Bell's palsy occasionally recurs and is very rarely bilateral. (Bell's *phenomenon* is the upward conjugate eye movement that occurs when the eyes are closed.)

Management and course

Weakness progresses over hours or several days. Spontaneous improvement usually begins during the second week. Thereafter, recovery continues but may take 12 months.

FURTHER READING

Love S, Coakham HB. Trigeminal neuralgia: pathology and pathogenesis. *Brain* 2001; **124**: 2347–2360.

FURTHER READING

Hato N, Murakami S, Gyo K. Steroid and antiviral treatment for Bell's palsy. *Lancet* 2008; **371**: 1818–1820.

Nerve stimulation (EMG, p. 1121) is of help in predicting outcome but seldom necessary. After the third week, absence of evoked potentials from facial muscles indicates that full recovery is unlikely.

Steroids (e.g. prednisolone 60 mg daily, reducing to nil over 10 days) should be given. Antivirals (valaciclovir) and prednisolone are given for severe palsy, starting within 3 days of onset only, although the evidence for giving antivirals is poor. To prevent corneal exposure, adhesive tape to keep the eyelids closed is invaluable; suturing the upper to the lower lid (tarsorrhaphy) is considered if the eye cannot be closed for more than a week. Residual, unsightly paralysis occurs in more than 10%; cosmetic surgery is sometimes helpful.

Ramsay Hunt syndrome

This is herpes zoster (shingles) of the geniculate ganglion. There is a facial palsy (identical to Bell's) with herpetic vesicles around the external auditory meatus and/or the soft palate (sensory twigs from VII). Deafness and a Vth nerve lesion may occur. Complete recovery is less likely than in Bell's. Treatment for shingles should be given (e.g. famciclovir, p. 101).

Hemifacial spasm

This is an irregular, painless clonic spasm of facial muscles, usually occurring after middle age, and more common in women. It varies from a mild inconvenience to a severe, disfiguring spasm.

Causes are:

- idiopathic
- following Bell's palsy
- acoustic neuroma
- Paget's disease within the skull base
- pressure from vessels in the cerebellopontine angle (postulated).

There are clonic spasms of facial muscles on one side. Mild LMN facial weakness is common.

Management

Mild cases require no treatment. Decompression of the VIIth nerve in the CPA is sometimes helpful. Local botulinum toxin injection into facial muscles reduces the spasm for some months, and can be repeated. Drugs (e.g. carbamazepine) are of little value.

Myokymia

Myokymia describes a rare, continuous, fine, sinuous or wave-like movement of the lower face seen in brainstem lesions (e.g. MS, brainstem glioma). Myokymia is also used to describe the common innocent twitching around the eye.

VIII: VESTIBULO-COCHLEAR NERVE

Cochlear nerve

Auditory fibres from the spiral organ of Corti within the cochlea pass to the cochlear nuclei in the pons. Fibres from these nuclei cross the midline and pass upwards via the medial lemnisci to the medial geniculate bodies and then to the temporal gyri.

Symptoms of a cochlear nerve lesion are deafness and tinnitus (p. 1080). The deafness is termed sensorineural (or perceptive). Clinical detection is by tuning fork tests (256 or 512 Hz, not 128 Hz, Rinne's and Weber's) that distinguish conductive from sensorineural deafness (p. 1078).

Basic investigations of cochlear lesions

- Pure tone audiometry and auditory thresholds.
- Auditory evoked potentials (recording responses from repetitive clicks via scalp electrodes; lesion levels determined from response pattern).

Causes of deafness are listed in Tables 21.10 and 20.1 (see p. 1078).

Vestibular nerve

Nerve impulses generated by movement of sensory epithelia within the three semicircular canals, saccule and utricle pass to vestibular nuclei in the pons. Vestibular nuclei are connected to the cerebellum, nuclei of the ocular muscles, PPRF, extrapyramidal system, reticular formation, temporal lobes and spinal cord. Balance and posture also depend upon the interaction of proprioceptive impulses from the neck, spinal muscles and limbs. The *main symptoms* of vestibular lesions are vertigo and loss of balance. Vomiting frequently accompanies any acute vertigo. Nystagmus is the principal sign.

Vertigo

Vertigo, the definite illusion of movement of the subject or surroundings, typically rotatory, indicates a disturbance of the vestibular nerve, brainstem or, very rarely cortical function (Table 21.10). Deafness and tinnitus accompanying vertigo indicate involvement of the ear or cochlear nerve (p. 1078).

Nystagmus

Nystagmus is rhythmic oscillation of eye movement, and a sign of disease in the retina, oculomotor and/or vestibular

Table 21.10 Some causes of vertigo and hearing loss

Benign paroxysmal positional vertigo (V)
Vestibular neuritis (V)
Ménière's disease (V, D)
Alcohol, antiepileptic drug intoxication (V)
Cerebellar lesions (V)
Partial (temporal lobe) seizures (V)
Migraine (V + phonophobia)
Brainstem ischaemia (V, occasionally D)
Multiple sclerosis (V, occasionally D)
Mumps, intrauterine rubella and congenital syphilis (D)
Advancing age (presbyacusis) and otosclerosis (D)
Acoustic trauma (D)
Congenital e.g. Pendred's syndrome (D)
Gentamicin, furosemide (V, D)
Middle and external ear disease (V, D)
Cerebellopontine angle lesions e.g. acoustic neuroma (V, D)
Carcinomatous meningitis, sarcoid and tuberculous meningitis (V, D)

V, Vertigo; D, hearing loss.

systems and their connections. Nystagmus is either *jerk* or *pendular*. Nystagmus must be sustained within binocular gaze to be of diagnostic value – a few beats at the extremes of gaze are normal.

Jerk nystagmus

Jerk nystagmus (usual in neurological disease) is a fast/slow oscillation. This is seen in vestibular end-organ, VIIIth nerve, brainstem, cerebellar and (very rarely) cortical lesions. Direction of nystagmus is decided by the fast component, a reflex attempt to correct the slower, primary movement. Nystagmus is both a common and valuable indication of abnormality as a whole but there are difficulties using the direction of jerk nystagmus as a localizing sign. The following are useful starting points:

- **Horizontal or rotary jerk nystagmus** may be either of peripheral (middle ear) or central origin (VIIIth nerve, brainstem, cerebellum and connections).
  - In peripheral lesions, nystagmus is usually acute and transient (minutes or hours) and associated with severe prostrating vertigo.
  - In central lesions nystagmus tends to be long-lasting (weeks, months or more). Vertigo caused by central lesions tends to wane after days or weeks, the nystagmus outlasting it.
- **Vertical jerk nystagmus** is caused typically by central lesions.
- **Down-beat jerk nystagmus** is a rarity caused by lesions around the foramen magnum (e.g. meningioma, cerebellar ectopia).

Pendular nystagmus

Pendular describes movements to and fro, similar in velocity and amplitude. Pendular nystagmus is almost always binocular, horizontal and present in all directions of gaze. The causes are usually ocular (e.g. poor visual fixation from long-standing, severe visual impairment), or congenital, sometimes associated with head-nodding. Fine pendular, jelly nystagmus is seen (rarely) in MS (brainstem plaque) and brainstem glioma.

Other ocular movement disorders include ocular bobbing and opsoclonus – disorganized dancing eye movements seen typically in children with neuroblastoma and as a postinfective phenomenon (p. 455).

Basic investigation of vestibular problems

High-definition MRI provides the best structural imaging. Hallpike's test (thrusting the head quickly to one side, p. 1079) may provoke nystagmus. Caloric tests assess labyrinth function, recording duration of evoked nystagmus when ice-cold, then warm, water is run into the external meatus. In the normal caloric test:

- ice-cold water in the left ear causes nystagmus with the fast movement to the right
- warm water in the left ear causes nystagmus with the fast movement to the left.

Decreased or absent nystagmus indicates ipsilateral labyrinth, VIIIth nerve or central involvement. Caloric tests are also used in confirmation of brainstem death (p. 920). There are difficulties relating abnormalities in caloric tests to clinical symptoms of dizziness. Other specialized neuro-otological tests are available. Vestibular positioning techniques (e.g. the Epley manouevre) and exercises involving neck movement (e.g. Cawthorne–Cooksey exercises) are sometimes useful treatments.

Vestibular and auditory lesions – central, VIIIth nerve and end organ

Drugs (e.g. anticonvulsant toxicity), alcohol intoxication, MS and brainstem vascular disease are common central causes of vertigo. It is possible to recognize lesions from clinical features at six levels.

Cerebral cortex. Vertigo can be part of the aura of a partial (temporal lobe) seizure. Vertigo is also a psychological and perceptual sensation at heights, in panic attacks and agoraphobia. Deafness is rare in acquired cortical disease; bilateral lesions are necessary.

Pons and brainstem. Vertigo is common when lesions (MS, vascular, tumour, syrinx) involve vestibular nuclei and their connections. A VIth or VIIth nerve lesion, internuclear ophthalmoplegia, lower cranial nerve lesions or contralateral hemiparesis help localization. Nystagmus is frequently present, while deafness is rare. Transient vertigo occurs in basilar migraine (p. 1137), preceding syncope, and occasionally in hypoglycaemic attacks. Its site of origin is often difficult to ascertain.

Cerebellum. Nystagmus develops towards the side of a cerebellar mass (tumour, haemorrhage, infarct or abscess). Limb ataxia is usually present. Bilateral cerebellar, or cerebellar connection disease (e.g. MS) causes bilateral nystagmus. Deafness does not occur.

Cerebellopontine angle. Ipsilateral sensorineural deafness, tinnitus and vertigo occur, with emergence of VIth, VIIth and Vth nerve lesions, followed by cerebellar signs (ipsilateral) and later pyramidal signs (contralateral). Nystagmus is often present. Causes include acoustic neuroma (p. 1160), meningioma and metastases, carcinomatous meningitis and inflammatory lesions (Table 21.10).

Petrous temporal bone. Facial weakness often accompanies the VIIIth nerve lesion. Causes include trauma, middle ear infection, secondary neoplasm and Paget's disease of bone (see Gradenigo's syndrome, p. 1106).

End organs – cochlea, semicircular canals, middle and external ear.

Vestibular neuronitis

Vestibular neuronitis is a common, poorly understood problem. It is an acute attack of isolated vertigo with nystagmus, often with vomiting, believed to follow viral infections. The disturbance lasts for several days or weeks, is self-limiting and rarely recurs. Vestibular neuritis is sometimes followed by benign positional vertigo (p. 1079). Deafness is absent. Acute treatment is with vestibular sedatives. Similar symptoms can be caused by MS or brainstem vascular lesions. Other signs are usually apparent.

FURTHER READING

Baloh RW. Vestibular neuritis. *New England Journal of Medicine* 2003; **348**: 1027–1032.

LOWER CRANIAL NERVES IX, X, XI, XII

The glossopharyngeal (IX), vagus (X) and accessory (XI) nerves arise in the medulla and leave the skull through the jugular foramen. The hypoglossal (XII) arises in the medulla, to leave the skull base via the anterior condylar foramen. Outside the skull, the four cranial nerves lie together, close to the carotid artery and sympathetic trunk.

Glossopharyngeal (IX)

This nerve is largely sensory, supplying sensation from the tonsillar fossa and pharynx (afferent pathway of gag reflex), taste from the tongue (posterior third) and fibres from the carotid sinus. Motor fibres supply stylopharyngeus, and autonomic fibres the parotid.

Vagus (X)

The vagus is a mixed nerve, largely motor, which supplies striated muscle of the pharynx (efferent gag reflex pathway), larynx (including vocal cords via recurrent laryngeal nerves) and upper oesophagus. There are sensory fibres from the larynx. Parasympathetic fibres supply the heart and abdominal viscera.

Accessory (XI)

The accessory nerve, a complex motor nerve, supplies trapezius and sternomastoid.

Hypoglossal (XII)

The hypoglossal nerve is a motor nerve to tongue muscles.

IXth and Xth nerve lesions

Principal causes of IXth, Xth, XIth and XIIth nerve lesions are listed in Table 21.11.

Isolated lesions of IXth and Xth nerves are unusual, since disease at the jugular foramen affects both nerves and sometimes XI.

A unilateral IXth nerve lesion causes diminished sensation on the same side of the pharynx, hard to recognize in isolation. A Xth nerve palsy produces ipsilateral failure of voluntary and reflex elevation of the soft palate, which is drawn to the opposite side, and ipsilateral vocal cords.

Bilateral lesions of IXth and Xth nerves cause weakness of palatal elevation and depression of palatal sensation with depression of the gag reflex, dysphonia and choking. *Bulbar palsy* is a general term describing palatal, pharyngeal and tongue weakness of LMN or muscle origin.

Table 21.11 Principal causes of IXth, Xth, XIth and XIIth nerve lesions

Within brainstem
Infarction
Syringobulbia
Motor neurone disease (motor fibres)
Poliomyelitis (motor fibres)
Around skull base
Carcinoma of nasopharynx
Glomus tumour
Neurofibroma
Jugular venous thrombosis
Trauma
Within neck and nasopharynx
Carcinoma of nasopharynx
Metastases
Carotid artery dissection (XII)
Polyneuropathy
Trauma
Lymph node biopsy in posterior triangle (XI)

Recurrent laryngeal nerve lesions. Paralysis of this branch of each vagus causes hoarseness (dysphonia) and failure of the forceful, explosive part of voluntary and reflex coughing. There is no visible palatal weakness; vocal cord paralysis is seen endoscopically. Bilateral acute lesions (e.g. postoperatively) cause respiratory obstruction – an emergency.

The left recurrent laryngeal nerve (looping beneath the aorta) is damaged more commonly than the right.

Causes of recurrent laryngeal nerve lesions include:

- mediastinal primary tumours (e.g. thymoma)
- secondary spread from bronchial carcinoma
- aortic aneurysm
- trauma or surgery of neck or thorax.

Glossopharyngeal neuralgia. This is intensely painful, paroxysmal neuralgic spasm within the pharynx triggered repeatedly by swallowing. There are no physical signs. Treatment of this rarity is with carbamazepine (p. 1107) or IXth nerve section.

XIth nerve lesions

XIth nerve lesions cause weakness of sternomastoid (rotation of the head and neck to the opposite side) and trapezius (shoulder shrugging). Nerve section (e.g. following lymph node biopsy in the neck posterior triangle) is followed by persistent neuralgic pain.

XIIth nerve lesions

LMN lesions of XII lead to unilateral tongue weakness, wasting and fasciculation. The protruded tongue deviates towards the weaker side. Bilateral supranuclear (UMN) lesions produce slow, limited tongue movements and a stiff tongue that cannot be protruded far. Fasciculation is absent.

BRAINSTEM LESIONS

Bulbar and pseudobulbar palsy

Bulbar palsy describes LMN weakness of muscles whose cranial nerve nuclei lie in the medulla (the bulb). Paralysis of bulbar muscles is caused by disease of lower cranial nerve nuclei, lesions of IXth, Xth and XIIth nerves (Table 21.11), malfunction of their neuromuscular junctions (e.g. myasthenia gravis, botulism) or disease of muscles themselves (e.g. dystrophies).

Pseudobulbar palsy describes bilateral supranuclear (UMN) lesions of lower cranial nerves producing weakness of the tongue and pharyngeal muscles. This resembles, superficially, a bulbar palsy, hence pseudobulbar. Findings are a stiff, slow, spastic tongue (not wasted), dysarthria with a stiff, slow voice sounding dry and gravelly, and dysphagia. Gag and palatal reflexes are preserved and the jaw jerk exag-

Ask the authors
www.studentconsult.com

gerated. Emotional lability (inappropriate laughing or crying) often accompanies pseudobulbar palsy.

Principal causes are:

- motor neurone disease – often both UMN and LMN lesions (i.e. elements of both pseudobulbar and bulbar palsy)
- MS, mainly as a late event
- cerebrovascular disease, typically following multiple infarcts
- following severe brain injury.

Difficulty swallowing, dysarthria and a slow tongue also develop in late stages of Parkinson's disease.

MOTOR CONTROL SYSTEMS

There are three systems, each of which interacts by feedback loops with the other two, with sensory input and the reticular formation:

- The *corticospinal* (or *pyramidal*) system enables *purposive, skilled, intricate, strong* and *organized* movements. Defective function is recognized by a distinct pattern of signs – loss of skilled voluntary movement, spasticity and reflex change – seen, for example in a hemiparesis, hemiplegia or paraparesis.
- The *extrapyramidal* system facilitates *fast*, *fluid* movements that the corticospinal system has generated. Defective function produces slowness (bradykinesia), stiffness (rigidity) and/or disorders of movement (rest tremor, chorea and other dyskinesias). One feature (e.g. stiffness, tremor or chorea) will often predominate.
- The *cerebellum* and its connections have a role coordinating *smooth* and *learned* movement, initiated by the pyramidal system, and in posture and balance control. Cerebellar disease leads to unsteady and jerky movements (*ataxia*), with characteristic limb signs of past pointing, action tremor and incoordination, gait ataxia and/or truncal ataxia.

CORTICOSPINAL (PYRAMIDAL) SYSTEM

The corticospinal tracts originate in neurones of the cortex and terminate at motor nuclei of cranial nerves and spinal cord anterior horn cells. The pathways of particular clinical significance (Fig. 21.8) congregate in the internal capsule and cross in the medulla (decussation of the pyramids), passing to the contralateral cord as the lateral corticospinal tracts. This is the pyramidal system, disease of which causes upper motor neurone (UMN) lesions. 'Pyramidal' is simply a descriptive term that draws together anatomy and characteristic physical signs, and is used interchangeably with the term UMN.

A proportion of the corticospinal outflow is uncrossed (anterior corticospinal tracts). This is of no relevance in practice.

Characteristics of pyramidal lesions

(Table 21.12)

Signs of an early pyramidal lesion may be minimal. Weakness, spasticity or changes in superficial reflexes can predominate, or be present in isolation.

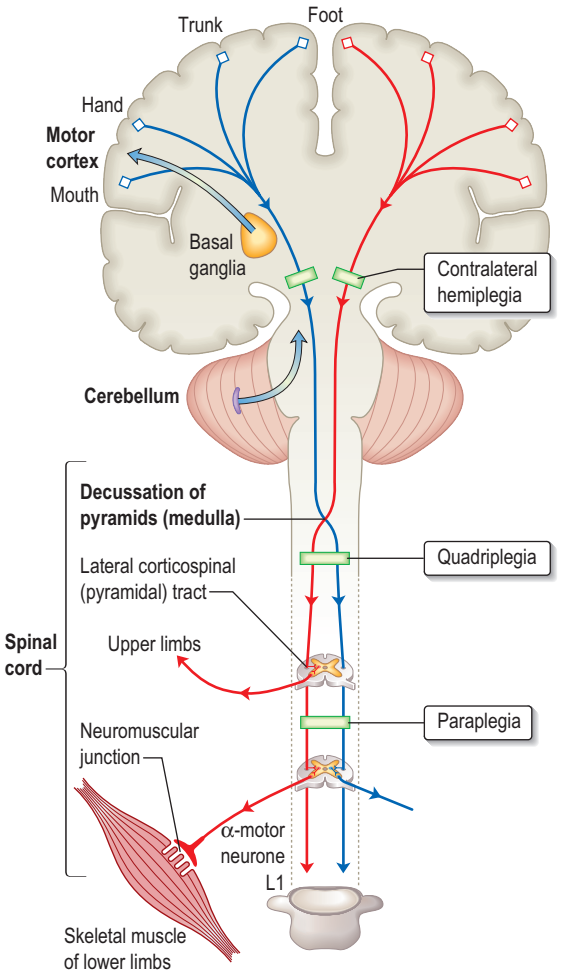


Fig. 21.8 **The motor system.**

Table 21.12 Features of upper motor neurone lesions

Drift of upper limb
Weakness with characteristic distribution
Changes in tone: flaccid→spastic
Exaggerated tendon reflexes
Extensor plantar response
Loss of skilled finger/toe movements
Loss of abdominal reflexes
No muscle wasting
Normal electrical excitability of muscle

Pyramidal drift of an upper limb

Normally, the outstretched upper limbs are held symmetrically, when the eyes are closed. With a pyramidal lesion, when both upper limbs are held outstretched, palms uppermost, the affected limb drifts downwards and medially. The forearm tends to pronate and the fingers flex slightly. This sign is often first to emerge, sometimes before weakness and/or reflex changes become apparent.

Weakness and loss of skilled movement

A unilateral pyramidal lesion above the decussation in the medulla causes weakness of the opposite limbs. When acute and complete, this weakness will be immediate and total, e.g. a hemiplegia following an internal capsule infarct. With slowly

progressive lesions (e.g. a hemisphere glioma) a characteristic pattern of weakness emerges – a hemiparesis.

In the upper limb, flexors remain stronger than extensors, whereas in the lower limb, extensors remain stronger than flexors. In the upper arm, weaker movements are thus shoulder abduction and elbow extension; in the forearm and hand, wrist and finger extensors and abductors are weaker than their antagonists. In the lower limb, weaker movements are hip flexion and abduction, knee flexion, ankle dorsiflexion and eversion. There is also loss of skilled movement – fine finger and toe control diminishes. Wasting (except from disuse) is not a feature. Muscles remain normally excitable electrically.

When a UMN lesion is below the decussation of the pyramids, e.g. in the cervical cord, hemiparesis is on the same side as the lesion, an unusual situation.

Changes in tone and tendon reflexes

An acute lesion of one pyramidal tract (e.g. internal capsule stroke) causes initially flaccid paralysis with loss of tendon reflexes. Increase in tone follows, usually within several days due to loss of inhibitory effects of the corticospinal pathways and an increase in spinal reflex activity. This increase in tone (spasticity) is detectable most easily in stronger muscles. Spasticity is characterized by sudden *changing resistance* to passive movement – the clasp-knife effect. Relevant tendon reflexes become exaggerated; clonus may emerge.

Changes in superficial reflexes

The normal flexor plantar response becomes extensor (a positive Babinski). In a severe acute lesion, this extensor response can be elicited from a wide area of the foot. As recovery progresses, the receptive area diminishes until the lateral posterior third of the sole remains receptive to an orange-stick stimulus (an appropriate instrument). An extensor plantar is certain when great toe dorsiflexion is accompanied by abduction of adjacent toes. Abdominal (and cremasteric) reflexes are abolished on the side affected.

Patterns of UMN disorders

There are three main patterns:

- *Hemiparesis* means weakness of the limbs on one side; it is usually caused by a lesion in the brain and occasionally in the cord.
- *Paraparesis* means weakness of both lower limbs and is usually diagnostic of a cord lesion; bilateral brain lesions occasionally cause paraparesis.
- *Tetraparesis* (syn. *quadriparesis*) means weakness of four limbs.

Hemiplegia, paraplegia and tetraplegia indicate (strictly) *total* paralysis, but are often used to describe severe weakness.

Hemiparesis

The level within the corticospinal system is recognized by particular features.

Motor cortex. Weakness and/or loss of skilled movement confined to one contralateral limb (an arm or a leg – monoparesis) or part of a limb (e.g. a clumsy hand) is typical of an isolated motor cortex lesion (e.g. a secondary neoplasm). A defect in cognitive function (e.g. aphasia) and focal epilepsy may occur.

Table 21.13 Causes of spastic paraparesis

Spinal lesions
Spinal cord compression (Table 21.47)
Multiple sclerosis
Myelitis, e.g. varicella zoster
Motor neurone disease
Subacute combined degeneration of the cord
Syringomyelia
Syphilis
Familial or sporadic paraparesis
Vascular, e.g. cord infarction, arteriovenous malformation
Paraneoplastic syndromes
Tropical spastic paraparesis
HIV-associated myelopathy
Rarities, e.g. lathyrism, copper deficiency

Cerebral lesions (rare)
Parasagittal cortical lesions:
  meningioma
  sagittal sinus thrombosis
Hydrocephalus
Multiple cerebral infarction

Internal capsule. Corticospinal fibres are tightly packed in the internal capsule (about 1 cm^2), thus a small lesion causes a large deficit. A middle cerebral artery branch infarction (p. 1129) produces a sudden, dense, contralateral hemiplegia.

Pons. A pontine lesion (e.g. an MS plaque) is rarely confined to the corticospinal tract. Adjacent structures, e.g. VIth and VIIth nuclei, MLF and PPRF (p. 1104) are involved – diplopia, facial weakness, internuclear ophthalmoplegia (INO) and/or a lateral gaze palsy occur with contralateral hemiparesis.

Spinal cord. An isolated lesion of one lateral corticospinal tract (e.g. a cervical cord injury) causes an ipsilateral UMN lesion, the level indicated by changes in reflexes (e.g. absent biceps, C5/6), features of a Brown–Séquard syndrome (p. 1118) and muscle wasting at the level of the lesion (p. 1163).

Paraparesis (Table 21.13)

Paraparesis indicates bilateral damage to corticospinal pathways. Cord compression (p. 1163) or cord diseases are the usual causes; cerebral lesions occasionally produce paraparesis. Paraparesis is a feature of many neurological conditions; finding the cause is crucial (p. 1164).

EXTRAPYRAMIDAL SYSTEM

The extrapyramidal system is a general term for basal ganglia motor systems, i.e. corpus striatum (caudate nucleus + globus pallidus + putamen), subthalamic nucleus, substantia nigra and parts of the thalamus. In basal ganglia/extrapyramidal disorders, two features (either or both) become apparent, in limbs and axial muscles:

- reduction in speed (bradykinesia, meaning slow movement) or akinesia (no movement), with muscle rigidity
- involuntary movements (e.g. tremor, chorea, hemiballismus, athetosis, dystonia).

Extrapyramidal disorders are classified broadly into *akinetic–rigid syndromes* (p. 1145) where poverty of movement pre-

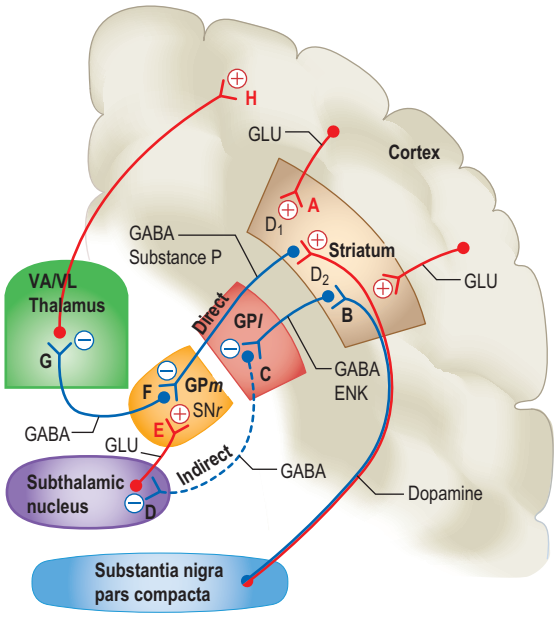


Fig. 21.9 **Extrapyramidal system: connections and neurotransmitters.** The inhibitory pathways are in blue (B, C, D, F, G) and excitory in red. VA/VL, ventral anterior and ventro-lateral thalamic nuclei. GP*l*, lateral globus pallidus. GP*m*: medial globus pallidus. SN*r*, substantia nigra pars reticulata. GLU, glutamate; ENK, enkephalin; GABA, gamma-aminobutyric acid.

dominates, and *dyskinesias* where there are involuntary movements (p. 1146).

The most common extrapyramidal disorder is Parkinson's disease.

Essential anatomy

The corpus striatum lies close to the substantia nigra, thalami and subthalamic nuclei, lateral to the internal capsule (Figs 21.8 and 21.9).

Function and dysfunction

Overall function of this system is modulation of cortical motor activity by a series of servo loops between cortex and basal ganglia (Fig. 21.9). In involuntary movement disorders there are specific changes in neurotransmitters (Table 21.14) rather than focal lesions seen on imaging or at autopsy.

Proposed model of principal pathways

1. Direct pathway from striatum to medial globus pallidus (GP*m*) and substantia nigra pars reticulata (SN*r*). Inhibitory **synapse F**, GABA and substance P.
2. Indirect pathway from striatum to globus pallidus; via lateral globus pallidus (GP*l*; inhibitory **synapse C**, GABA, enkephalin) and subthalamic nucleus inhibitory **synapse D**, GABA). Terminates in GP*m*–SN*r* (in excitatory **synapse E**, glutamate).
3. Direct pathways, both inhibitory and excitatory from substantia nigra pars compacta (SN*c*) to striatum. **Synapse A**, dopamine, D_1, excitatory; and **synapse B**, D_2, inhibitory.
4. GP*m* and SN*r* to thalamus. Inhibitory **synapse G**, GABA.
5. Thalamus to cortex. Excitatory, **synapse H**.
6. Cortex to striatum. Excitatory, glutamate.

Table 21.14 Changes in neurotransmitters in Parkinson's and Huntington's diseases

Condition	Site	Neurotransmitter
Parkinson's	Putamen	Dopamine ↓ 90% Norepinephrine (noradrenaline)↓ 60% 5-HT ↓ 60%
	Substantia nigra	Dopamine ↓90% GAD + GABA ↓↓
	Cerebral cortex	GAD + GABA ↓↓
Huntington's	Corpus striatum	Acetylcholine ↓↓ GABA ↓↓ Dopamine: normal GAD + GABA ↓↓

GABA, γ-aminobutyric acid; GAD, glutamic acid decarboxylase, the enzyme responsible for synthesizing GABA; 5-HT, 5-hydroxytryptamine

The model helps explain how basal ganglia disease can either reduce excitatory thalamo-cortical activity at synapse H, i.e. movement – causing bradykinesia, or increase it, causing dyskinesias.

Parkinson's disease (PD). This is characterized by slowness, stiffness and rest tremor (p. 1146). Degeneration in SN*c* causes loss of dopamine activity in the striatum. Dopamine is excitatory for synapse A and inhibitory for synapse B. Through the direct pathway there is reduced activity at synapse F, leading to increased inhibitory output (G) and decreased cortical activity (H).

Also in PD, in the indirect pathway, dopamine deficiency results in disinhibition of neurones synapsing at C. This leads to reduced activity at D, and to increased activity of neurones in the subthalamic nucleus. There is excess stimulation at synapse E, enhancing further inhibitory output of GP*m*–SN*r*.

The net effect via both pathways is to inhibit the ventral anterior (VA) and ventrolateral (VL) nuclei of the thalamus at synapse G. Cortical (motor) activity at H is thus reduced.

Levodopa helps slowness and tremor in PD (p. 1147) but induces unwanted dyskinesias by increasing dopamine activity at synapses A and B, it is thought by reversing sequences in both direct and indirect pathways.

Huntington's disease (HD). HD is an inherited dementia (p. 1149) with progressive chorea. Chorea results from damage to neurones (GABA, enkephalin) in the indirect pathway from striatum to GP*l*, reducing activity at synapse C. In turn, there is increased inhibition of subthalamic neurones at D, reduced stimulation at E and decreased inhibition of VA/VL at G. Cortical activity at H increases.

Hemiballismus. Wild, flinging (ballistic) limb movements are caused by a lesion in the subthalamic nucleus, typically an infarct. This reduces excitatory activity at synapse E, reduces inhibition at G, with increased thalamo-cortical neuronal activity, and increases activity at H.

CEREBELLUM

The third system of motor control modulates coordination and learned movement patterns, rather than speed. Ataxia, i.e. unsteadiness, is characteristic.

The cerebellum receives afferents from:

- proprioceptive receptors (joints and muscles)
- vestibular nuclei
- basal ganglia
- the corticospinal system
- olivary nuclei.

Efferents pass from the cerebellum to:

- each red nucleus
- vestibular nuclei
- basal ganglia
- corticospinal system.

Each lateral cerebellar lobe coordinates movement of the ipsilateral limbs. The vermis (a midline structure) is concerned with maintenance of axial (midline) posture and balance.

Cerebellar lesions (Table 21.15)

Expanding lesions obstruct the aqueduct to cause hydrocephalus, with severe pressure headaches, vomiting and papilloedema. Coning of the cerebellar tonsils (p. 1161) through the foramen magnum leads to respiratory arrest, sometimes within minutes/hours. Rarely, tonic seizures (attacks of limb stiffness) occur.

Lateral cerebellar lobes

A lesion within one cerebellar lobe (e.g. tumour or infarction) causes disruption of the normal sequence of movements (dyssynergia) on the same side.

Posture and gait. The outstretched arm is held still in the early stages of a cerebellar lobar lesion (cf. the drift of pyramidal lesions) but there is rebound upward overshoot when the limb is pressed downwards and released. Gait becomes broad and ataxic, faltering towards the side of the lesion.

Tremor and ataxia. Movement is imprecise in direction, force and distance (dysmetria). Rapid alternating movements (tapping, clapping or rotary hand movements) become disorganized (dysdiadochokinesis). Intention tremor (action tremor with past-pointing) is seen, but speed of fine movement is preserved, cf. extrapyramidal and pyramidal lesions.

Nystagmus. Coarse horizontal nystagmus (p. 1109) develops with a lateral cerebellar lobe lesion. The fast component is always towards the side of the lesion.

Dysarthria. Halting, jerking speech develops – scanning speech.

Other signs. Titubation – rhythmic head tremor as either forward and back (yes–yes) movements or rotary (no–no) movements – can occur, mainly when cerebellar connections are involved (e.g. in essential tremor and MS, pp. 1149 and 1151). Hypotonia (floppy limbs) and depression of reflexes (with slow, pendular reflexes) are also sometimes seen.

Midline cerebellar lesions

Cerebellar vermis lesions have dramatic effects on trunk and axial muscles. There is difficulty standing and sitting unsupported (truncal ataxia), with a rolling, broad, ataxic gait. Lesions of the flocculonodular region cause vertigo and vomiting with gait ataxia if they extend to the roof of the IVth ventricle. Table 21.15 summarizes the main causes of cerebellar disease.

Table 21.15 Principal causes of cerebellar syndromes

Tumours	Haemangioblastoma Medulloblastoma Secondary neoplasm Compression by acoustic neuroma
Vascular	Haemorrhage Infarction Arteriovenous malformation
Infection	Abscess HIV Prion diseases
Developmental	Arnold–Chiari malformation Basilar invagination Cerebral palsy
Toxic and metabolic	Antiepileptic drugs Chronic alcohol abuse Carbon monoxide poisoning Lead poisoning Solvent abuse
Inherited	Friedreich's and other spinocerebellar ataxias Ataxia telangiectasia Essential tremor
Miscellaneous	Multiple sclerosis Hydrocephalus Postinfective (childhood) Hypothyroidism Paraneoplastic syndromes High altitude cerebral oedema

TREMOR

Tremor means a regular and sinusoidal oscillation of the limbs, head or trunk.

Postural tremor

Everyone has a physiological tremor (often barely perceptible) of the outstretched hands at 8–12 Hz. This is increased with anxiety, caffeine, hyperthyroidism and drugs (e.g. sympathomimetics, sodium valproate, lithium) and occurs in mercury poisoning. A coarser, postural tremor is seen in benign essential tremor (usually at 5–8 Hz) and in chronic alcohol abuse.

Intention tremor

Tremor exacerbated by action, with past-pointing and accompanying incoordination of rapid alternating movement (dysdiadochokinesis), occurs in cerebellar lobe disease and with lesions of cerebellar connections. Titubation and nystagmus may be present.

Rest tremor

Seen typically in Parkinson's disease, this tremor is noticeably worse at rest, usually 4–7 Hz (pill-rolling, between thumb and forefinger).

Other tremors

Coarse tremor is seen following lesions of the red nucleus (e.g. infarction, multiple sclerosis) and rarely with frontal lesions.

LOWER MOTOR NEURONE (LMN) LESIONS

The LMN is the pathway from anterior horn cell (or cranial nerve nucleus) via a peripheral nerve to muscle motor endplates. The motor unit consists of one anterior horn cell, its single fast-conducting axon that leaves the cord via the anterior root, and the group of muscle fibres (100–2000) supplied via the nerve. Anterior horn cell activity is modulated by impulses from:

- corticospinal tracts
- extrapyramidal system
- cerebellum
- afferents via posterior roots.

Signs of lower motor neurone lesions

These are seen in voluntary muscles that depend upon an intact nerve supply both for contraction and metabolic integrity. Signs follow rapidly if the LMN is interrupted (Table 21.16).

Causes

Examples of LMN lesions at various levels are:

- cranial nerve nuclei (Bell's palsy) and anterior horn cell (motor neurone disease)
- spinal root – cervical and lumbar disc protrusion, neuralgic amyotrophy (p. 1177)
- peripheral (or cranial) nerve – trauma, entrapment (p. 1171), polyneuropathy (p. 1172).

SPINAL REFLEX ARC

Components are illustrated in Figure 21.10. The stretch reflex is the physiological basis for all tendon reflexes. In the knee jerk, a tap on the patellar tendon activates stretch receptors in the quadriceps. Impulses in first-order sensory neurones pass directly to LMNs (L3 and L4) that contract quadriceps. Loss of a tendon reflex is caused by a lesion anywhere along the spinal reflex path. The reflex lost indicates its level (Table 21.17).

Table 21.16 Features of lower motor neurone lesions

Weakness
Wasting
Hypotonia
Reflex loss
Fasciculation
Fibrillation potentials (EMG)
Muscle contractures
Trophic changes in skin and nails

Table 21.17 Spinal levels, recording/interpretation of tendon reflexes

Level	Reflex	Symbol for reflex	Meaning
C5–6	Supinator	0	Absent
C5–6	Biceps	+/–	Present with reinforcement
C7	Triceps	+	Normal
L3–4	Knee	++	Brisk, normal
S1	Ankle	+++	Exaggerated (abnormal)
		CL	Clonus

Reinforcement. Distraction of the patient's attention, clenching teeth or pulling interlocked fingers enhances reflex activity. Such reinforcement manoeuvres should be done before a reflex is recorded as absent.

SENSORY PATHWAYS AND PAIN

Peripheral nerves and spinal roots

Peripheral nerves carry all modalities of sensation from either free or specialized nerve endings to dorsal roots and thence to the cord. Sensory distribution of spinal roots (dermatomes) is shown in Figure 21.11.

Spinal cord

Posterior columns

Axons in the posterior columns whose cell bodies are in the ipsilateral gracile and cuneate nuclei in the medulla carry sensory modalities of vibration, joint position (proprioception), light touch and two-point discrimination. Axons from second-order neurones then cross in the brainstem to form the medial lemniscus, passing to the thalamus (Fig. 21.12).

Spinothalamic tracts

Axons carrying pain and temperature sensation synapse in the dorsal horn of the cord, cross within the cord and pass in the spinothalamic tracts to the thalamus and reticular formation.

Sensory cortex

Fibres from the thalamus pass to the parietal region sensory cortex (Fig. 21.12). Connections exist between the thalamus, sensory cortex and motor cortex.

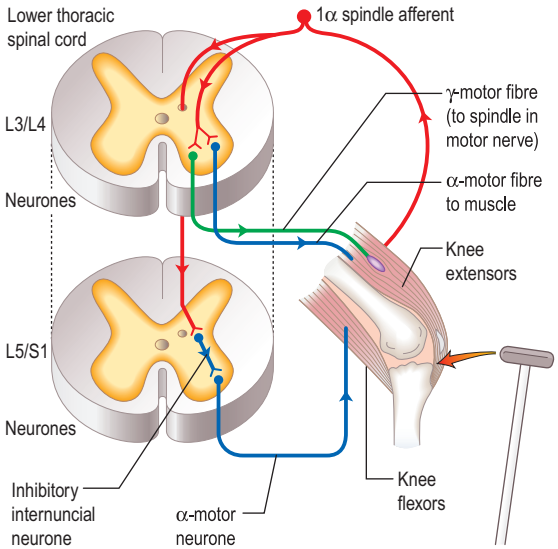


Fig. 21.10 **Knee jerk: a spinal reflex arc.** Sudden patellar tendon stretching generates sensory action potentials in 1α muscle spindle afferents that synapse with γ motor fibres to spindles and α motor fibres. Motor action potentials cause brisk extensor muscle contraction; there is also inhibition of knee flexors.

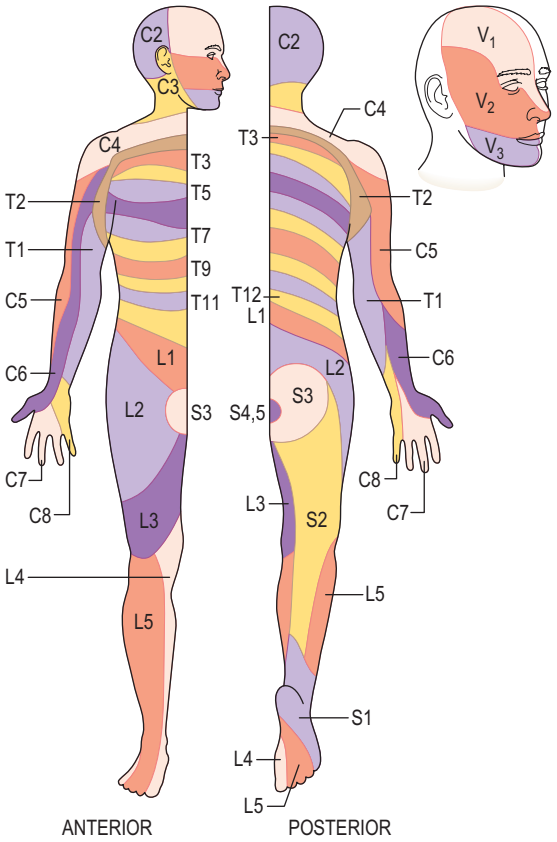


Fig. 21.11 **Dermatomes of spinal roots and Vth nerve.**

LESIONS OF THE SENSORY PATHWAYS

Altered sensation (paraesthesia), tingling, clumsiness, numbness and pain are the principal symptoms of sensory lesions. The pattern and distribution point to the site of pathology (Fig. 21.13).

Peripheral nerve lesions

Symptoms are felt within the distribution of a peripheral nerve (p. 1171). Section of a sensory nerve is followed by complete sensory loss. Nerve entrapment (p. 1171) causes numbness, pain and tingling. Tapping the site of compression sometimes causes a sharp, electric-shock-like pain in the distribution of the nerve, known as Tinel's sign, e.g. in carpal tunnel syndrome (p. 508).

Neuralgia

Neuralgia refers to pain, usually of great severity, in the distribution of a damaged nerve. Examples are:

- trigeminal neuralgia (p. 1106)
- postherpetic neuralgia (p. 1157)
- complex regional pain syndrome type II (causalgia) – chronic burning pain that occasionally follows nerve section.

Spinal root lesions

Root pain

Pain of root compression is felt in the myotome supplied by the root, and there is also a tingling discomfort in the dermatome. The pain is worsened by manoeuvres that either stretch the root (e.g. straight leg raising in lumbar disc prolapse) or increase pressure in the spinal subarachnoid space (coughing and straining). Cervical and lumbar disc protrusions (p. 1177) are common causes of root lesions.

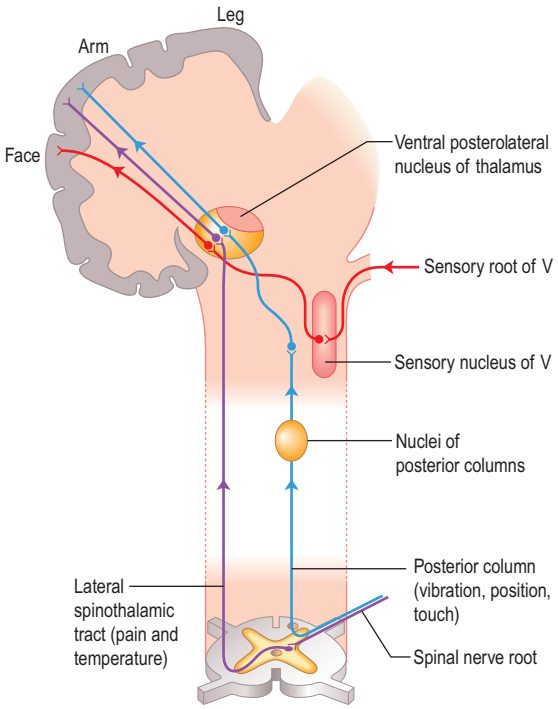


Fig. 21.12 **Principal sensory pathways.** Posterior columns remain uncrossed until the medulla. Spinothalamic tracts cross close to their entry into the cord.

Dorsal spinal root lesions

Section of a dorsal root causes loss of all modalities of sensation within a dermatome (Fig. 21.11). However, overlap between adjacent dermatomes makes it difficult to detect anaesthesia when a single root is destroyed.

Lightning pains. Tabes dorsalis (rare in the UK) is a form of neurosyphilis that causes low-grade inflammation of dorsal roots and spinal cord root entry zones. Irregular, sharp, stabbing pains (like lightning) involve one or two spots, typically in a calf, thigh or ankle.

Spinal cord lesions

Posterior column lesions

These cause:

- tingling
- electric-shock-like sensations
- clumsiness
- numbness
- band-like sensations.

These symptoms, though lateralized, are often felt vaguely without a clear sensory level. Position sense, vibration sense, light touch and two-point discrimination are diminished below the lesion. Position sense loss produces a stamping gait (sensory ataxia, p. 1097).

Lhermitte's phenomenon

Electric-shock-like sensations radiate down the trunk and limbs on neck flexion. This points to a cervical cord lesion.

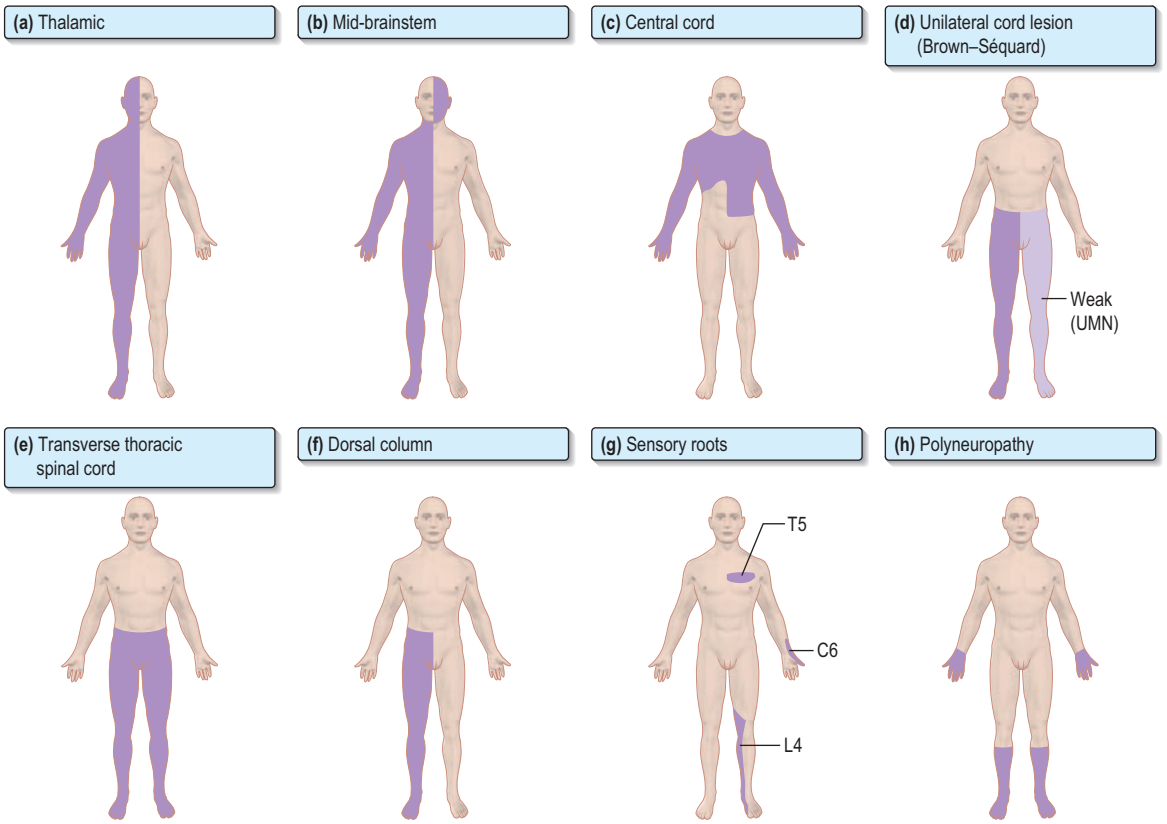


Fig. 21.13 **Principal patterns of loss of sensation. (a) Thalamic lesion:** sensory loss throughout opposite side (rare). **(b) Brainstem lesion:** contralateral sensory loss below face and ipsilateral loss on face. **(c) Central cord lesion,** e.g. syrinx: 'suspended' areas of loss, often asymmetrical and 'dissociated', i.e. pain and temperature loss but light touch intact. **(d) Hemisection of cord/unilateral cord lesion** = Brown–Séquard syndrome: contralateral spinothalamic (pain and temperature) loss with ipsilateral weakness and dorsal column loss below lesion. **(e) Transverse cord lesion:** loss of all modalities, including motor, below lesion. **(f) Dorsal column lesion,** e.g. MS: loss of proprioception, vibration and light touch. **(g) Individual sensory root lesions,** e.g. C6, T5, L4. **(h) Polyneuropathy:** distal sensory loss.

Lhermitte's is common in acute exacerbations of MS (p. 1152), and also occurs in cervical myelopathy (p. 1176), subacute combined degeneration of the cord (p. 1175), radiation myelopathy (p. 1165) and cord compression.

Spinothalamic tract lesions

Pure spinothalamic spinal lesions cause contralateral loss of pain and temperature sensation with a clear level below the lesion. This is called *dissociated* sensory loss – pain and temperature are dissociated from light touch, which remains preserved. This is seen typically in syringomyelia where a cavity occupies the central cord (p. 1164).

The spinal level is modified by lamination of fibres within the spinothalamic tracts. Fibres from lower spinal roots lie superficially and are damaged first by compressive lesions from outside the cord. As an external compressive lesion (e.g. a midthoracic extradural meningioma; Fig. 21.14) enlarges, the spinal sensory level ascends as deeper fibres become involved. Conversely, a central cord lesion (e.g. a syrinx, p. 1165) affects deeper fibres first. Spinothalamic tract lesions cause loss of pain and temperature perception (e.g. painless burns). Perforating ulcers and neuropathic (Charcot) joints develop.

Spinal cord compression (Fig. 21.14)

Cord compression causes progressive spastic paraparesis (or tetraparesis/quadriparesis) with sensory loss below the level of compression. Sphincter disturbance is common.

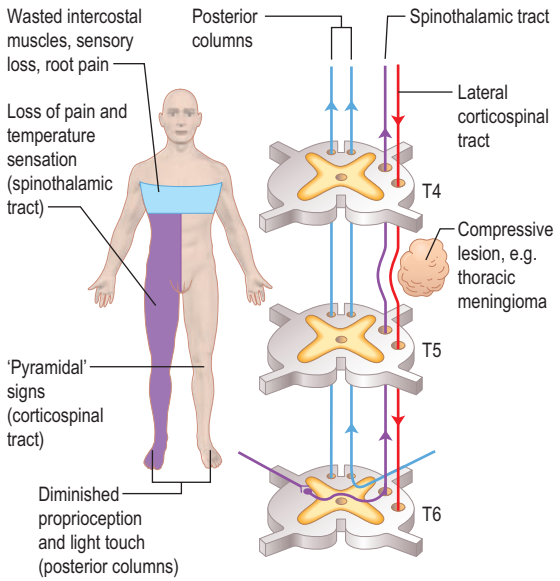


Fig. 21.14 **Spinal cord compression: features and anatomy.**

Root pain is frequent but not invariable, felt characteristically at the level of compression. With thoracic cord compression (e.g. an extradural meningioma), pain radiates around the chest, exacerbated by coughing and straining, as meningeal root sheaths are stretched.

Damage to one spinothalamic tract (contralateral loss of pain and temperature) with the ipsilateral corticospinal tract is known as the *Brown–Séquard syndrome* (originally, cord hemisection). The patient complains of numbness on one side and weakness on the other. Paraparesis/spinal cord lesions are discussed on page 1163.

Pontine lesions

Since lesions (e.g. an MS plaque) lie above the decussation of the posterior columns, and both medial lemniscus and spinothalamic tracts are close together, there is loss of all forms of sensation on the side opposite the lesion. Combinations of III, IV, V, VI and VII cranial nerve nuclei are seen, and may indicate a level (Fig. 21.12).

Thalamic lesions

Thalamic pain (also called *central post-stroke pain* or *thalamic syndrome*) follows a small thalamic infarct. The patient has a stroke (hemiparesis and sensory loss). Weakness improves, but deep-seated constant pain in the paretic limbs develops. Choreo-athetotic movements occur. Secondary depression may lead to self-harm. Thalamic lesions can also cause diminished sensation alone, on the opposite side; this is less usual.

Parietal cortex lesions

Sensory loss, neglect of one side, apraxia (p. 1097) and subtle disorders of sensation occur. Pain is not a feature of destructive cortical lesions. Irritative phenomena (e.g. partial sensory seizures from a parietal cortex glioma) cause tingling sensations in a limb, or elsewhere.

PAIN

Pain is an unpleasant, unique physical and psychological experience. Acute pain serves a biological purpose (e.g. withdrawal) and is typically self-limiting, ceasing as healing ensues. Some forms of chronic pain (e.g. causalgia, p. 1116) outlast the period required for healing, and may be permanent.

Essential physiology of pain

Pain perception is mediated by free nerve endings, terminations of finely myelinated A-delta and of non-myelinated C fibres. Chemicals released following injury produce pain either by direct stimulation or by sensitizing nerve endings. A-delta fibres give rise to perception of sharp, immediate pain, then slower-onset, more diffuse and prolonged pain is mediated by slower-conducting C fibres.

Sensory impulses enter the cord via dorsal spinal roots. Impulses ascend either in each dorsal (posterior) column or in each spinothalamic tract. Grey matter neurones in the cord are arranged in laminae labelled I to X (dorsal to ventral). A fibres terminate in laminae I and V and excite second-order neurones that project to the contralateral side via the anterior commissure and via the anterolateral column of the direct spinothalamic tract. C fibres mostly terminate in the substantia gelatinosa (laminae II and III); axons then pass through the anterior commissure to the contralateral side and rostrally, up the spino-reticulo-thalamic tract.

The spinothalamic tracts carry impulses that localize pain. Thalamic pathways to and from the cortex mediate emotional components. Sympathetic activity increases pain, e.g. hyperaemia in a painful limb.

Gate theory of pain

Gate theory, a useful way of thinking about pain, proposes that entry of afferent impulses is monitored first by the substantia gelatinosa. This gate determines whether or not sufficient activity penetrates to fire secondary neurones in the dorsal horn. This, and subsequent gates, are influenced by brain and cord regulators that alter how far a gate remains open.

Animal studies suggest that downstream regulators binding to calcium-regulated transcription factors *control* CNS endogenous opiate precursors (e.g. pro-dynorphin). These modulate pain.

Endogenous opiates

Endorphin peptides have opioid activity and probably account for placebo phenomena, effects of stress and acupuncture, and may explain in part why some patients develop chronic pain syndromes. Endorphins are CNS neurotransmitters acting at inhibitory synapses via δ, κ and μ receptors. Dynorphin (from the precursor, pro-dynorphin) is thought to modulate nociceptive (pain) impulses entering the cord.

Management of chronic pain

Chronic pain is gravely disabling, distressing, and taxing to treat (p. 494). Multidisciplinary pain-relief clinics provide specific and supportive therapy.

Management plans for intractable pain have seven components:

Diagnostic

Rigorous attention must be paid to the diagnosis, reviewing the entire history and investigations. A specific surgical approach may become apparent (e.g. spinal stenosis, trigeminal neuralgia, glossopharyngeal neuralgia, or syringomyelia).

Psychological

Chronic pain influences quality of life. Depression (p. 1198) is commonly associated with pain when the pathology is benign; antidepressants can help. Of patients suffering pain from secondary cancer about one-third are clinically depressed.

Analgesics

Perseverance and compliance with therapy is a common problem. The WHO analgesic ladder (p. 494) is useful.

Co-analgesics

Co-analgesics have a primary use other than for pain but help either alone or when added to analgesics. Examples are tricyclic antidepressants and anticonvulsants (p. 495). Calcium-channel blockers (nifedipine) improve sympathetically mediated pain in, for example, Raynaud's disease. Muscle relaxants (spasticity), antibiotics (infection) and steroids (inflammatory arthropathy) can each reduce pain.

Stimulation

Acupuncture, ice, heat, ultrasound, massage, transcutaneous electrical nerve stimulation (TENS) and spinal cord stimu-

FURTHER READING

McMahon SB, Koltzenburg M (eds). *Wall and Melzack's Textbook of Pain,* 5th edn. Edinburgh: Churchill Livingstone, 2005.

lation all achieve analgesia by gating effects on large myelinated nerve fibres.

Nerve blocks

Pain pathways can be blocked either temporarily by local anaesthetic or permanently with phenol, or radiofrequency lesions:

- somatic blocks:
  (a) peripheral nerve and plexus injections
  (b) epidural and spinal analgesia
- sympathetic blocks:
  (a) sympathetic ganglia injections
  (b) central epidural and spinal sympathetic blockade.

Neurosurgery

Highly specialized techniques have a place alongside drugs. Examples are dorsal rhizotomy, sympathectomy, cordotomy and neurostimulation.

BLADDER CONTROL AND SEXUAL DYSFUNCTION

Changes in micturition and failure of normal sexual activity due to neurological conditions are seen in sacral, spinal cord and cortical disease.

Essential functions and anatomy

The bladder has two functions – storage and voiding. Afferent pathways (T12–S4) respond to pressure within the bladder and sensation from the genitalia. As the bladder distends, continence is maintained by suppression of parasympathetic and reciprocal activation of sympathetic outflow. Both are under some voluntary control. Voiding takes place by parasympathetic activation of the detrusor, and relaxation of the internal sphincter (Table 21.18).

Cortical awareness of bladder fullness is located in the post-central gyrus, parasagittally, while initiation of micturition is in the pre-central gyrus. Voluntary control of micturition is located in the frontal cortex, parasagittally.

Neurological disorders of micturition

Urogenital tract disease is dealt with largely by urologists. Incontinence is common and easy to recognize; neurological causes are sometimes not obvious. These are:

Table 21.18 Efferents to bladder and genitalia

Nerve supply	Function
Parasympathetic	Bladder wall: contraction
S2–4	Internal sphincter: relaxation Penis/clitoris: engorgement
Sympathetic	Bladder wall: relaxation
T12–L2	Internal sphincter: contraction Orgasm, ejaculation
Pudendal nerves	External sphincter (skeletal muscle)

Cortical:

- Post-central lesions cause loss of sense of bladder fullness.
- Pre-central lesions cause difficulty initiating micturition.
- Frontal lesions cause socially inappropriate micturition.

Spinal cord. Bilateral UMN lesions (pyramidal tracts) cause urinary frequency and incontinence. The bladder is small and hypertonic, i.e. sensitive to small changes in intravesical pressure. Frontal lesions can also cause a hypertonic bladder.

LMN. Sacral lesions (conus medullaris, sacral root and pelvic nerve – *bilateral*) cause a flaccid, atonic bladder that overflows (cauda equina, p. 1177), often unexpectedly.

Management. Assessment of both urological causes (e.g. calculi, prostatism, gynaecological problems) and potential neurological causes of incontinence is necessary. Intermittent self-catheterization is used by many patients, with for example spinal cord lesions.

Male erectile dysfunction

Failure of penile erection often has mixed organic and psychological causes. Depression is common. Endocrine aspects of sexual dysfunction are described on page 1000. Erectile dysfunction is sometimes helped by phosphodiesterase type 5 inhibitors, e.g. sildenafil.

FURTHER READING

Rees PM, Fowler C, Maas C. Sexual function in men and women with neurological disorders. *Lancet* 2007; **369**: 512–525.

NEUROLOGICAL TESTS

NEURO-IMAGING

Skull and spinal X-rays

These show:

- fractures of the vault or base
- vault and skull base disease (e.g. metastases, osteomyelitis, Paget's disease, abnormal skull foramina, fibrous dysplasia)
- enlargement or destruction of the pituitary fossa – intrasellar tumour, raised intracranial pressure
- intracranial calcification – tuberculoma, oligodendroglioma, wall of an aneurysm, cysticercosis.

Spinal X-rays show fractures, congenital and destructive lesions (bone cysts, infection, metastases) and degenerative spondylosis.

Imaging brain and spinal cord

Brain CT and MRI are widely available world-wide. Myelography (contrast imaging of the cord and ventricles (ventriculography)) is obsolete.

Brain computed tomography (CT)

A collimated X-ray beam moves synchronously across a brain slice 2–13 mm thick. Transmitted X-irradiation from a pixel, an element <1 mm^2, is computer-processed to assign a Hounsfield number to its density (air = –1000 units; water = 0; bone = +1000 units). The digital data are converted to cross-sectional images to reconstruct brain anatomy. Helical

CT (spiral or volumetric CT) provides greater definition in a shorter time. It is used for two-dimensional reconstruction.

Differences in attenuation (density) between bone, brain and CSF enable recognition of normal and infarcted tissue, tumour, blood, intracerebral haemorrhage, free subarachnoid blood, subdural and extradural haematoma and oedema.

Enhancement with i.v. contrast delineates areas of altered blood supply (CT angiography).

Safety. The irradiation involved is relatively small. There are occasional reactions to contrast.

Limitations of brain CT

- Lesions under 1 cm in diameter may be missed.
- Lesions with attenuation close to bone may be missed, if near the skull.
- Lesions with attenuation similar to brain are poorly displayed (e.g. MS plaques, isodense subdural haematoma).
- CT is not good at detecting posterior fossa lesions because of surrounding bone.
- Patient cooperation: an anaesthetic is very occasionally needed.

Magnetic resonance imaging: MRI

A hydrogen nucleus is a proton whose electrical charge creates a local electrical field. Protons are aligned by sudden strong magnetic impulses and then imaged with radiofrequency waves at right angles to their alignment. The protons resonate and spin, then revert to their normal alignment. As they do so, images are made at different phases of relaxation, known as T1, T2, T2 STIR, FLAIR, diffusion-weighted imaging (DWI) and other sequences. From these sequences, often referred to as different weightings, recorded images are compared. Gadolinium is used as i.v. contrast to show areas of increased vascularity.

Advantages of MRI

- MR distinguishes between white and grey matter in the brain and cord.
- Cord and nerve roots are imaged directly.
- Pituitary imaging.
- MRI has resolution superior to CT (around 0.5 cm).
- No radiation is involved.
- MR angiography (MRA) images blood vessels without contrast.
- MR images soft tissues.
- Tumours, infarction, haemorrhage, MS plaques, posterior fossa, foramen magnum and cord are demonstrated well by MRI.

Limitations are principally time and cost. Patients need to keep still within a narrow tube: claustrophobia is an issue; open machines are available that are less claustrophobic. Patients with pacemakers or metallic fragments in the brain cannot be imaged. MR imaging for some days after lumbar puncture (p. 1121) frequently shows diffuse meningeal enhancement with gadolinium.

Doppler studies

B-mode and colour ultrasound are valuable in detection of carotid stenosis.

Digital cerebral and spinal angiography

Where advanced MRI and CT are available, these techniques are little used. Contrast is injected intra-arterially or intravenously. Angiography carries a mortality and stroke risk (<1%). Images of aorta, carotid, vertebral and brain arteries demonstrate occlusion, stenoses, atheroma, aneurysms and arteriovenous malformations (AVMs). Spinal angiography images cord AVMs.

Positron emission tomography (PET), single proton emission computed tomography (SPECT), dopamine transporter imaging (DAT) and functional MRI (fMRI)

These functional imaging techniques track uptake of radioisotopes and/or metabolites. PET is used principally in the detection of occult neoplasms, outside the CNS. SPECT is now little used in cerebrovascular disease and traumatic brain injury – there are issues of reliability. DAT is used in basal ganglia disease. fMRI is largely a research tool for mapping brain function, in health and disease.

Isotope bone scanning

The radioisotope [^{99m}Tc]-pertechnetate is given intravenously. The technique is used principally in detection of vertebral, skull and bone metastases.

Electroencephalography (EEG)

EEG (Fig. 21.15) recorded from scalp electrodes (16 channels simultaneously) is of value in epilepsy and diffuse brain diseases. Videotelemetry, combining EEG with video, is invaluable in attacks difficult to diagnose.

Epilepsy

Spikes, or spike-and-wave abnormalities, are hallmarks of epilepsy, but it should be emphasized that patients

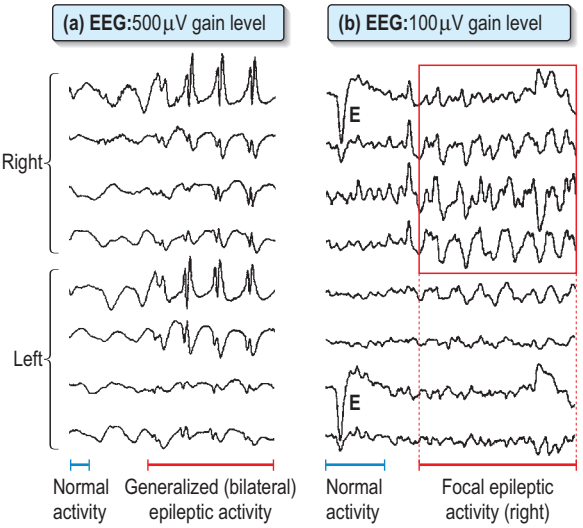


Fig. 21.15 **Examples of EEG traces. (a)** Normal activity followed by generalized epileptic spikes in all leads. **(b)** Normal activity followed by focal spikes in right leads. E, eye-movement artefact.

with epilepsy often have a normal EEG between seizures (p. 1141).

Diffuse brain disorders

Slow-wave EEG abnormalities appear in encephalitis, prion (Creutzfeldt–Jakob) diseases and metabolic states (e.g. hypoglycaemia, hepatic coma).

Brain death

The EEG is isoelectric (flat); EEG is no longer necessary to confirm brain death (p. 920).

Electromyography and conduction studies

Electromyography

A concentric needle electrode is inserted into voluntary muscle. Amplified recordings, on an oscilloscope, are also heard through a speaker. The main EMG features are:

- normal interference pattern
- denervation and reinnervation changes
- myopathic, myotonic and myasthenic changes (p. 1178).

Peripheral nerve conduction (Fig. 21.16)

Four measurements are of principal value in neuropathies and nerve entrapment:

- mean nerve (motor and sensory) conduction velocity
- distal motor latency
- sensory action potentials
- muscle action potentials.

Measurements differentiate between axonal and demyelinating damage and determine whether pathology is focal or diffuse.

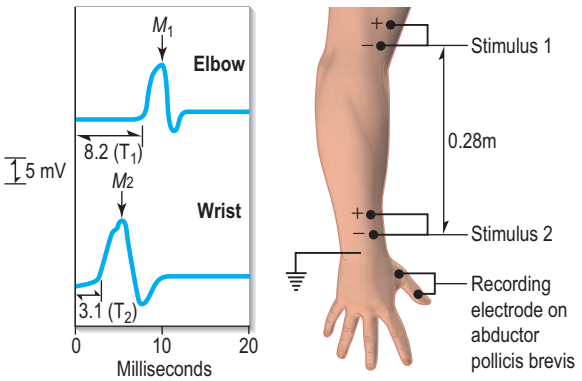


Fig. 21.16 **Measurement of motor conduction velocity (MCV) in ulnar nerve.** The recording electrode on abductor pollicis brevis measures muscle action potential (M) from ulnar nerve stimulation at elbow (Stimulus 1) and at wrist (Stimulus 2).
MCV calculation:

M_1 = muscle action potential (MAP) from Stimulus 1
M_2 = MAP from Stimulus 2
T_1 = time from elbow to recording electrode
T_2 = time from wrist to recording electrode

$$MCV = \frac{\text{distance (metres)}}{T_1 - T_2} \times 1000 = m/s$$

Cerebral-evoked potentials

Visual-evoked potentials (VEPs) record the interval visual stimuli take to reach the occipital cortex, and the amplitude of response. VEPs are used to confirm previous retrobulbar neuritis (p. 1102); this leaves a delayed latency despite clinical recovery.

Similar techniques for auditory and somatosensory potentials (from a limb) are also used.

Lumbar puncture and CSF examination (Table 21.19 and Practical box 21.3)

Indications for lumbar puncture (LP) are:

- diagnosis of meningitis and encephalitis
- diagnosis of subarachnoid haemorrhage (sometimes)
- measurement of CSF pressure, e.g. idiopathic intracranial hypertension (p. 1162)
- removal of CSF therapeutically, e.g. idiopathic intracranial hypertension
- diagnosis of various conditions, e.g. MS, neurosyphilis, sarcoidosis, Behçet's, neoplastic involvement, polyneuropathies
- intrathecal injection/drugs.

Meticulous attention should focus on microbiology in suspected CNS infection. Close liaison between clinician and microbiologist is essential. Specific techniques (e.g. polymerase chain reaction to identify bacteria) are invaluable. Repeated CSF examination is often necessary in chronic infection such as tuberculosis. Post-LP headaches, worse on standing, are a common complaint for several days (or more). Prolonged headaches can be treated by an 'autologous intrathecal blood patch' – injection of 20 mL of the patient's venous blood into the CSF.

FURTHER READING

Serpell MG, Rawal N. Headaches after diagnostic dural punctures. *British Medical Journal* 2000; **321**: 973–974.

Biopsy

Interpretation of brain, tonsillar, muscle and nerve histology requires specialist neuropathology services.

Brain and meninges

Brain biopsy (e.g. of a non-dominant frontal lobe) is sometimes used to diagnose inflammatory and degenerative brain diseases. CT and MR stereotactic biopsies of intracranial lesions are standard procedures.

Tonsillar biopsy

This is used in the diagnosis of variant Creutzfeldt–Jakob disease.

Table 21.19 The normal CSF

Appearance	Crystal clear, colourless
Pressure	60–150 mm of CSF, recumbent
Cell count	<5/mm^3 No polymorphs Mononuclear cells only
Protein	0.2–0.4 g/L
Glucose	⅔ to ½ of blood glucose
IgG	<15% of total CSF protein
Oligoclonal bands	Absent

Practical Box 21.3 Lumbar puncture

The procedure should be explained carefully to the patient, and consent obtained. LP should not be performed in the presence of raised intracranial pressure or when an intracranial mass lesion is a possibility.

Technique

- The patient is placed on the edge of the bed in the left lateral position with knees and chin as close together as possible.
- The third and fourth lumbar spines are marked. The fourth lumbar spine usually lies on a line joining the iliac crests.
- Using sterile precautions, 2% lidocaine is injected into the dermis by raising a bleb in either the third or fourth lumbar interspace.
- The LP needle is pushed through the skin in the midline, steadily forwards and slightly towards the head, with the head and spine bolstered horizontally with pillows.
- When the needle is felt to penetrate the dura, the stylet is withdrawn and a few drops of CSF allowed to escape.
- The CSF pressure can then be measured with a manometer connected to the needle. The patient's head must be on the same level as the sacrum. Normal CSF pressure is 60–150 mm of CSF. The level rises and falls with respiration and heart beat, and rises on coughing.
- CSF specimens are collected in three sterile bottles and a sample for CSF glucose, together with a simultaneous blood glucose sample.
- Record CSF naked-eye appearance: clear, cloudy, yellow (xanthochromic), red.
- The patient is asked to lie flat after the procedure to avoid subsequent headaches, but this manoeuvre is probably of little value.
- Analgesics may be required for post-LP headaches.

Contraindications

- Suspicion of a mass lesion in the brain or cord. Caudal herniation of the cerebellar tonsils (coning) may occur if an intracranial mass is present and the pressure below is reduced by removal of CSF.
- Any cause of raised intracranial pressure.
- Local infection near the LP site.
- Congenital lesions in the lumbosacral region (e.g. meningomyelocele).
- Platelet count $<40\times10^9$/L and other clotting abnormalities, including anticoagulant drugs.
- Unconscious patients and those with papilloedema must have a CT scan before lumbar puncture.

Notes

- Contraindications are relative; there are circumstances when LP is carried out despite them.
- Composition of normal CSF is shown in Table 21.19.

Muscle

Biopsy, with light and electron microscopy and biochemical analysis, elucidates diagnosis of inflammatory, metabolic and dystrophic disorders (p. 1178).

Peripheral nerve

Biopsy, usually of a sural nerve (ankle) or superficial branch of a radial nerve, aids diagnosis in polyneuropathies.

Psychometric assessment

Psychometric testing assesses cognitive function. Preservation of verbal IQ (a measure of past attainments) with deterioration of performance IQ (a measure of present abilities) indicates decline of cognitive function, seen for example following brain injury or in dementia. Low subtest scores (e.g. block design, various aspects of memory, visual, speech and constructional skills) indicate impaired function of specific brain regions.

Depression and lack of attention also reduce scores – a substantial problem. Opinions sometimes vary between psychologists about interpretation of tests, particularly after brain injury, limiting the value of the tests.

Table 21.20 Value of routine investigations in neurology

Test	Yield	Condition
Urinalysis	Glycosuria Ketones Bence Jones protein	Polyneuropathy Coma Cord compression
Blood picture	↑ MCV ↑ ↑ ESR	B_{12} deficiency Giant cell arteritis
Blood glucose	Hypoglycaemia Hyperglycaemia	Coma Coma
Serum electrolytes	Hyponatraemia Hypokalaemia	Coma Weakness
Serum calcium	Hypocalcaemia	Tetany, spasms
Serum CPK	Raised	Muscle disease
Chest X-ray	Lytic bone or mass lesion	Bronchial cancer, thymoma

Routine tests

See Table 21.20.

Specialized tests in specific diseases

Various tests are employed to diagnose individual (sometimes rare) diseases. Examples are:

- anti-cardiolipin and lupus anticoagulant antibody and detailed clotting studies in stroke (p. 544)
- antibody to acetylcholine receptor protein and anti-Mu SK antibodies in myasthenia gravis (p. 1180)
- serum copper and caeruloplasmin in Wilson's disease (p. 357)
- blood lactate studies (failure to rise on exercise) in McArdle's syndrome (p. 1182)
- anti-neuronal antibodies (paraneoplastic syndromes)
- serum phytanic acid (elevated) in Refsum's disease (p. 1174)
- serum long-chain fatty acid (present) in adrenoleucodystrophy (p. 1152)
- genetic studies – e.g. Huntington's disease, hereditary sensorimotor neuropathies and ataxias (p. 1149 and p. 1175).

UNCONSCIOUSNESS AND COMA

The ascending reticular formation, extending from lower brainstem to thalamus, influences the state of arousal. Our state of consciousness is the product of complex interactions between parts of the reticular formation, cortex and brainstem, and all sensory stimuli.

Disturbed consciousness: definitions

Each term simply describes a *recognizable* state:

- **Consciousness** means wakefulness with awareness of self and surroundings.

- **Clouding of consciousness** – used more in psychiatry than neurology – means reduced wakefulness and/or self-awareness, sometimes with confusion.
- **Confusion** means that the subject is bewildered and misinterprets his/her surroundings.
- **Delirium** is a state of confusion, sometimes with visual hallucination, and often high arousal (e.g. *delirium tremens*, p. 1213).
- **Sleep** is normal mental and physical inactivity: the subject can be roused.
- **Stupor** is abnormal; a sleepy state from which the subject can be aroused by vigorous or repeated stimuli. The term is also used for psychiatric states, e.g. catatonic and depressive stupor (p. 1216).
- **Coma** means unrousable unresponsiveness.

The Glasgow Coma Scale, designed for recording coma following head injuries, is shown in Table 21.21.

Mechanisms of coma

Altered consciousness is produced by three mechanisms affecting brainstem, reticular formation and cortex.

- *Diffuse brain dysfunction*. Generalized severe metabolic or toxic disorders (e.g. alcohol, sedatives, uraemia, septicaemia) depress overall brain function.
- *Direct effect within the brainstem*. A brainstem lesion inhibits the reticular formation.
- *Pressure effect on the brainstem*. A mass lesion within the brain compresses the brainstem, inhibiting the reticular formation.

A single focal hemisphere (or cerebellar) lesion does not produce coma unless it compresses or damages the brainstem. Cerebral oedema frequently surrounds masses, increasing their effects.

Other states of unresponsiveness must be distinguished from coma; this sometimes causes difficulty. Features of these are listed in Table 21.24 and illustrated in Figure 21.17.

Table 21.21 Glasgow Coma Scale

	Score
Eye opening (*E*)	
Spontaneous	4
To speech	3
To pain	2
No response	1
Motor response (*M*)	
Obeys	6
Localizes	5
Withdraws	4
Flexion	3
Extension	2
No response	1
Verbal response (*V*)	
Orientated	5
Confused conversation	4
Inappropriate words	3
Incomprehensible sounds	2
No response	1

Glasgow Coma Scale = $E + M + V$
(GCS minimum = 3: maximum = 15)

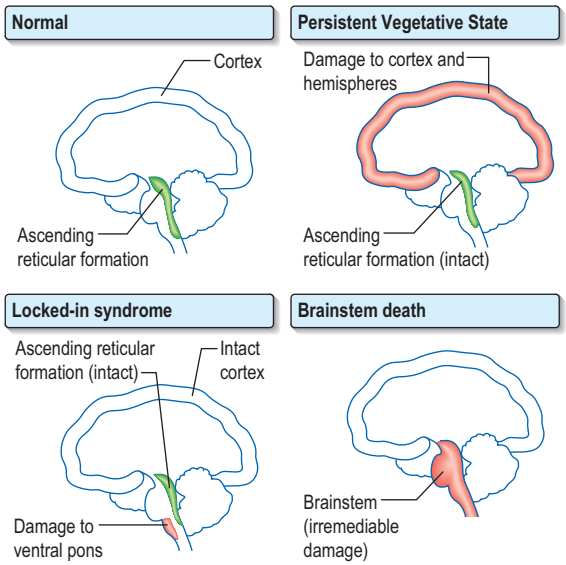


Fig. 21.17 **Anatomy of vegetative state, locked-in syndrome and brainstem death.** Adapted from Bates D. Coma and brainstem death. *Medicine* 2004; **32**(10), with permission of Elsevier.

Table 21.22 Principal causes of coma – examples of mechanisms

Diffuse brain dysfunction
Drug overdose, alcohol abuse
CO poisoning, anaesthetic gases
Hypoglycaemia, hyperglycaemia
Hypoxic/ischaemic brain injury
Hypertensive encephalopathy (p. 800)
Severe uraemia (p. 626)
Hepatocellular failure (p. 342)
Respiratory failure with CO_2 retention (p. 836)
Hypercalcaemia, hypocalcaemia
Hypoadrenalism, hypopituitarism and hypothyroidism
Hyponatraemia, hypernatraemia
Metabolic acidosis
Hypothermia, hyperpyrexia
Trauma to brain
Epilepsy
Encephalitis, cerebral malaria, septicaemia
Subarachnoid haemorrhage
Metabolic rarities, e.g. porphyria
Cerebral oedema from chronic hypoxia (p. 910)
Direct effect within brainstem
Brainstem haemorrhage or infarction
Brainstem neoplasm, e.g. glioma
Brainstem demyelination
Wernicke–Korsakoff syndrome
Trauma
Pressure effect on brainstem
Hemisphere tumour, infarction, haematoma, abscess, encephalitis or trauma
Cerebellar mass

Causes of coma and stupor

Principal causes of coma and stupor are shown in Table 21.22. A common cause of coma in the UK is self-poisoning; world-wide, cerebral malaria is a frequent cause.

THE UNCONSCIOUS PATIENT

Immediate assessment

Actions that take seconds save lives, see Table 21.23.

- Assess airway; resuscitate.
- Obtain a history – circumstances, witnesses, paramedics, police. Many people carry drug details and identification data.

General and neurological examination

- Record depth of coma (Table 21.21).
- Examine the patient as fully as possible.

Many clues can point to the cause.

Temperature

Body temperature is raised in infection and hyperpyrexia, and is subnormal in hypothermia. *Feel* the abdomen and record the temperature.

Skin, skull, spine and injuries

Look for skull trauma/burr holes, neck stiffness, other injuries, cyanosis, jaundice, purpura, rashes, pigmentation and injection marks. In hypothyroidism, the skin is coarse/dry.

Breath

Sniff – for ketones, alcohol, hepatic and uraemic fetor.

Respiration

Depressed but regular respiration occurs in many states of stupor and coma. As coma deepens, respiration becomes irregular. Patterns of abnormal respiration are of significance:

- *Cheyne–Stokes (periodic) respiration* is alternating hyperpnoea and periods of apnoea. This may point to bilateral cerebral or upper brainstem dysfunction and may be a sign of incipient coning. A similar pattern can occur in normal people during deep sleep, in metabolic comas, with CO_2 retention, and chronic hypoxia at altitude.
- *Acidotic (Kussmaul) respiration* is sustained, deep, sighing hyperventilation seen in diabetic ketoacidosis and uraemia.
- *Central neurogenic hyperventilation* is sustained, rapid, respiration seen with pontine lesions. It may switch off and on abruptly.
- *Ataxic respiration* is shallow, halting, irregular respiration that frequently precedes death. The medullary respiratory centre is failing.
- *Vomiting, hiccup and excessive yawning* in stupor point to a lower brainstem lesion.

Table 21.23 Unconsciousness: actions

Examine	Issue	Potential action
Airway	Clear?	Maintain; intubate?
Pulse	Absent?	Cardiopulmonary resuscitation
Pupils	Fixed, dilated?	Record, resuscitate
Head	Trauma?	Observe and investigate
Spine	Trauma?	Immobilize

Pupils

Record size and reaction to light.

- *Dilatation of one pupil* that then becomes fixed to light indicates compression of the IIIrd nerve (coning of the temporal lobe uncus). This is a potential neurosurgical emergency (e.g. extradural haematoma).
- *Horner's syndrome* (ipsilateral pupillary constriction and ptosis, p. 1103) occurs with hypothalamic damage and also, rarely, in coning.
- *Bilateral mid-point reactive pupils* (i.e. normal pupils) are characteristic of metabolic comas and follow coma due to sedative drugs except opiates.
- *Bilateral light-fixed, dilated pupils* are one cardinal sign of brain death. They can occur in deep coma of any cause, but particularly in barbiturate intoxication and hypothermia.
- *Bilateral pinpoint, light-fixed pupils* occur with pontine lesions (e.g. haemorrhage) and with opiates.
- *Bilateral mid-position light-fixed or slightly dilated light-fixed pupils* (4–6 mm), sometimes irregular, are seen in brainstem lesions.
- Mydriatic drugs and previous pupillary surgery can cause diagnostic difficulty.

Fundi

Look for papilloedema and retinal haemorrhage.

Ocular movements

In light coma, slow, roving, side-to-side eye movements are seen; *ocular axes* are often slightly divergent.

Vestibulo-ocular reflexes. Passive head turning produces conjugate ocular deviation away from the direction of rotation (doll's head reflex). This reflex disappears in deep coma, in brainstem lesions and in brain death.

Calorics. Slow tonic ocular deviation towards the irrigated ear is seen when ice-cold water is run into the external auditory meatus – caloric/vestibulo-ocular reflexes, indicating an intact brainstem. Calorics are used in the confirmation of brain death (p. 920).

Abnormalities of conjugate gaze (p. 1105):

- *Sustained conjugate lateral deviation* towards a destructive frontal lesion – eyes look towards the normal limbs. An irritative lesion in one frontal lobe (e.g. an epileptic focus) drives the eyes away – deviation is away from the lesion. In the pons lesion, when one PPRF is damaged (p. 1104), sustained conjugate lateral deviation occurs away from the lesion, towards the weak limbs, because the opposite PPRF is active.
- *Skew deviation* (one eye up, the other down) indicates a brainstem or cerebellar lesion.
- *Ocular bobbing*, i.e. sudden, brisk, downward-diving eye movements, occur in pontine (or cerebellar) haemorrhage.

Other spontaneous eye movements (other than roving movements) are distinctly unusual in true coma.

Lateralizing signs

Coma makes it difficult to recognize lateralizing signs. These are helpful:

- *Asymmetry of response to visual threat* in a stuporose patient: suggests hemianopia.
- *Asymmetry of face*: drooping or dribbling on one side, blowing in and out of mouth when the paralysed cheek does not move.
- *Asymmetry of tone*. Unilateral flaccidity or spasticity may be the only sign of hemiparesis.
- *Asymmetry of decerebrate and decorticate posturing.*
- *Asymmetrical response to painful stimuli.*
- *Asymmetry of tendon reflexes and plantar responses*. Both plantars are often extensor in deep coma.

Unresponsiveness of psychological origin is also a prominent cause of apparent coma.

Investigations in coma

Often, the cause is evident (e.g. head injury, cerebral haemorrhage, self-poisoning); if no cause is evident, further investigations are essential.

Blood and urine

- *Drugs screen* (e.g. salicylates, diazepam, narcotics, amfetamines).
- *Routine biochemistry* (urea, electrolytes, glucose, calcium, liver biochemistry).
- *Metabolic and endocrine studies* (TSH, cortisol).
- *Blood cultures.*
- *Other*, e.g. cerebral malaria (request *thick* blood film) and porphyria (p. 1073).

Imaging

CT or MR brain imaging may indicate an unsuspected mass lesion or intracranial haemorrhage.

CSF examination

Lumbar puncture should be performed in coma only after careful risk assessment. It is contraindicated when an intracranial mass lesion is a possibility: CT is essential to exclude this. CSF examination is likely to alter therapy only if undiagnosed meningoencephalitis or other infection is present.

Electroencephalography

EEG is of some value in the diagnosis of metabolic coma and encephalitis.

Management

Comatose and stuporose patients – at home or outside, on a trolley, in a ward or ITU – need immediate careful nursing, meticulous attention to the airway, and frequent monitoring of vital functions.

Longer-term essentials are:

- skin care – turning (to avoid pressure sores and pressure palsies), removal of jewellery
- oral hygiene – mouthwashes, suction
- eye care – prevention of corneal damage (lid taping, irrigation)
- fluids – intragastric or i.v.
- calories – liquid diet through a fine intragastric tube, 1255 kJ (3000 kcal) daily
- sphincters – catheterization when essential (use Paul's tubing if possible); rectal evacuation.

Differential diagnosis of the vegetative state

Prolonged coma and similar states require accurate diagnosis to assess the nature of the problem, prognosis and management (Table 21.24).

Table 21.24 Differential diagnosis of the vegetative state

Feature	Vegetative	Minimally conscious	Locked-in	Coma	Brain death
Awareness	Absent	Present	Present	Absent	Absent
Sleep–wake cycles	Present	Present	Present	Absent	Absent
Noxious stimuli	Response	Response	Response (eyes only)	Response ±	Response: nil
Glasgow Coma Scale	E4, M1–4, V1–2	E4, M1–5, V1–4	E4, M1, V1	E1–2, M1–4, V1–2	E1, M1–3, V1
Motor function	No purposeful movement	Consistent or inconsistent sounds/ movements	Vertical eye movements/blinking	No purposeful movement	None/only reflex spinal movement
Respiration	Preserved	Preserved	Preserved	Variable	Absent
EEG	Slow waves, usually	Data insufficient	Normal, usually	Slow waves, usually	Typically absent
Cerebral metabolism (PET)	Severely reduced	Data insufficient	Mildly reduced	Reduced	Severely reduced/ absent
Prognosis	Varies: usually continued VS or death	Varies	Varies: full recovery unlikely	Recovery, VS, or death	Already dead

EEG, electroencephalography; PET, positron emission tomography; VS, vegetative state
EEG and PET are not required to confirm brain death
Reproduced from Royal College of Physicians. *The vegetative state: guidance on diagnosis and management.* Report of a Working Party. London: RCP, 2003. Copyright © 2003 Royal College of Physicians. Reproduced by permission.

- *Vegetative state* (VS) is a sequel of, for example, widespread cortical damage after brain injury. Brainstem function is normal. VS implies loss of sentient behaviour. The patient perceives little or nothing but lies apparently awake, breathing spontaneously.
- *Minimally conscious state* (MCS) describes patients with some sentient behaviour, e.g. apparent, vague pain perception. A patient may emerge from VS into MCS.
- *Locked-in syndrome* is a state of unresponsiveness due to massive brainstem damage below the level of the IIIrd nerve nuclei. The patient has a functioning cerebral cortex and is fully aware, unlike the VS/MCS patient. They cannot move or communicate except by vertical eye movement.
- *Brainstem death* is discussed on page 920.

Distinction between these states is essential before addressing issues of prognosis and cessation of supportive care.

FURTHER READING

Royal College of Physicians Working Party Report. The vegetative state: guidance on diagnosis and management. *Clinical Medicine* 2003; **3**: 249–254.

STROKE AND CEREBROVASCULAR DISEASE

Stroke is the second commonest cause of death (9%) and a major cause of disability world-wide. Although data are difficult to obtain, approximately two-thirds of the global burden of strokes is in middle- and low-income countries. WHO has introduced a strokes surveillance study in developing countries (STEPS surveillance). The age-adjusted annual death rate from strokes is about 200 per 100 000 in the UK (12% of all deaths). Rates are higher in Asian and black African populations than in Caucasians. Stroke is uncommon below the age of 40 and commoner in males. The death rate following stroke is 20–25%. Hypertension is the most treatable stroke risk factor; stroke is decreasing in the 40–60 age group because hypertension is treated. In the elderly, it remains a major cause of morbidity and mortality. Thromboembolic infarction (80%), cerebral and cerebellar haemorrhage (10%) and subarachnoid haemorrhage (SAH) (about 5%) are the main causes; arterial dissection and arteriovenous malformations also contribute.

Definitions

- **Stroke**. To the public, stroke means weakness, usually permanent on one side, often with loss of speech. Stroke is *defined* as a syndrome of rapid onset of cerebral deficit (usually focal) lasting >24 hours or leading to death, with no cause apparent other than a vascular one. Hemiplegia following middle cerebral arterial thromboembolism is the typical example.
- **Completed stroke** means the deficit has become maximal, usually within 6 hours.
- **Stroke-in-evolution** describes progression during the first 24 hours.
- **Minor stroke**. Patients recover without significant deficit, usually within a week.
- **Transient ischaemic attack (TIA)** means a sudden focal deficit, e.g. a weak limb, aphasia or loss of vision lasting from seconds to 24 hours with complete recovery. This definition is unsatisfactory as after 1 hour ischaemic damage has already occurred (see below). TIAs have a tendency to recur, and may herald thromboembolic stroke.

Avoid the vague term 'cerebrovascular accident'.

Pathophysiology

Different pathological processes cause similar clinical events in cerebrovascular disease.

Completed stroke

This is caused by:

- arterial embolism from a distant site (usually heart, carotid, vertebral or basilar arteries) and subsequent brain infarction
- arterial thrombosis in atheromatous carotid, vertebral or cerebral arteries with subsequent infarction
- haemorrhage (intracranial or SAH).

Less commonly, other processes cause stroke:

- venous infarction
- carotid or vertebral artery dissection
- polycythaemia and hyperviscosity syndromes
- fat and air embolism (see diving, p. 956)
- multiple sclerosis – demyelinating plaque(s)
- mass lesions (e.g. brain tumour, abscess, subdural haemorrhage).

Transient ischaemic attacks

TIAs are usually the result of microemboli, but different mechanisms produce similar clinical events. For example, TIAs may be caused by a fall in cerebral perfusion (e.g. a cardiac dysrhythmia, postural hypotension or decreased flow through atheromatous arteries). Infarction is usually averted by autoregulation (p. 1127). Small areas of infarction following thrombosis or haemorrhage may occasionally cause a clinical TIA, and 25% of TIAs have MRI changes showing small infarcts. Rarely, tumours and subdural haematomas cause episodes indistinguishable from thromboembolic TIAs: a clinical TIA is thus not a reliable indicator of thromboembolism. Principal sources of emboli to the brain are cardiac thrombi and atheromatous plaques/thrombi within the great vessels, carotid and vertebral systems. Cardiac thrombi (mural and valvular) follow atrial fibrillation, often secondary to valvular disease, or myocardial infarction. Polycythaemia is also a cause. Thromboembolism from sources outside the brain generates 80% of TIAs and also 70% of strokes.

Carotid and vertebral artery dissection

Dissection accounts for around 1 in 5 strokes below age 40 and is sometimes a sequel of head or neck trauma. Stroke, TIA or sudden headache/migraine-like symptoms occur, sometimes with neck pain at the site of dissection.

Risk factors and prevention

The principal risk factors and effects of altering these are shown in Table 21.25.

Low-dose aspirin is used in primary prevention in men and women with a 10-year risk of coronary heart disease greater than 10%; its effect on stroke prevention is unclear.

For brain haemorrhage, hypertension, bleeding disorders, anticoagulants and antiplatelet drugs, pre-existing cerebral aneurysms and AVMs are risk factors.

www.studentconsult.com
Ask the authors

Table 21.25 Factors reducing stroke risk

Risk factor	Action	Reduction in stroke risk		
		Infarction	Haemorrhage	SAH
Hypertension	Treat and monitor	++	++	Probable
Smoking	Stop	++	+	Probable
Lifestyle	More active	+	0	0
Alcohol	Moderate intake	+	0	0
High cholesterol	Statins, diet	+	0	0
Raised haematocrit	Reduce	+	0	0
Atrial fibrillation	Anticoagulate	+	*Increases* risk slightly	0
Obesity	Weight reduction	Probable	Probable	0
Diabetes	Good control	Probable	0	0
Severe carotid stenosis	Surgery	++	0	0
Sleep apnoea	Treat	+	0	0

++: Major correlation with reduced stroke risk, +: moderate correlation, SAH: subarachnoid haemorrhage.

Rarer risk factors and other causes of stroke

- Thrombocythaemia and thrombophilia (protein C deficiency, factor V Leiden) are weakly associated with arterial stroke but predispose to cerebral venous thrombosis.
- Anti-cardiolipin and lupus anticoagulant antibodies (antiphospholipid syndrome, p. 544) predispose to arterial thrombotic strokes in young patients.
- Endocarditis (p. 769) – thromboembolic stroke may be the presenting feature.
- Low-dose oestrogen-containing oral contraceptives do not increase stroke risk significantly in healthy women but probably do so with other risk factors, e.g. uncontrolled hypertension or smoking.
- Migraine is a rare cause of cerebral infarction (p. 1136).
- Vasculitis (systemic lupus erythematosus (SLE), polyarteritis, giant cell arteritis, granulomatous CNS angiitis) is a rare cause of stroke.
- Amyloidosis can present as recurrent cerebral haemorrhage (p. 1072).
- Hyperhomocysteinaemia predisposes to thrombotic strokes. Folic acid therapy does not reduce the incidence.
- Neurosyphilis, SLE, mitochondrial disease.
- Drugs of abuse e.g. cocaine, and possibly over-the-counter cold remedies containing vasoconstrictors.
- Cox-2 inhibitors are associated with a slightly increased incidence of stroke (p. 517).

CADASIL (cerebral dominant arteriopathy with subcortical infarcts and leucoencephalopathy) is a rare inherited cause of stroke/vascular dementia. There is a mutation in the *NOTCH3* gene on chromosome 19. Characteristic damage to small brain arterioles follows with multiple infarcts, which can be seen on MRI in the subcortical white matter. A skin biopsy shows granular osmophilic material (GOM). Familial migraine and depression occur in youth, progressing to TIA and stroke in the third and fourth decades, and dementia in the sixth.

Vascular anatomy

Knowledge of normal arterial anatomy and likely sites of atheromatous plaques and stenoses helps understanding of the main stroke syndromes.

The circle of Willis is supplied by the two internal carotid arteries and by the basilar. The distribution of the anterior, middle and posterior cerebral arteries that supply the cerebrum is shown in Figures 21.18–21.20.

Stenoses and plaques proximal to the circle of Willis are seen typically at four extracranial sites (1–4: see Fig. 21.20):

- origin of common carotid artery (1)
- origin of internal carotid artery (2)
- origin of vertebral artery (3)
- subclavian artery (4) and carotid artery syphon – within the cavernous sinus.

Autoregulation

Usually, constant cerebral blood flow (CBF) is maintained at mean arterial blood pressures between 60 and 120 mmHg, smooth muscle in small arteries responding directly to changes in pressure. CBF is normally independent of perfusion pressure, i.e. there is autoregulation.

In disease, CBF autoregulation can fail. Contributory causes are:

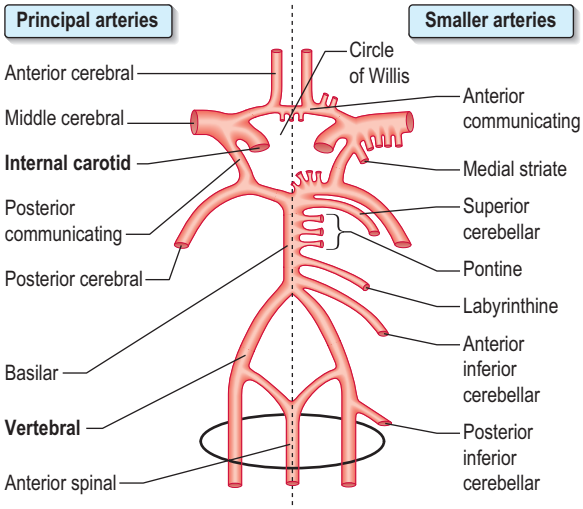


Fig. 21.18 **Arteries supplying brain and cord.**

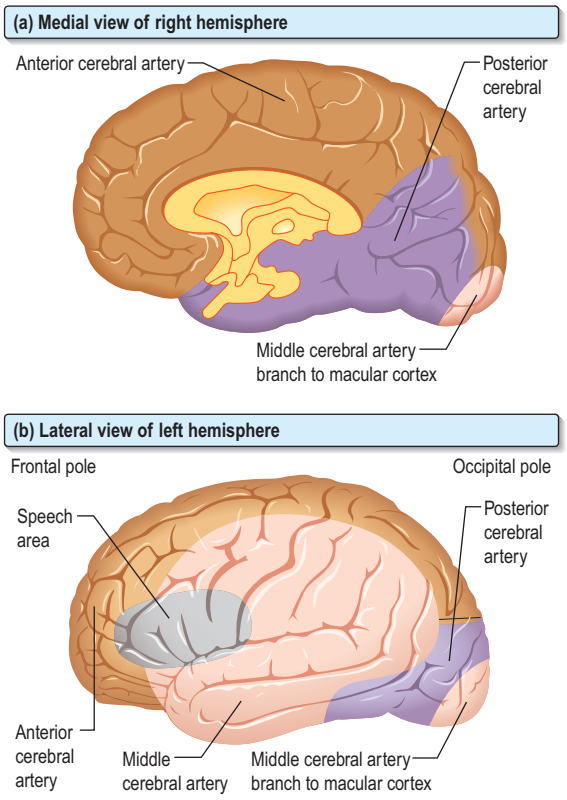


Fig. 21.19 **The three major cerebral arteries.**

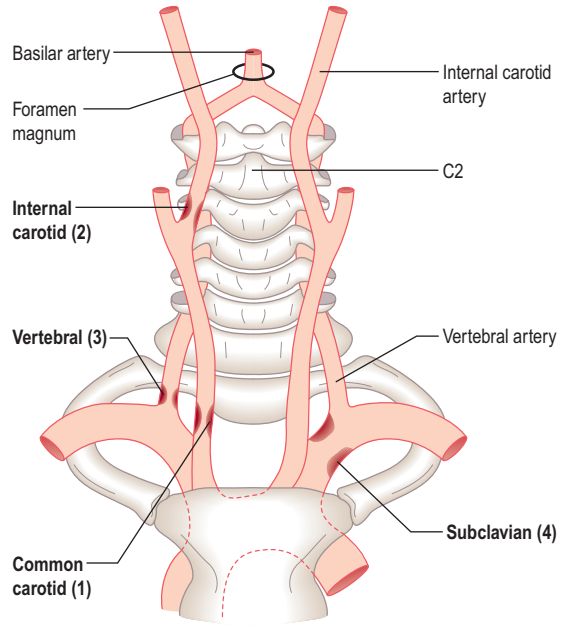


Fig. 21.20 **Principal sites of stenoses in extracerebral arteries.**

- severe hypotension with systolic BP <75 mmHg
- severe hypertension with systolic BP >180 mmHg
- increase in blood viscosity, e.g. polycythaemia
- raised intracranial pressure
- increase in arterial PCO_2 and/or fall in arterial PO_2.

CLINICAL SYNDROMES

Transient ischaemic attacks (TIAs)

Features

TIAs cause sudden loss of function, usually within seconds, and last for minutes or hours (but by classical definition <24 hours). The site is often suggested by the type of attack. Clinical features of the principal forms of TIA are given in Table 21.26. Hemiparesis and aphasia are the commonest; two other events are mentioned briefly here.

Amaurosis fugax

This is a sudden transient loss of vision in one eye. When due to the passage of emboli through the retinal arteries, arterial obstruction is sometimes visible through an ophthalmoscope during an attack. A TIA causing an episode of amaurosis fugax is often the first clinical evidence of internal carotid artery stenosis – and forerunner of a hemiparesis. Amaurosis fugax also occurs as a benign event in migraine.

Transient global amnesia

Episodes of amnesia/confusion lasting several hours, occurring principally in people over 65 and followed by complete recovery, are presumed to be due to posterior circulation ischaemia. The exact cause of this striking but unusual event is unknown. Episodes rarely recur and there is little evidence that they are actually thromboembolic.

Clinical findings in TIA

Diagnosis of TIA is often based solely upon its description. It is unusual to witness an attack as they are so brief. Consciousness is usually preserved in TIA. There may be clinical evidence of a source of embolus, e.g.:

- carotid arterial bruit (stenosis)
- atrial fibrillation or other dysrhythmia
- valvular heart disease/endocarditis
- recent myocardial infarction
- difference between right and left brachial BP.

An underlying condition may be evident:

- atheroma
- hypertension
- postural hypotension
- bradycardia or low cardiac output
- diabetes mellitus

Table 21.26 **Features of transient ischaemic attacks**

Anterior circulation	Posterior circulation
Carotid system	***Vertebrobasilar system***
Amaurosis fugax	Diplopia, vertigo, vomiting
Aphasia	Choking and dysarthria
Hemiparesis	Ataxia
Hemisensory loss	Hemisensory loss
Hemianopic visual loss	Hemianopic visual loss
	Bilateral visual loss
	Tetraparesis
	Loss of consciousness (rare)
	Transient global amnesia (possibly)

- rarely, arteritis, polycythaemia, neurosyphilis, HIV
- antiphospholipid syndrome (p. 544).

Differential diagnosis

TIAs can be distinguished, usually on clinical grounds, from other transient episodes (p. 1144). Occasionally, events identical to TIAs are produced by mass lesions. Focal epilepsy is usually recognized by its positive features (e.g. limb jerking and loss of consciousness) and progression over minutes. In a TIA, involuntary limb movements do occur occasionally; deficit is usually instantaneous. A focal prodrome in migraine sometimes causes diagnostic difficulty. Headache, common but not invariable in migraine, is rare in TIA. Typical migrainous visual disturbances are not seen in TIA. Investigations and management are discussed with stroke (p. 1131).

Prognosis

Prospective studies show that 5 years after a single thromboembolic TIA:

- 30% have had a stroke, a third of these in the first year
- 15% have suffered a myocardial infarct.

TIAs in the anterior cerebral circulation carry a more serious prognosis than those in the posterior circulation (see Table 21.26).

Cerebral infarction

Major thromboembolic cerebral infarction usually causes an obvious stroke. Some small infarcts cause TIAs; others are silent. The clinical picture is thus very variable, depending on the infarct site and extent. Whilst the general site can be deduced from physical patterns (e.g. cortex, internal capsule, brainstem), clinical estimations of precise vascular territories are often inaccurate, when compared to imaging.

Following vessel occlusion brain ischaemia occurs, followed by infarction. The infarcted region is surrounded by a swollen area which does not function but is structurally intact. This is the *ischaemic penumbra*, which is detected on MRI and can regain function with neurological recovery.

Within the ischaemic area, hypoxia leads to neuronal damage. There is a fall in ATP with release of glutamate, which opens calcium channels with release of free radicals. These alterations lead to inflammatory damage, necrosis and apoptotic cell death.

Clinical features

The stroke most typically seen is caused by infarction in the internal capsule following thromboembolism in a middle cerebral artery branch (Fig. 21.21). A similar picture is caused by internal carotid occlusion (see Fig. 21.20). Limb weakness on the opposite side to the infarct develops over seconds, minutes or hours (occasionally longer). There is a contralateral hemiplegia or hemiparesis with facial weakness. Aphasia is usual when the dominant hemisphere is affected. Weak limbs are at first flaccid and areflexic. Headache is unusual. Consciousness is usually preserved. Exceptionally, an epileptic seizure occurs at the onset. After a variable interval, usually several days, reflexes return, becoming exaggerated. An extensor plantar response appears. Weakness is maximal at first; recovery occurs gradually over days, weeks or many months.

Fig. 21.21 **CT: Massive middle cerebral artery infarct.**

Brainstem infarction

This causes complex signs depending on the relationship of the infarct to cranial nerve nuclei, long tracts and brainstem connections (Table 21.27).

- *The lateral medullary syndrome* (posterior inferior cerebellar artery (PICA) thrombosis and Wallenberg's syndrome) is a common example of brainstem infarction presenting as acute vertigo with cerebellar and other signs (Table 21.28 and Fig. 21.22). It follows thromboembolism in the PICA or its branches, vertebral artery thromboembolism or dissection. Features depend on the precise structures damaged.
- *Coma* follows damage to the brainstem reticular activating system.
- *The locked-in syndrome* is caused by upper brainstem infarction (p. 1126).
- *Pseudobulbar palsy* (p. 1110) can follow lower brainstem infarction.

Table 21.27 Features of brainstem infarction

Clinical feature	Structure involved
Hemiparesis or tetraparesis	Corticospinal tracts
Sensory loss	Medial lemniscus and spinothalamic tracts
Diplopia	Oculomotor system
Facial numbness	Vth nerve nuclei
Facial weakness	VIIth nerve nucleus
Nystagmus, vertigo	Vestibular connections
Dysphagia, dysarthria	IXth and Xth nerve nuclei
Dysarthria, ataxia, hiccups, vomiting	Brainstem and cerebellar connections
Horner's syndrome	Sympathetic fibres
Coma, altered consciousness	Reticular formation

Table 21.28 Clinical signs in the lateral medullary syndrome (PICA thrombosis)

Ipsilateral	Contralateral
Facial numbness (Vth)	Spinothalamic sensory loss
Diplopia (VIth)	Hemiparesis (mild, unusual)
Nystagmus	
Ataxia (cerebellar)	
Horner's syndrome	
IXth and Xth nerve lesions	

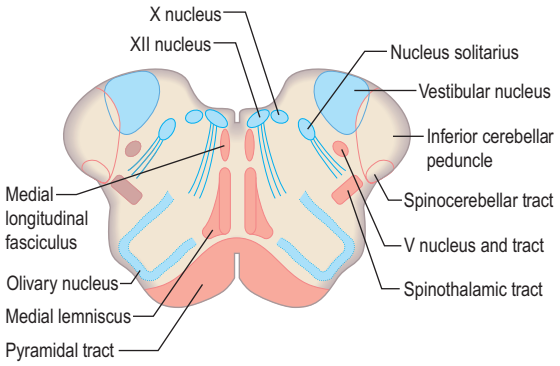


Fig. 21.22 **Medulla (cross-section): structures at risk after brainstem infarction.**

Other patterns of infarction

Lacunar infarction

Lacunes are small (<1.5 cm^3) infarcts seen on MRI or at autopsy. Hypertension is commonly present. Minor strokes (e.g. pure motor stroke, pure sensory stroke, sudden unilateral ataxia and sudden dysarthria with a clumsy hand) are syndromes caused typically by single lacunar infarcts. Lacunar infarction is often symptomless.

Hypertensive encephalopathy (p. 800)

This is due to cerebral oedema, causing severe headaches, nausea and vomiting. Agitation, confusion, fits and coma occur if the hypertension is not treated. Papilloedema develops, either due to ischaemic optic neuropathy or following the brain swelling due to multiple acute infarcts. MRI shows oedematous white matter in the parieto-occipital regions.

Multi-infarct dementia (vascular dementia)

Multiple lacunes or larger infarcts cause generalized intellectual loss seen with advanced cerebrovascular disease. In the late stages, there is dementia, pseudobulbar palsy and a shuffling gait – the *marche à petits pas* (small steps), sometimes called atherosclerotic parkinsonism. *Binswanger's disease* is a term for widespread low attenuation in cerebral white matter, usually with dementia, TIAs and stroke episodes in hypertensive patients (the changes being seen on imaging/autopsy).

Visual cortex infarction

Posterior cerebral artery infarction or infarction of the middle cerebral artery macular branch causes combinations of hemianopic visual loss and cortical blindness (Anton's syndrome, Fig. 21.19 and p. 1128).

Weber's syndrome

Ipsilateral IIIrd nerve palsy with contralateral hemiplegia is due to a unilateral infarct in the midbrain. Paralysis of upward gaze is usually present.

Watershed (borderzone) infarction

Cortical infarcts, often multiple, follow prolonged periods of low perfusion (e.g. hypotension after massive myocardial infarction or cardiac bypass surgery). Infarcts occur in the borderzones, between areas supplied by the anterior, middle and posterior cerebral arteries. Cortical visual loss, memory loss and intellectual impairment are typical. In some cases a vegetative state or minimal conscious state follows (p. 1125).

Acute stroke: immediate care, and thrombolysis

Paramedics and members of the public are encouraged to make the diagnosis of stroke on a simple history and examination – FAST:

- **F**ace – sudden weakness of the face
- **A**rm – sudden weakness of one or both arms
- **S**peech – difficulty speaking, slurred speech
- **T**ime – the sooner treatment can be started, the better.

Dedicated units with multidisciplinary, organized teams deliver higher standards of care than a general hospital ward, reducing stroke mortality and long-term disability. Evidence-based guidelines have contributed to clear protocols. Admission to hospital should proceed without delay, for imaging, care and investigation.

Following a stroke, immediate, continued and meticulous attention to the airway and to swallowing is essential. Management of unconscious or stuporose patients is outlined on page 1125.

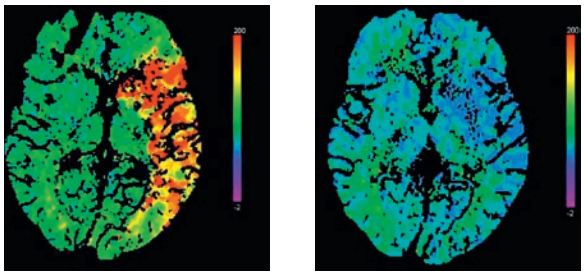
In cerebral infarction, the issue is thrombolysis. Table 21.29 outlines the current proposals for stroke thrombolysis. The benefit of thrombolysis is shown on CT perfusion scans (Fig. 21.23 a and b) and decreases with time, even within the time window of 4.5 hours. *Every minute counts*. If thrombolysis is not given, aspirin 300 mg daily should be given as soon as a diagnosis of ischaemic stroke or thromboembolic TIA is confirmed, reducing to 75 mg after several days. Following thrombolysis aspirin should not be started until 24–48 hours later.

Investigations

The purpose of investigations in stroke is:

- to confirm the clinical diagnosis and distinguish between haemorrhage and thromboembolic infarction
- to look for underlying causes and to direct therapy
- to exclude other causes, e.g. tumour.

Sources of embolus should be sought (e.g. carotid bruit, atrial fibrillation, valve lesion, evidence of endocarditis, previous emboli or TIA) and hypertension/postural hypotension assessed. Brachial BP should be measured on each side; >20 mmHg difference is suggestive of subclavian artery stenosis.



(a) **(b)**

Fig. 21.23 **CT perfusion scans: (a)** pre-thrombolysis; **(b)** post-thrombolysis showing reperfusion of the ischaemic site. (Courtesy of Professor Adrian Dixon, Cambridge Radiology Department, UK.)

Table 21.29 Thrombolysis in acute ischaemic stroke

Eligibility

Age ≥18 years
Clinical diagnosis of acute ischaemic stroke
Assessed by experienced team
Measurable neurological deficit
Blood tests: results available
CT or MRI consistent with acute ischaemic stroke
Timing of onset well established
Thrombolysis should commence as soon as possible and up to 4.5 hours after acute stroke

Exclusion criteria

Historical

Stroke or head trauma within the prior 3 months
Any prior history of intracranial haemorrhage
Major surgery within 14 days
Gastrointestinal or genitourinary bleeding within the previous 21 days
Myocardial infarction in the prior 3 months
Arterial puncture at a non-compressible site within 7 days
Lumbar puncture within 7 days

Clinical

Rapidly improving stroke syndrome
Minor and isolated neurological signs
Seizure at the onset of stroke if the residual impairments are due to postictal phenomena
Symptoms suggestive of subarachnoid haemorrhage, even if the CT is normal
Acute MI or post-MI pericarditis
Persistent systolic BP >185, diastolic BP >110 mmHg, or requiring aggressive therapy to control BP
Pregnancy or lactation
Active bleeding or acute trauma (fracture)

Laboratory

Platelets <100 000/mm^3
Serum glucose <2.8 mmol/L or >22.2 mmol/L
INR >1.7 if on warfarin
Elevated partial thromboplastin time if on heparin

Dose of i.v. alteplase (tissue plasminogen activator)

Total dose 0.9 mg/kg (maximum 90 mg)
10% of total dose by initial i.v. bolus over 1 minute
Remainder infused i.v. over 60 minutes

Source: amended from Adams HP et al. *Stroke* 2007; **381**: 1655

Routine investigations in thromboembolic stroke and TIA with their potential yields are listed in Box 21.1.

Imaging in acute stroke

CT and MRI:

- *Non-contrast CT* will demonstrate haemorrhage immediately but cerebral infarction is often not detected or only subtle changes are seen initially (Fig. 21.24a).
- MRI shows changes early in infarction (Fig. 21.25a) and a later MRI shows the full extent of the damaged area or penumbra (Fig. 21.25b).
- *Diffusion-weighted MRI (DWI)* can detect cerebral infarction immediately (see Fig. 21.24b) but is as accurate as CT for the detection of haemorrhage.

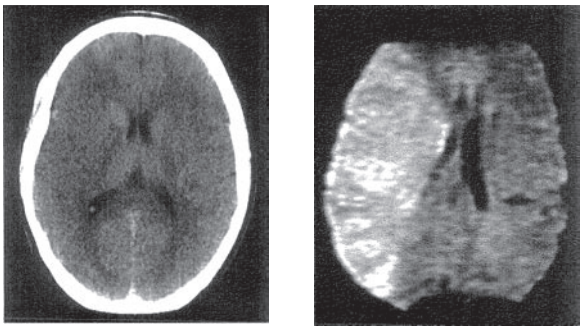
Box 21.1 Stroke: further investigation and management

Further investigations

- Routine bloods (for ESR, polycythaemia, infection, vasculitis, thrombophilia, syphilitic serology, clotting studies, autoantibodies, lipids)
- Chest X-ray
- ECG
- Carotid Doppler studies
- MR angiography, if appropriate

Further management

- Drugs for hypertension, heart disease, diabetes, other medical conditions
- Other antiplatelet agents, e.g. dipyridamole
- Question of endarterectomy
- Question of anticoagulation – Table 21.30
- Speech therapy, dysphagia care, physiotherapy, occupational therapy
- Specific issues, e.g. epilepsy, pain, incontinence
- Preparations for future care



(a) (b)

Fig. 21.24 **Middle cerebral artery infarction. (a)** CT performed initially shows only very subtle low density changes in right MCA territory. **(b)** Diffusion-weighted MRI done at the same time shows full extent of the area of ischaemia. (Courtesy of Dr Paul Jarman.)

CT is still more widely available than MRI and should be performed if MRI is unavailable so that there is no delay in giving thrombolysis for cerebral infarction.

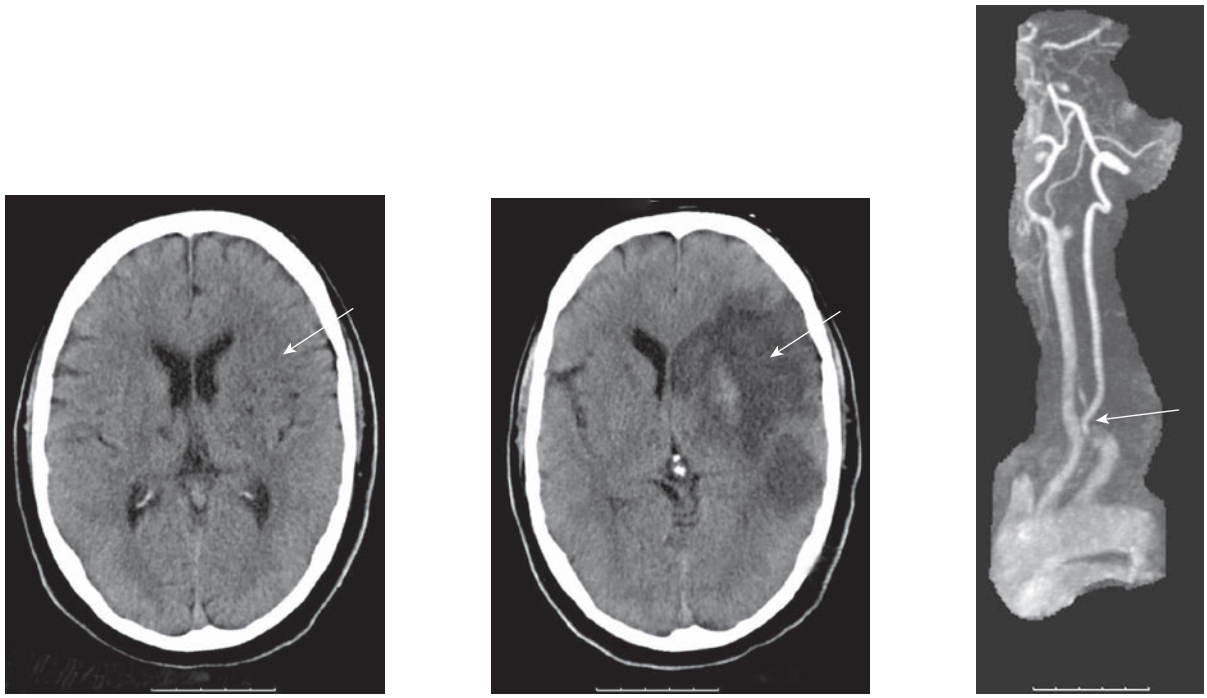
More detailed studies involving perfusion-weighted images and diffusion-weighted MRI will differentiate the infarct core and the penumbral area which is potentially recoverable.

Treatment of acute stroke

This is shown in Box 21.2. Thrombolysis has been shown to improve outcome and should be used immediately if there are no contraindications. In a massive middle cerebral artery infarct, hemispheric swelling occurs with oedema (Fig. 21.26). Decompressive hemicraniectomy reduces the intracranial pressure and the mortality but extensive neurological deficits remain.

Later investigations

MR angiography (MRA) or CT angiography is valuable in anterior circulation TIAs to confirm surgically accessible arterial stenoses, mainly internal carotid stenosis (Fig. 21.25c). If ultrasound suggests carotid stenosis, normotensive patients with TIA or stroke in the anterior circulation should have vascular imaging.



(a) (b) (c)

Fig. 21.25 **MRI of cerebral infarction** and subsequent MRA performed in the same patient. **(a)** Early changes. **(b)** Changes after 5 days. **(c)** MR angiography showing internal carotid artery occlusion. (Courtesy of Dr Paul Jarman.)

Box 21.2 Stroke: immediate management

1. Admit to multidisciplinary stroke unit
2. General medical measures
   Airway: confirm patency and monitor
   Continue care of the unconscious or stuporose patient
   Oxygen by mask
   Monitor BP
   Look for source of emboli
   Assess swallowing
3. Is thrombolysis appropriate?
   If so (Table 21.29) *immediate* brain imaging is essential.
4. Brain imaging
   CT should always be available. This will indicate haemorrhage, other pathology and sometimes infarction. MR is better overall, if immediately available
5. Cerebral infarction
   If CT excludes haemorrhage, give immediate thrombolyte therapy. Aspirin, 300 mg/day is given if thrombolytic therapy is contraindicated
6. Cerebral haemorrhage
   If CT shows haemorrhage, give no drugs that could interfere with clotting. Neurosurgery may occasionally be needed

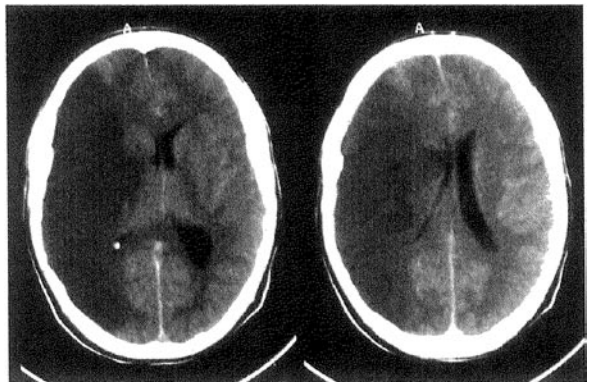


Fig. 21.26 **CT scan of massive infarction in the middle carotid artery territory.** There is hemisphere swelling with brain shift. (Courtesy of Dr Paul Jarman.)

Carotid Doppler and duplex scanning. These screen for carotid (and vertebral) stenosis and occlusion: in skilled hands they demonstrate accurately the degree of internal carotid stenosis.

Long-term management

Medical therapy

Risk factors (Table 21.25) should be identified and addressed.

Antihypertensive therapy

Recognition and good control of high blood pressure is the major factor in primary and secondary stroke prevention. Transient hypertension, often seen following stroke, usually does not require treatment provided diastolic pressure does not rise >100 mmHg. Sustained severe hypertension needs treatment (p. 1805); BP should be lowered slowly to avoid any sudden fall in perfusion.

Antiplatelet therapy (see also p. 443)

Long-term soluble aspirin (75 mg daily) reduces substantially the incidence of further infarction following thromboembolic TIA or stroke. Aspirin inhibits cyclo-oxygenase, which converts arachidonic acid to prostaglandins and thromboxanes; predominant therapeutic effects are reduction of platelet aggregation. Clopidogrel and dipyridamole are also used (p. 443). Combined aspirin 75 mg daily and dipyridamole 200 mg twice daily possibly provide optimal prophylaxis against further thromboembolic stroke or TIA.

Anticoagulants

Heparin and warfarin should be given when there is atrial fibrillation, other paroxysmal dysrhythmias or when there are cardiac valve lesions (uninfected) or cardiomyopathies. Brain haemorrhage must be excluded by CT/MRI. Patients must be aware of the small risk of cerebral (and other) haemorrhage. Anticoagulants are potentially dangerous in the two weeks following infarction because of the risk of provoking cerebral haemorrhage; there are wide differences in clinical practice. Table 21.30 outlines the issues in secondary stroke prevention.

Other measures

Polycythaemia and any clotting abnormalities should be treated (p. 419). Statin therapy should be given for all.

Surgical approaches

Internal carotid endarterectomy

Surgery is usually recommended in TIA or stroke patients with internal carotid artery stenosis >70%. Successful surgery reduces the risk of further TIA/stroke by around 75%. Endarterectomy has a mortality around 3%, and a similar risk of stroke. Percutaneous transluminal angioplasty (stenting) is an alternative. The value of surgery for asymptomatic carotid stenosis is debatable.

Stroke in the elderly

Whilst in the elderly the yield of investigation in stroke diminishes, age is no barrier to recovery; elderly patients benefit the most from good rehabilitation. Consider social isolation, pre-existing cognitive impairment, nutrition, skin and sphincter care, and reassess swallowing. Carotid endarterectomy over 75 years carries little more risk than in younger cases.

Table 21.30 Anticoagulants and stroke prevention

Indication	Comment
Valvular heart disease	Heparin/warfarin of benefit in chronic rheumatic heart disease, particularly mitral stenosis
Recent MI	
Intracardiac thrombus	Heparin/warfarin if there is evidence of intracardiac thrombus
Atrial fibrillation	Anticoagulants long term reduce stroke incidence in atrial fibrillation
Acute internal carotid artery thrombus Acute basilar artery thrombus Internal carotid artery dissection Extracranial vertebral artery dissection	Anticoagulants reserved for imaging-confirmed cases of arterial thrombosis or dissection. They have not been shown to be beneficial in stroke prevention after thromboembolism from carotid or vertebrobasilar sources
Prothrombic states, e.g. protein C deficiency	Consider anticoagulation, in consultation with haematologist
Recurrent TIAs or stroke on full antiplatelet therapy	If no remediable cause, a trial of anticoagulants may be justified
Cerebral venous thrombosis including sinus thrombosis	Benefits of anticoagulation outweigh risks of haemorrhage

Modified from Brown M. *Medicine* 2000; **28**(7): 64.

Rehabilitation: multidisciplinary approach

Skilled physiotherapy has particular value in the first few weeks after stroke to relieve spasticity, prevent contractures and teach patients to use walking aids. The benefits of physiotherapy for longer-term outcome are still inadequately researched. Baclofen and/or botulinum toxin are sometimes helpful in the management of severe spasticity.

Speech therapists have a vital understanding of aphasic patients' problems and frustration. Return of speech is hastened by conversation generally. If swallowing is unsafe because of the risk of aspiration, either nasogastric feeding or percutaneous gastrostomy will be needed. Video-fluoroscopy while attempting to swallow is helpful.

Physiotherapy, occupational and speech therapy have a vital role in assessing and facilitating the future care pathway. Stroke is frequently devastating and, particularly during working life, alters radically the patient's remaining years. Many become unemployable, lose independence and are financially embarrassed. Loss of self-esteem makes depression common.

At home, aids and alterations may be needed: stair and bath rails, portable lavatories, hoists, sliding boards, wheelchairs, tripods, stair lifts, electric blinds and modified sleeping arrangements, kitchen, steps, flooring and doorways. Liaison between hospital-based and community care teams, and primary care physician is essential.

Prognosis

About 25% of patients die within 2 years of a stroke, nearly 10% within the first month. This early mortality is higher following intracranial haemorrhage than thromboembolic infarction. Poor outcome is likely when there is coma, a defect in conjugate gaze and hemiplegia. Many complications, particularly in the elderly, are preventable – for example, aspiration. Coordinated care reduces deaths.

Recurrent strokes are, however, common (10% in the first year) and many patients die subsequently from myocardial infarction. Of initial stroke survivors, some 30–40% remain alive at 3 years.

Gradual improvement usually follows stroke, although late residual deficits are typically substantial. One-third of survivors return to independent mobility and one-third have disability requiring institutional care.

The outlook for recovery of language varies: in general, if language is intelligible at all at 3 weeks, prognosis for fluent speech is good, but many are left with word-finding difficulties.

Intracranial haemorrhage

This comprises:

- intracerebral and cerebellar haemorrhage
- subarachnoid haemorrhage
- subdural and extradural haemorrhage/haematoma.

Intracerebral haemorrhage

Aetiology

Intracerebral haemorrhage causes approximately 10% of strokes. Rupture of microaneurysms (Charcot–Bouchard

aneurysms, 0.8–1.0 mm diameter) and degeneration of small deep penetrating arteries are the principal pathology. Such haemorrhage is usually massive, often fatal, and occurs in chronic hypertension and at well-defined sites – basal ganglia, pons, cerebellum and subcortical white matter.

In normotensive patients, particularly over 60 years, *lobar* intracerebral haemorrhage occurs in the frontal, temporal, parietal or occipital cortex. Cerebral amyloid angiopathy (rare) is the cause in some of these haemorrhages, and the tendency to rebleed is associated with particular apolipoprotein E genotypes.

Recognition

At the bedside, there is no entirely reliable way of distinguishing between haemorrhage and thromboembolic infarction. Both produce stroke. Intracerebral haemorrhage tends to be dramatic with severe headache. It is more likely to lead to coma than thromboembolism.

Brain haemorrhage is seen on CT imaging immediately (cf. infarction, p. 1132) as intraparenchymal, intraventricular or subarachnoid blood. Routine MRI may not identify an acute small haemorrhage correctly in the first few hours but MRI diffusion weighted (MRI DW) is as good as CT.

Management: haemorrhagic stroke

The principles are those for cerebral infarction. The immediate prognosis is less good. Antiplatelet drugs and, of course, anticoagulants are contraindicated. Control of hypertension is vital. Urgent neurosurgical clot evacuation is occasionally necessary when there is deepening coma and coning (particularly in cerebellar haemorrhage).

Cerebellar haemorrhage (Fig. 21.27)

There is headache, often followed by stupor/coma and signs of cerebellar/brainstem origin (e.g. nystagmus, ocular palsies). Gaze deviates towards the haemorrhage. Skew deviation (p. 1105) may develop. Cerebellar haemorrhage sometimes causes acute hydrocephalus, a potential surgical emergency.

Subarachnoid haemorrhage (SAH)

SAH means spontaneous arterial bleeding into the subarachnoid space, and is usually clearly recognizable clinically from its dramatic onset. SAH accounts for some 5% of strokes and has an annual incidence of 6 per 100 000.

Causes

The causes of SAH are shown in Table 21.31; it is unusual to find any contributing disease.

Saccular (berry) aneurysms (Fig. 21.28)

Saccular aneurysms develop within the circle of Willis and adjacent arteries. Common sites are at arterial junctions:

- between posterior communicating and internal carotid artery – posterior communicating artery aneurysm
- between anterior communicating and anterior cerebral artery – anterior communicating and anterior cerebral artery aneurysm
- at the trifurcation or a bifurcation of the middle cerebral artery – middle cerebral artery aneurysm.

Other aneurysm sites are on the basilar, posterior inferior cerebellar, intracavernous internal carotid and ophthalmic arteries. Saccular aneurysms are an incidental finding in 1% of autopsies and can be multiple.

Aneurysms cause symptoms either by spontaneous rupture, when there is usually no preceding history, or by direct pressure on surrounding structures; for example, an enlarging unruptured posterior communicating artery aneurysm is the commonest cause of a painful IIIrd nerve palsy (p. 1105).

Arteriovenous malformation (AVM)

AVMs are collections of arteries and veins of developmental origin. An AVM may also cause epilepsy, often focal. Once an AVM has ruptured, the tendency is to rebleed – 10% will then do so annually.

Cavernous haemangiomas (cavernomas) are common (<0.5% of population) and consist of collections of capillary vessels; they are frequently symptomless (Fig. 21.29) and seen incidentally on imaging. Cavernomas occasionally have a genetic basis, with linkage to two loci on chromosome 7q. Rarely cavernomas cause seizures or bleed; exceptionally they cause sudden death from massive haemorrhage. Surgery is rarely appropriate.

Table 21.31 Underlying causes of subarachnoid haemorrhage

Saccular (berry) aneurysms	70%
Arteriovenous malformation (AVM)	10%
No arterial lesion found	15%
Rare associations (<5%)	
Bleeding disorders	
Mycotic aneurysms – endocarditis	
Acute bacterial meningitis	
Tumours, e.g. metastatic melanoma, oligodendroglioma	
Arteritis (e.g. SLE)	
Spinal AVM → spinal SAH	
Coarctation of the aorta	
Marfan's, Ehlers–Danlos syndrome	
Polycystic kidneys	

Fig. 21.27 **CT: Cerebellar haemorrhage.**

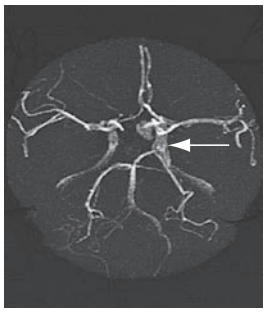


Fig. 21.28 **Digital subtraction angiogram: posterior communicating artery aneurysm.**

Fig. 21.29 **MR T2: Symptomless frontal cavernoma (arrow).**

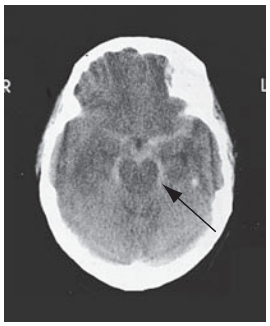


Fig. 21.30 **Subarachnoid haemorrhage CT showing blood around the brainstem (arrow).**

Clinical features of subarachnoid haemorrhage

There is a sudden devastating headache, often occipital. Headache is usually followed by vomiting and often by coma and death. Survivors may remain comatose or drowsy for hours, days, or longer. Less severe headaches, without loss of consciousness, can cause difficulty (p. 1096); SAH is a *possible* diagnosis in any sudden headache.

Following major SAH there is neck stiffness and a positive Kernig's sign. Papilloedema is sometimes present, with retinal and/or subhyaloid haemorrhage (tracking beneath the retinal hyaloid membrane). Minor bleeds cause few signs, but almost invariably headache.

Investigations

CT imaging is the immediate investigation needed (Fig. 21.30). Subarachnoid and/or intraventricular blood is usually seen. Lumbar puncture is not necessary if SAH is confirmed by CT, but should be performed if doubt remains. CSF becomes yellow (xanthochromic) several hours after SAH. Visual inspection of supernatant CSF is usually sufficiently reliable for diagnosis. Spectrophotometry to estimate bilirubin in the CSF released from lysed cells is used to define SAH with certainty. MR angiography is usually performed in all potentially fit for surgery, i.e. generally below 65 years and awake. In some, no aneurysm or source of bleeding is found, despite a definite SAH.

Differential diagnosis

SAH must be differentiated from migraine. This is sometimes difficult. *Thunderclap headache* is used (confusingly) to describe either SAH or a sudden (benign) headache for which no cause is ever found. Acute bacterial meningitis occasionally causes a very abrupt headache, when a meningeal microabscess ruptures; SAH also occasionally occurs at the onset of acute bacterial meningitis. Cervical arterial dissection can present with a sudden headache.

Complications

Blood in the subarachnoid space can lead to obstructive hydrocephalus, seen on CT. Hydrocephalus can be asymptomatic but may cause deteriorating consciousness following after SAH. Shunting may be necessary. Arterial spasm (visible on angiography and a cause of coma or hemiparesis) is a serious complication of SAH and a poor prognostic feature.

Management

Immediate treatment of SAH is bed rest and supportive measures. Hypertension should be controlled. Dexamethasone is often prescribed to reduce cerebral oedema; it also is believed to stabilize the blood–brain barrier. Nimodipine, a calcium-channel blocker, reduces mortality.

All SAH cases should be discussed urgently with a neurosurgical centre. Nearly half of SAH cases are either dead or moribund before reaching hospital. Of the remainder, a further 10–20% rebleed and die within several weeks. Failure to diagnose SAH, e.g. mistaking SAH for migraine, contributes to this mortality.

Patients who remain comatose or who have persistent severe deficits have a poor outlook. In others, who are less impaired, and where angiography demonstrates aneurysm, either a direct approach to clip the aneurysm neck or intravascular coiling is carried out. In selected cases, surgical results are excellent. For AVMs, surgery, microembolism and focal radiotherapy (gamma knife) are used, when appropriate.

Subdural and extradural bleeding

These conditions can cause death following head injuries unless treated promptly.

Subdural haematoma (SDH)

SDH means accumulation of blood in the subdural space following rupture of a vein. This usually follows a head injury, sometimes trivial. The interval between injury and symptoms can be days, or extend to weeks or months. Chronic, apparently spontaneous SDH is common in the elderly, and also occurs with anticoagulants.

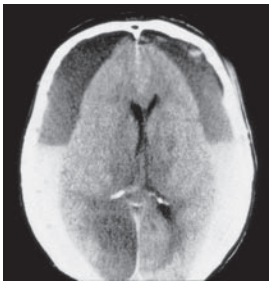
Headache, drowsiness and confusion are common; symptoms are indolent and can fluctuate. Focal deficits, e.g. hemiparesis or sensory loss, develop. Epilepsy occasionally occurs. Stupor, coma and coning may follow.

Extradural haemorrhage (EDH)

EDH typically follows a linear skull vault fracture tearing a branch of the middle meningeal artery. Extradural blood accumulates rapidly over minutes or hours. A characteristic picture is of a head injury with a brief duration of unconsciousness, followed by improvement (the lucid interval). The patient then becomes stuporose, with an ipsilateral dilated pupil and contralateral hemiparesis, with rapid transtentorial coning. Bilateral fixed dilated pupils, tetraplegia and respiratory arrest follow. An acute progressive SDH presents similarly.

Management

Possible extradural or subdural bleeding needs immediate imaging. CT (Fig. 21.31a) is the most widely used investigation because of its immediate availability. MRI is more sensi-



(a)

(b)

Fig. 21.31. **Bilateral subdural haematomas (a) CT scan** (courtesy of Dr Paul Jarman). **(b) MR T1.**

tive for the detection of small haematomas. T1 weighted MRI (Fig. 21.31b) shows bright images due to the presence of methaemoglobin.

EDHs require urgent neurosurgery: if it is performed early, the outlook is excellent. When far from a neurosurgeon, e.g. in wartime or at sea, drainage through skull burr-holes has been lifesaving when an EDH has been diagnosed clinically.

Subdural bleeding usually needs less immediate attention but close neurosurgical liaison is necessary. Even large collections can resolve spontaneously without drainage. Serial imaging is needed to assess progress.

CORTICAL VENOUS THROMBOSIS AND DURAL VENOUS SINUS THROMBOSIS

Intracranial venous thromboses are usually (>50%) associated with a pro-thrombotic risk factor, e.g. oral contraceptives, pregnancy, genetic or acquired pro-thrombotic states and dehydration. Head injury is also a cause. Infection, e.g. from a paranasal sinus, may be present. Venous thromboses can also arise spontaneously.

Cortical venous thrombosis

The venous infarct leads to headache, focal signs (e.g. hemiparesis) and/or epilepsy, often with fever.

Dural venous sinus thromboses

Cavernous sinus thrombosis causes ocular pain, fever, proptosis and chemosis. External and internal ophthalmoplegia with papilloedema develops.

Sagittal and lateral sinus thrombosis cause raised intracranial pressure with headache, fever, papilloedema and often epilepsy.

Management

MRI, MRA and MR venography (MRV) show occluded sinuses and/or veins. Treatment is with heparin initially, though its value is questioned, followed by warfarin for 6 months. Anticonvulsants are given if necessary.

FURTHER READING

Chalela JA, Kidwell CS, Nentwich LM et al. Magnetic resonance imaging and computed tomography in emergency assessment of patients with suspected acute stroke: a prospective comparison. *Lancet* 2007; **369**: 293–298.

Dobkin BH. Rehabilitation after stroke. *New England Journal of Medicine* 2005; **352**: 1677–1684.

Donnan GA, Fisher M, Macleod M et al. Stroke. *Lancet* 2008; **371**:1612–1623.

Hacke W, Kaste M, Bluhmki E et al. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. *New England Journal of Medicine* 2008; **359**: 1319–1329.

Khaja AM, Grotta JC. Established treatments for acute ischaemic stroke. *Lancet* 2007; **369**: 319–330.

Royal College of Physicians' Intercollegiate Stroke Working Party. *National Clinical Guidelines for Stroke*, 2nd edn. London: Royal College of Physicians, 2008.

Stam J. Thrombosis of the cerebral veins and sinuses. *New England Journal of Medicine* 2005; **352**: 1791–1798.

Strong K, Mather C, Bonita R. Preventing stroke; saving lives around the world. *Lancet Neurology* 2007; **6**:182–187 + Series.

van Gijn J, Kerr RS, Rinkel GJ. Subarachnoid haemorrhage. *Lancet* 2007; **369**: 306–318.

HEADACHE, MIGRAINE AND FACIAL PAIN

Headache is one of the commonest symptoms in primary care and a frequent reason for hospital referral. Most headaches are innocent, reflecting no sinister intracranial disease. Headache symptoms are unpleasant, disabling, common world-wide and have a substantial economic impact because of time lost from work (p. 1096 and Table 21.1). There is an internationally agreed classification for headaches that defines all headache patterns in detail.

Tension headache

The vast majority of chronic daily headaches and recurrent headaches are thought to be generated by neurovascular irritation and referred to scalp muscles and soft tissues, although the exact pathogenesis remains unclear. Tight band sensations, pressure behind the eyes, throbbing and bursting sensations are common. What is clear is that almost all headaches with these features are benign.

There may be obvious precipitating factors such as worry, noise, concentrated visual effort or fumes. Depression is also a frequent co-morbid feature. Tension headaches are often attributed to cervical spondylosis, refractive errors or high blood pressure: evidence for such associations is poor. Headaches also follow even minor head injuries. Tenderness and tension in neck and scalp muscles are the only physical signs. Analgesic overuse is a prominent cause of headache.

Management

Headache management involves:

- explanation (imaging is often needed)
- avoiding evident causes, e.g. bright lights
- analgesic withdrawal
- physical treatments – massage, ice packs, relaxation
- antidepressants – when indicated
- drugs for recurrent headache/migraine.

Migraine

Migraine is recurrent headache associated with visual and gastrointestinal disturbance. The borderland between migraine and tension headaches can be indistinct. Over 20% of any population world-wide report migrainous symptoms; in 90%, these began before 40 years of age.

Mechanisms

Precise mechanisms remain unclear. Genetic factors play some part – a rare form of familial migraine is associated with mutation in the alpha-1 subunit of the P/Q-type voltage-gated calcium channel on chromosome 19.

The pathophysiology of migraine is now thought to involve changes in the brainstem blood flow which have been found on PET scanning during migraine attacks. This leads to an unstable trigeminal nerve nucleus and nuclei in the basal thalamus. This results in release of calcitonin-related peptide (CGR8), substance P and other vasoactive peptides, leading to neurogenic inflammation, which gives rise to pain, and vasodilation of cerebral and dural vessels which also contribute towards the headache.

Cortical spreading depression is also proposed as a mechanism for the aura.

Some patients recognize precipitating factors:

- weekend migraine (a time of relaxation)
- chocolate (high in phenylethylamine)
- cheese (high in tyramine)
- noise and irritating lights
- association with premenstrual symptoms.

Migraine is common around puberty and at the menopause and sometimes increases in severity or frequency with hormonal contraceptives, in pregnancy and occasionally with the onset of hypertension or following minor head trauma. Migraine is not suggestive of any serious intracranial lesion. However, since migraine is so common, an intracranial mass and migraine sometimes occur together by coincidence.

Clinical patterns

Migraine attacks vary from intermittent headaches indistinguishable from tension headaches to discrete episodes that mimic thromboembolic cerebral ischaemia. Distinction between variants is somewhat artificial. Migraine can be separated into phases:

- well-being before an attack (occasional)
- prodromal symptoms
- the main attack – headache, nausea, vomiting
- sleep and feeling drained afterwards.

Migraine with aura (classical migraine)

Prodromal symptoms are usually visual and related to depression of visual cortical function or retinal function. Unilateral patchy scotomata (retina), hemianopic symptoms (cortex), teichopsia (flashes) and fortification spectra (jagged lines like battlements) are common. Transient aphasia sometimes occurs, with tingling, numbness, vague weakness of one side and nausea. The prodrome persists for a few minutes to about an hour. Headache then follows. This is occasionally hemicranial (i.e. splitting the head) but often begins locally and becomes generalized. Nausea increases and vomiting follows. The patient is irritable and prefers the dark. Superficial scalp arteries are engorged and pulsating. After several hours the migraine settles, sometimes with a diuresis. Deep sleep often ensues.

Migraine without aura (common migraine)

This is the usual variety. Prodromal visual symptoms are vague. There is a similar headache often accompanied by nausea and malaise.

Basilar migraine

Prodromal symptoms include circumoral and tongue tingling, vertigo, diplopia, transient visual disturbance (even blindness), syncope, dysarthria and ataxia. These occur alone or progress to a typical migraine.

Hemiparetic migraine

This rarity is classical migraine with hemiparetic features, i.e. resembling a stroke, but with recovery within 24 hours. Exceptionally, cerebral infarction occurs.

Ophthalmoplegic and facioplegic migraine

These rarities are a IIIrd, VIth or VIIth nerve palsy with a migraine, and they are difficult to diagnose without investigation to exclude other conditions.

Differential diagnosis

A sudden migraine headache may resemble SAH or the onset of meningitis. Hemiplegic, visual and hemisensory symptoms must be distinguished from thromboembolic TIAs (p. 1128). In TIAs maximum deficit is present immediately and headache is unusual. Unilateral tingling or numbness may resemble sensory epilepsy (partial seizures). In epilepsy, distinct march (progression) of symptoms is usual.

Management

General measures include:

- explanation
- avoidance of dietary factors – rarely helpful.

Patients taking hormonal contraceptives may benefit from a brand change, or trying without. Depot oestrogens are sometimes used. Severe hemiparetic symptoms are a potential reason to stop hormonal contraceptives. Premenstrual migraine sometimes responds to diuretics.

At the start of an attack. Paracetamol or other analgesics should be taken, with an antiemetic such as metoclopramide if necessary. Repeated use of analgesics leads to further headaches. Triptans ($5HT_1$ agonists) are also widely used, sometimes aborting an attack effectively. Sumatriptan was the first marketed; almotriptan, eletriptan, frovatriptan, naratriptan, rizatriptan and zolmitriptan are now available, with various routes of administration. Triptans should be avoided when there is vascular disease, and not overused.

Prophylaxis. The following are used continuously when attacks are frequent:

- pizotifen (5HT antagonist) 0.5 mg at night for several days, increasing to 1.5 mg (common side-effects are weight gain and drowsiness)
- propranolol 10 mg three times daily, increasing to 40–80 mg three times daily
- amitriptyline 10 mg (or more) at night.

Sodium valproate, methysergide, SSRIs, verapamil, topiramate, nifedipine and naproxen are also used. Gap junction blockers are being used in trials.

Other benign headaches

- *Ice-cream headache*. Sufferers describe intense, retropharyngeal head pain lasting for a few seconds or minutes following ice-cream or very cold foods.
- *Primary cough headache* is a sudden sharp head pain on coughing. No underlying cause is found but intracranial pathology should be excluded. The problem often resolves spontaneously. Very rarely, for severe headache, a lumbar puncture with removal of CSF can help.
- *Primary low CSF volume headaches,* seen typically on standing up, are also well recognized. The patient may give a history of some event, such as a vigorous Valsalva, straining or orgasm, but these headaches sometimes arise spontaneously. Treatment with an autologous intrathecal blood patch can be helpful. Secondary low CSF volume can follow lumbar puncture (p. 1121).

- *Primary sex headache* describes varieties of head pain that typically rise to a crescendo at orgasm, largely in males. Treatment with propranolol or diltiazem is said to be helpful, but these pains often resolve spontaneously after several attacks. Exceptionally, sex headaches occur with an unruptured intracranial aneurysm.
- Many other varieties of primary headache are listed in the international classification, e.g. hemicrania continua, primary stabbing headache, primary exertional headache, hypnic headache, and primary thunderclap headache.
- *Post-traumatic headache* is also a common problem. Headaches do sometimes follow a minor blow to the head; they tend to resolve, typically within 6–8 weeks. However, when there is third party involvement, and especially with litigation, these headaches can persist for long periods. Opinions vary about their cause.

Facial pain

The face has many pain-sensitive structures: teeth, gums, sinuses, temporomandibular joints, jaw and eyes. Facial pain is also caused by specific neurological conditions.

Trigeminal neuralgia (p. 1106), trigeminal nerve lesions (p. 1106) and postherpetic neuralgia (p. 1157) are described elsewhere.

Cluster headache

Cluster headache (migrainous neuralgia and trigeminal autonomic cephalalgia type I), is distinct from migraine. It describes recurrent bouts of excruciating unilateral pain that typically wake the patient. Attacks cluster around one eye. Cluster headache affects adults, mostly males aged between 30 and 40. Alcohol sometimes provokes an attack, and experimentally, glyceryl trinitrate. During an attack, changes in the hypothalamus appear on MRI and PET. PET, presumed to be due to activation of the trigeminal-vascular and autonomic systems. Severe pain rises to an even worse crescendo over some minutes, lasting for several hours. Vomiting can occur. One cheek and nostril become congested. Transient ipsilateral Horner's syndrome is common. One bout of cluster attacks, with pain every few nights, usually lasts one to two months. Despite excruciating pain there are no sequelae. Bouts recur at intervals over several years but tend to disappear after the age of 55.

Management. Analgesics are unhelpful. Subcutaneous sumatriptan is the drug of choice. Oxygen inhalation sometimes aborts an attack. Most prophylactic migraine drugs are unhelpful. Verapamil, topiramate, lithium carbonate and/or a short course of steroids sometimes help bring to an end a bout of cluster headaches.

Paroxysmal hemicrania

Episodic paroxysmal hemicrania (trigeminal autonomic cephalalgia type II) is a rare condition describing unilateral sudden, brief (<20 min) pains with some characteristics of cluster headache. Paroxysms can occur many times each day; typically they respond to indometacin.

SUNCT and SUNA (trigeminal autonomic cephalalgia type III) are eponyms for other, rare and exceedingly painful sudden headaches.

- In SUNCT *(short-lasting unilateral neuralgiform headache with conjunctival injection and tearing)* attacks last from 5 seconds to 2 minutes.
- In SUNA *(short-lasting unilateral neuralgiform headache with cranial autonomic symptoms)* the pain is similar to SUNCT but there is no tearing or conjunctival injection.

Diagnosis is made from the clinical features. Preventative treatment with lamotrigine, topiramate or carbamazepine is sometimes helpful. In hospital, immediate i.v. lidocaine often aborts an attack.

Exceptionally, SUNCT-like symptoms have been described with pituitary tumours, and with posterior fossa tumours.

Atypical facial pain

Facial pain for which no cause can be found is seen in the elderly, mainly in women, and is believed (on little evidence) to be a somatic equivalent of depression. Tricyclic antidepressants are sometimes helpful.

Other causes of facial pain

Facial pain occurs in variants of migraine and in giant cell arteritis (see below).

Giant cell arteritis (temporal arteritis; cranial arteritis)

These are granulomatous arteritides seen almost exclusively in people over 50 (p. 550).

Clinical features

- **Headache** is almost invariable in giant cell arteritis (GCA). Pain develops over inflamed superficial temporal and/or occipital arteries. Touching the skin over an inflamed vessel (e.g. combing hair) causes pain. Arterial pulsation is soon lost; the artery becomes hard, tortuous and thickened. The scalp over inflamed vessels may become red. Rarely, gangrenous patches appear.
- **Facial pain**. Pain in the face, jaw and mouth is caused by inflammation of facial, maxillary and lingual branches of the external carotid artery in GCA. Pain is characteristically worse on eating (jaw claudication). Mouth opening and protruding the tongue become difficult. A painful, ischaemic tongue occurs rarely.
- **Visual problems**. Visual loss from arterial inflammation and occlusion occurs in 25% of untreated GCA cases. Posterior ciliary artery occlusion causes anterior ischaemic optic neuropathy in three-quarters of these. Other mechanisms are central retinal artery occlusion, cilioretinal artery occlusion and posterior ischaemic optic neuropathy. There is sudden monocular visual loss (partial or complete), usually painless. Amaurosis fugax (p. 1128) may precede permanent blindness.

  When the posterior ciliary vessels are affected, ischaemic optic neuropathy causes the disc to become swollen and pale; retinal branch vessels usually remain normal. When the central retinal artery is occluded, there is sudden permanent unilateral blindness, disc pallor and visible retinal ischaemia. Bilateral blindness may develop, with the second eye being affected in 1–2 weeks.

- **Rare complications**. Brainstem ischaemia, cortical blindness, ischaemic microangiopathic neuropathy of peripheral or cranial nerves, and involvement of the aorta, coronary, renal and mesenteric arteries sometimes occurs.

Management

The ESR is greatly elevated. The diagnosis should be established immediately by superficial temporal artery biopsy, because of the risk of blindness. Immediate high doses of steroids (prednisolone, initially 60–100 mg daily) should be started in a patient with typical features, even before biopsy (p. 550). Since the risk of visual loss persists, long-term treatment is recommended, for some years at least.

FURTHER READING

Dodick D. Chronic daily headache. *New England Journal of Medicine* 2006; **354**: 158–165.

Headache Classification Committee. The International Classification of Headache Disorders, 2nd edn. *Cephalgia* 2004; **24** (Suppl 1): 1–160.

Salvarani C, Cantini F, Boiardi L et al. Polymyalgia rheumatica and giant-cell arteritis. *Lancet* 2008; **372**: 234–245.

Wessman M. Migraine: a complex genetic disorder. *Lancet Neurology* 2007; **6**: 521–532.

EPILEPSY AND LOSS OF CONSCIOUSNESS

EPILEPSY

'Seizure' means a convulsion or other transient event caused by a paroxysmal discharge of cerebral neurones. Epilepsy is the continuing tendency to have seizures, even if a long interval separates attacks. A generalized convulsion (*grand mal fit*) is the most common recognized event.

Epilepsy is common. Its prevalence is 5 times higher in developing than developed (0.5%) countries, and the incidence is doubled. Over 2% of the population in developed countries have two or more seizures during their lives; in 0.5% epilepsy is an active problem and common in primary care. In the UK approximately 65 people suffer a probable first seizure each day; the lifetime risk of having a single seizure is 5%. Often no clear cause is found for seizures, although rarely epilepsy is caused by a brain tumour or follows a stroke. Approximately 250 000 people in the UK take anticonvulsant drugs, the mainstay of treatment. Neurosurgery (e.g. temporal lobectomy) is highly effective in carefully selected cases.

Mechanisms and definitions

Spread of electrical activity between neurones is normally restricted and synchronous discharge of neurones takes place in confined groups, producing normal EEG rhythms. During a seizure, large groups of neurones are activated repetitively, unrestrictedly and hypersynchronously; synaptic inhibition between them fails. High-voltage spike-and-wave activity is the result, epilepsy's EEG hallmark.

A *partial seizure* is epileptic activity confined to one area of cortex with a recognizable clinical pattern (Fig. 21.32). This activity either remains focal or spreads to generate epileptic activity in both hemispheres and thus a *generalized* seizure. This spread is called *secondary generalization*. The focal nature of a seizure may not be apparent clinically because of rapid secondary generalization – an *apparent* generalized tonic–clonic seizure may either have started as a focal seizure or be a *primary* generalized convulsion. Brain becomes epileptogenic either because neurones are predisposed to be hyperexcitable, for example following abnormal neuronal migration in utero, or because they acquire this tendency. Trauma and brain neoplasms are examples of acquired conditions that alter neuronal seizure threshold.

Aura means a stereotyped perception caused by initial focal electrical events before a partial seizure, such as a smell, tingling in one limb, or strange recognizable inner feelings.

Seizure threshold

Each person has a seizure threshold. Experimentally some chemicals (e.g. pentylenetetrazol, a toxic gas) induce seizures in everyone. A *low seizure threshold* describes the tendency, for example, to have seizures provoked by flashing lights; it is a concept, not a measurement.

Classification

Various terms are used to classify epilepsy, e.g. idiopathic, symptomatic, cryptogenic, primary, focal or generalized. Here attacks are described by *clinical pattern* (Table 21.32).

- *Generalized* means bilateral abnormal electrical activity, with bilateral motor manifestations and impaired consciousness.
- *A partial (focal) seizure* means the electrical abnormality is localized to one part of the brain:
  (a) simple – without loss of awareness, e.g. one limb jerking (Jacksonian seizure)
  (b) complex – with loss of awareness, e.g. a temporal lobe attack.

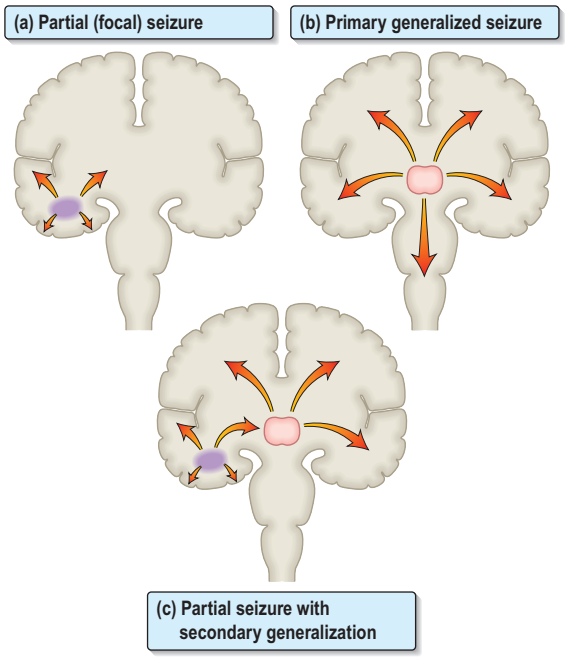


Fig. 21.32 **Seizure types.**

Table 21.32 The commoner types of seizure

Abridged from International League against Epilepsy

1. **Generalized seizure types**
   - A Absence seizures
     - (a) Typical absences with 3 Hz spike-and-wave discharge (petit mal)
     - (b) Atypical absences with other EEG changes
   - B Generalized tonic–clonic seizures (grand mal)
   - C Myoclonic seizures
   - D Tonic seizures
   - E Akinetic seizures
2. **Partial seizure types**
   These start by neurone activation in one part of one hemisphere (focal seizures)
   - A Simple partial seizures (without impaired awareness, e.g. Jacksonian seizures)
   - B Complex partial seizures (with impaired awareness)
   - C Partial seizures evolving to tonic–clonic seizures
   - D Apparent generalized tonic–clonic seizures, with EEG but not clinical evidence of focal onset
3. **Unclassifiable seizures**
   Seizures that do not fit a category above

Table 21.33 Epilepsy: aetiological factors

- Genetic predisposition
- Developmental, e.g. hamartomas, neuronal migration abnormalities
- Hippocampal sclerosis
- Brain trauma and surgery
- Pyrexia
- Intracranial mass lesions, e.g. tumour, neurocysticercosis
- Vascular, e.g. cerebral infarction, AVM
- Drugs and drug withdrawal
- Encephalitis and inflammatory conditions, e.g. herpes simplex, MS
- Metabolic abnormalities, e.g. porphyria, hypocalcaemia
- Neural degenerative disorders, e.g. Alzheimer's
- Provoked seizures, e.g. photosensitivity, sleep deprivation
- Drugs, e.g. ciclosporin, lidocaine, quinolones, SSRIs, interferons, cocaine, lithium, withdrawal of amfetamines, barbiturates
- Alcohol withdrawal

Generalized seizure types

Typical absence seizures (petit mal)

This generalized epilepsy almost invariably begins in childhood. Each attack is accompanied by 3 Hz spike-and-wave EEG activity (Fig. 21.15, p. 1120). Activity ceases, the patient stares and pales slightly for a few seconds. The eyelids twitch; a few muscle jerks may occur. After an attack, normal activity is resumed. Typical absence attacks are never due to acquired lesions such as tumours. They are a developmental abnormality of neuronal control. Children with petit mal attacks tend to develop generalized tonic–clonic seizures in adult life (known as primary generalized epilepsy). Petit mal describes these 3 Hz absence seizures only; clinically similar episodes can also be produced by temporal lobe foci (partial seizures), a source of confusion.

Generalized tonic–clonic seizures (GTCS, grand mal seizures)

Following a vague warning, the tonic phase commences. The body becomes rigid, for up to a minute. The patient utters a cry and falls, sometimes suffering serious injury. The tongue is usually bitten. There may be incontinence of urine or faeces.

The clonic phase then begins, with generalized convulsing, frothing at the mouth and bilateral, rhythmic jerking of muscles. This lasts from a few seconds to several minutes. Seizures are usually self-limiting, followed by drowsiness, confusion or coma for several hours.

Myoclonic, tonic and akinetic seizures

Myoclonic seizures describe isolated muscle jerking. *Juvenile myoclonic epilepsy* is one variety, seen in adolescence. Tonic means intense stiffening of the body, not followed by jerking. *Akinetic* means cessation of movement, falling and loss of consciousness.

Partial seizure types

Simple partial seizures

One example is a *focal motor* seizure (*Jacksonian*). These simple partial seizures originate within the motor cortex. Jerking typically begins on one side of the mouth or in one hand, sometimes spreading to involve the entire side. This visible spread of activity is called the *march* of a seizure. With a frontal lesion, conjugate gaze (p. 1105) deviates away from the epileptic focus and the head turns; this is known as an adversive seizure. Local temporary paralysis of the limbs affected sometimes follows – Todd's paralysis.

Temporal lobe seizures

These partial seizures, either simple or complex, can produce feelings of unreality (jamais vu) or undue familiarity (déjà vu) with the surroundings. Blank episodes of staring, vertigo, visual hallucinations (i.e. visions or faces) are other examples of temporal lobe seizures.

Many other types of partial seizure occur: panic attacks with overbreathing (temporal), strange smells (frontal), sensory disturbances (parietal) and crude visual shapes (occipital).

Aetiology and precipitants (Table 21.33)

A cause for epilepsy is found in less than one-third of cases in UK surveys: cerebrovascular disease accounts for less than 15%, tumours for 6%, alcohol-related seizures for 6% and post-traumatic epilepsy 2%. Hippocampal sclerosis is a cause of symptomatic epilepsy (resection may be possible); others are malformations of cortical development, AVMs, hamartomas, brain abscesses and tuberculomas. Neurocysticercosis is a major cause of epilepsy where cysticercosis is endemic (p. 170).

Genetic predisposition and developmental anomalies

Over 200 genetic disorders list epilepsy among their features, with complex syndromic classifications. These account for less than 2% of epilepsy. About 30% of epilepsy patients have first-degree relatives with seizures. Usually the mode of inheritance is uncertain. A low seizure threshold seems to run in some families. Petit mal and other primary generalized epilepsies are often inherited as an autosomal dominant trait with variable penetrance. Primary epilepsies are due to complex developmental abnormalities of neuronal control; there are abnormalities in synaptic connections, and anoma-

lies in neurotransmitter distribution and release. Neuronal migration defects in utero, dysplastic areas of cerebral cortex and hamartomas contribute to seizures both in infancy and adult life.

Trauma, hypoxia and surgery

Perinatal trauma (cerebral contusion and haemorrhage) and fetal anoxia are causes of childhood seizures. Hypoxic damage in childhood to the hippocampi leads to hippocampal sclerosis, another prominent cause. Brain injury is sometimes followed by seizures within the first week (early epilepsy) or many months or years later (late epilepsy). To cause epilepsy, the injury must (usually) be sufficient to cause coma. Early epilepsy, a depressed skull fracture, penetrating brain injury, cerebral contusion, dural tear or intracranial haematoma increase the incidence of late post-traumatic epilepsy.

Brain surgery is followed by seizures in 3–5%.

Brain tumours (and abscesses)

Mass lesions in the cortex cause epilepsy – either partial or secondary generalized seizures. If epilepsy develops in adult life, the chance of finding an unsuspected tumour is around 3%.

Hydrocephalus (p. 1162) also lowers seizure threshold.

Vascular and degenerative brain disorders

Seizures sometimes follow cerebral infarction; there is a peak in incidence late in life. A brain AVM can present with seizures, as occasionally can SAH. Seizures can occur in Alzheimer's disease and in many degenerative diseases. Epilepsy is three times commoner in MS patients than in the general population.

Encephalitis and inflammatory conditions

Seizures are frequent features of encephalitis, cerebral abscess, tuberculoma, cortical venous thrombosis and neurosyphilis. They also occur in chronic meningitis (e.g. TB) and may rarely be the first sign of acute bacterial meningitis. Neurocysticercosis is a prominent cause of seizures in countries where the pork tapeworm is endemic, e.g. India, South America, and is a major public health problem.

Alcohol and drugs

Chronic alcohol abuse is a common cause of seizures. These occur either while drinking heavily or during periods of withdrawal. Alcohol-induced hypoglycaemia also provokes attacks (p. 1060).

Phenothiazines, monoamine oxidase inhibitors, tricyclic antidepressants, SSRIs, amfetamines, lidocaine, propofol and nalidixic acid sometimes provoke fits, either in overdose or at therapeutic doses in individuals with a low seizure threshold.

Withdrawal of antiepileptic drugs (especially phenobarbital) and benzodiazepines may provoke seizures.

Metabolic abnormalities

Seizures can follow:

- hypocalcaemia, hypoglycaemia, hyponatraemia
- acute hypoxia
- uraemia, hepatocellular failure
- mitochondrial disease, porphyria.

Provoked seizures: photosensitivity, pyrexia

Seizures are occasionally precipitated by flashing lights or a flickering television screen. Photosensitivity may be recorded on occipital EEG electrodes. Rarely other stimuli (e.g. music) provoke attacks. High pyrexia can provoke convulsions in children under 5 years (febrile convulsions). In most, there is no recurrence; febrile convulsions are not labelled as epilepsy.

Sleep deprivation

A convulsion sometimes follows missing a night's sleep in a susceptible person.

Diagnosis

The history from a witness is crucial, and usually enables one to distinguish other causes of disturbed consciousness (p. 1144). The onset, setting and stages of attacks are of importance. Neurological examination may be normal or point to a clinical diagnosis. General screening, including serum calcium, and an ECG (rhythm, conduction abnormalities, QT interval) should be done.

Electroencephalography

EEG is a useful test, despite limitations. It is usually performed following a first fit, though the therapeutic yield is low.

- *During a seizure* the EEG is almost invariably abnormal, because spikes reach electrodes overlying brain.
- *EEG evidence of seizure activity* is shown typically by focal cortical spikes (e.g. over a temporal lobe) or by generalized spike-and-wave activity. Epileptic activity is continuous in status epilepticus.
- *3 Hz spike-and-wave* occurs specifically in petit mal (Fig. 21.15, p. 1120), always during an attack and frequently in between them.
- *A normal EEG between attacks* (interictal) does not exclude epilepsy; many people with epilepsy have normal interictal EEGs.
- *An abnormal interictal EEG* does not prove that a particular attack was epileptic.
- *EEG videotelemetry* is vital for studying attacks of uncertain nature (e.g. non-epileptic attacks, p. 1144).

CT and/or MR imaging

Imaging is indicated in all new cases. CT is a reasonable urgent screening test for tumours in adults. MR, with imaging of the hippocampi, is used routinely to study epilepsy.

Treatment

Emergency measures

When faced with a seizure it is best simply to ensure that the patient comes to as little harm as possible and that the airway is maintained both during a prolonged seizure and in postictal coma. Wooden mouth gags, tongue forceps and physical restraint cause injury.

Most seizures last only minutes and end spontaneously. A prolonged seizure (longer than 3 minutes) or repeated seizures are treated with rectal (10 mg) or i.v. diazepam or midazolam. If there is any suspicion of hypoglycaemia, take blood for glucose and give i.v. glucose. Serial epilepsy describes repeated seizures with brief periods of recovery. These may lead to status epilepticus. Sudden death in a seizure is unusual but does occur.

Ask the authors
www.studentconsult.com

Status epilepticus

This medical emergency (Practical box 21.4) means continuous seizures without recovery of consciousness. It consists of prolonged serial seizures (two or more) occurring with incomplete recovery of consciousness. Status epilepticus has a mortality of 10–15%. Over 50% of cases occur without a previous history of epilepsy. Some 25% with apparent refractory status have *pseudostatus* (non-epileptic attack disorder); iatrogenic errors and morbidity are significant.

Not all status is convulsive. In *absence status*, for example, status is non-convulsive – the patient is in a continuous, distant, stuporose state. *Focal status* also occurs. *Epilepsia partialis continua* is continuous seizure activity in one part of the body, such as a finger or a limb, without loss of consciousness. This is often due to a cortical neoplasm or, in the elderly, a cortical infarct.

Practical Box 21.4 Status epilepticus – management

Of prime importance:

- Treat convulsions quickly
- Accuracy of diagnosis
- Continued ITU monitoring and cardiorespiratory support

Several treatment schedules exist:

- At home, give immediate diazepam 10–20 mg i.v. at 5 mg/min and repeat once. If i.v. access is impossible, give rectal diazepam or paraldehyde
- Arrange immediate admission
- Administer oxygen, monitor ECG, BP, routine bloods (include alcohol, sugar, calcium, magnesium, drug screen, anticonvulsant levels)
- Give thiamin i.v. (250 mg) if nutrition is poor or alcohol abuse suspected. (In the UK, give vitamin B and C, high-potency ampoules, one pair i.v. over 10 mins)
- Antiepileptic drugs (AEDs):
  1. Give lorazepam i.v. 4 mg at 2 mg/min
  2. Reinstate previous AEDs. Measure levels urgently. Has the patient had phenytoin recently?
  3. If status continues, i.v. phenytoin or fosphenytoin is used
     - Phenytoin: give 15 mg/kg i.v. diluted to 10 mg/mL in 0.9% saline into a large vein at <50 mg/min (Phenytoin 250 mg 5 mL ampoule)
     - Fosphenytoin: this is a pro-drug of phenytoin and can be given faster than phenytoin. Doses are expressed in phenytoin equivalents (PE): fosphenytoin 1.5 mg = 1 mg phenytoin
     - Give 15 mg/kg (PE) fosphenytoin (15 mg × 1.5 = 22.5 mg) diluted to 10 mg/mL in 0.9% saline at 50–100 mg (PE)/min

     (Fosphenytoin sodium 750 mg 10 mL ampoule)
  4. If status continues, give phenobarbital 10 mg/kg diluted 1 in 10 in water for injection at <100 mg/min. (Phenobarbital 200 mg/mL 1 mL vial in propylene glycol 90% with water for injection 10%.) Intravenous clonazepam, paraldehyde and clomethiazole are also used
  5. If status persists >90 minutes, use thiopental or propofol anaesthesia with assisted ventilation
- EEG monitoring is valuable if there is doubt about the nature of status
- CT may reveal an underlying cause
- Remember: 25% of apparent status turns out to be pseudostatus
- Remember: potential unwanted effects of drugs, e.g. hypotension, cardiorespiratory arrest, and the need for continuous monitoring

Antiepileptic drugs (AEDs)

AEDs are indicated when there is a firm clinical diagnosis of epilepsy and a substantial risk of recurrent seizures. AED use carries the stigma of epilepsy. Concordance between doctor and patient in taking the drugs is essential, as is understanding of potential unwanted effects. For both partial and generalized seizures, monotherapy with an established first-line AED is the initial choice (Table 21.34). Doses should start low and be increased until control is achieved or tolerance exceeded. If control is not achieved, a second drug is added. There remain different views about the most appropriate drugs for each seizure type.

Drug levels. Serum levels of all AEDs can be measured. With phenytoin the therapeutic range is well defined and the level should usually be monitored every few months (Table 21.35). Monitoring serum levels of other AEDs is less useful. Levels are not usually measured routinely unless compliance or toxicity is an issue.

Unwanted effects of drugs. Intoxication with all AEDs causes ataxia, nystagmus and dysarthria. Chronic phenytoin causes gum hypertrophy, hypertrichosis, osteomalacia, folate deficiency, polyneuropathy and encephalopathy. A wide range of potential unwanted effects are known. See Table 21.36 for some serious idiosyncratic (i.e. non-dose-related) side-effects. The majority of severe skin reactions (e.g. toxic epidermal necrolysis) following phenytoin, carbamazepine, valproate (and phenobarbital) occur within the first 8 weeks of treatment.

Two drugs rarely commenced outside a specialist centre are phenobarbital and its derivative primidone (drowsiness, cognitive impairment), though phenobarbital still has a place in status. Vigabatrin can cause irreversible visual field defects and is no longer advocated.

Refractory epilepsy

Despite therapy, seizures persist in some 20–35% of cases. Rigorous attention to diagnosis, to concordance (compliance) and to trials of different drugs can sometimes reduce seizure frequency. Many cases of refractory seizures are treated in specialist units: the question of surgery can be reviewed.

Epilepsy in the elderly

Some 25% of new cases of epilepsy develop over the age of 65. Many patients have cerebrovascular disease, neurodegenerative conditions or brain tumours. The onset of seizures commonly leads to loss of independence and to physical injuries with their complications (e.g. fractures, subdural haematoma). With adequate AEDs, seizure control can be achieved in some 70% of this vulnerable population.

Women, epilepsy, pregnancy and AEDs

Fertility. There is some reduction of fertility in women with epilepsy. This is multifactorial. One-third of women with epilepsy have some ovarian abnormality: irregular menstrual cycles, anovulatory cycles and polycystic ovaries. These are possibly more frequent in patients taking valproate.

Birth defects. The overall risk of birth defects in babies of mothers who take one AED is around 7%, higher than the 3% in the population. Counselling before conception is essential. Some women choose to stop AEDs before becom-

Table 21.34 Principal antiepileptic drugs and common seizure types

	Generalized tonic–clonic seizures (grand mal) and partial	Petit mal	Myoclonic	Atypical absence, tonic and akinetic
First-line	Carbamazepine Lamotrigine Phenytoin Valproate	Ethosuximide Valproate	Valproate Clonazepam Lamotrigine	Valproate Lamotrigine Clonazepam
Second-line and/or add-ons	Phenobarbital Primidone Levetiracetam Topiramate Newer drugs (Table 21.35)	Clobazam Clonazepam Topiramate Lamotrigine	Clobazam Levetiracetam Topiramate	Topiramate Clobazam Levetiracetam Phenobarbital Acetazolamide
May worsen attacks		Carbamazepine Phenytoin	Carbamazepine Phenytoin	

Other drugs usually initiated in specialist centres are gabapentin, oxcarbazine, tiagabine, zonisamide, pregabalin.

Table 21.35 Antiepileptic drugs: doses and therapeutic levels

	Usual adult daily dose (mg)	Therapeutic range (μmol/L)
Phenytoin	300–450	40–80
Carbamazepine	400–1000	20–50
Valproate	800–2000	200–700

Table 21.36 Antiepileptic drugs: some idiosyncratic unwanted effects

Drug	Non-dose-related effects
Phenytoin	Rashes Blood dyscrasias Lymphadenopathy Systemic lupus erythematosus Toxic epidermal necrolysis
Carbamazepine	Rashes Blood dyscrasias, e.g. severe leucopenia Toxic epidermal necrolysis
Valproate	Anorexia Hair loss Liver damage
Lamotrigine	Toxic epidermal necrolysis

ing pregnant. If drugs are continued, monotherapy with a first-line drug is preferable with folic acid (5 mg/day) supplement. Careful antenatal screening (p. 45) is required. Vitamin K 20 mg orally should also be taken daily during the week before delivery to prevent neonatal haemorrhage (caused by inhibition of vitamin K transplacental transport).

Contraception. AEDs that induce enzymes (e.g. carbamazepine, phenytoin and phenobarbital) reduce efficacy of oral contraceptives; valproate does not. A combined contraceptive pill containing either ethinylestradiol 50 μg or mestranol 50 μg provides greater contraceptive security at the risk of side-effects. An IUCD or barrier methods of contraception are an alternative.

Breast-feeding. Mothers taking AEDs (Table 21.34) need not in general be discouraged from breast-feeding, though manufacturers are often hesitant in assuring that there is no risk to the baby.

Drug withdrawal

Epilepsy, when controlled, may remain in remission. AED withdrawal is sometimes possible, and this question is often raised. Withdrawal should not usually be considered until all fits have been absent for at least 2–3 years. Less than 50% of attempts to withdraw AEDs are successful. Recurrence of seizures can cause considerable difficulty, e.g. when a driving licence has been regained following successful AED therapy. The UK Driving Licensing Authority (DVLA Swansea) recommends that patients do not drive while AEDs are being reduced and for 6 months after stopping them. Careful discussion and full explanation are necessary.

Neurosurgical treatment

Several surgical approaches are used in epilepsy. Amputation of the non-dominant anterior temporal lobe can be performed in a patient with uncontrolled seizures and hippocampal sclerosis defined by imaging and confirmed by EEG. In these highly selected cases (under 1% of epilepsy patients) in specialist centres this surgical treatment is very effective, with cure rates (complete seizure cessation) over 50%. Section of the corpus callosum and hemispherectomy are also used.

Social consequences

The majority of patients with epilepsy are managed in primary care or as outpatients, and suffer infrequent seizures. In a small minority who have exceedingly frequent seizures, treatment in hospital or even residential care is necessary.

There remains a considerable stigma attached to the word epilepsy; this must be taken into account when the nature of attacks is uncertain. Employers are reluctant to accept people labelled with epilepsy.

Both adults and children with epilepsy should be encouraged to lead lives as unrestricted as reasonably possible, though with simple, sensible provisos such as avoiding swimming alone and dangerous sports such as rock-climbing or solo canoeing. Bicycling is felt to be reasonable by some but not by others. Advice at home includes leaving bathroom and lavatory doors unlocked. Support groups and information services are valuable.

Driving and epilepsy

Patients with epilepsy are not legally permitted to drive motor vehicles unless they have been seizure-free for 12 months in the UK and European Union, or as little as 3 months in parts of the USA. Rules for commercial vehicles are more strict. If attacks occur solely in sleep, driving may be permitted by the licensing anthorities.

Stringent regulations also exist for potential aircraft pilots, sea captains, divers and other defined activities. The correct diagnosis of an attack at any age is therefore of particular social and legal importance.

It is an essential medical requirement to inform patients of the law. The patient should then inform the licensing authorities.

FURTHER READING

Duncan, S. Epilepsy surgery. *Clinical Medicine* 2007; **7**: 137–142.

French JA, Pedley TA. Initial management of epilepsy. *New England Journal of Medicine* 2008; **359**: 166–176.

National Society for Epilepsy, UK patient support group. Chesham Lane, Chalfont St Peter, Bucks SL9 0RJ (http://www.epilepsynse.org.uk).

Shorvon SD. Epilepsy and related disorders. In: Clarke C, Howard R, Rossor M, Shorvon S. (eds) *Neurology: A Queen Square Textbook*. Oxford: Blackwell, 2009.

UK Drivers Medical Group, DVLA. *At a Glance Guide to the Current Medical Standards of Fitness to Drive*. Swansea: DVLA, 2007.

OTHER FORMS OF DISTURBED CONSCIOUSNESS (Table 21.37)

Transient disturbance of awareness, loss of consciousness and falls are common clinical problems. It is usually possible to distinguish between a fit (i.e. a seizure), a faint (i.e. syncope) and other types of attack from descriptions and an eye-witness account. *Falls* must be distinguished from episodes of disturbed consciousness. The precise cause of falls often remains ill-defined.

Syncope, postural hypotension and drop attacks

The simple faint that over half the population experiences at some time (particularly in childhood, in youth or in pregnancy) is due to sudden reflex bradycardia with vasodilatation of both peripheral and splanchnic vasculature. This *simple syncope* (*neurocardiogenic syncope*) is a common response to prolonged standing, fear, venesection or pain. Syncope almost never occurs in the recumbent posture. The subject falls to the ground and is unconscious for less than 2 minutes. Recovery is rapid. Jerking movements can occur. Incontinence of urine can occur.

Cardiac syncope (p. 689) is potentially serious and often treatable.

Syncope occurs after *micturition* in men, particularly at night, and in either sex when the venous return to the heart is obstructed by breath-holding and severe coughing. *Effort syncope* (on exertion) is of cardiac origin.

Postural hypotension (p. 689) can cause syncope and occurs in the elderly, in autonomic neuropathy, with phenothiazines, levodopa or tricyclic antidepressants. *Carotid sinus problems* (p. 689) can cause syncope. Syncope can occur with severe anaemia.

Table 21.37 Causes of sudden attacks, 'funny turns'

Epilepsy
Syncope
Simple faints
Cardiac dysrhythmias
Cough syncope
Effort syncope
Micturition syncope
Carotid sinus syncope
Autonomic failure
Basilar migraine
Non-epileptic attacks (pseudoseizures)
Panic attacks
Hyperventilation
Breath-holding
Choking attacks, apnoeic episodes
Transient ischaemic attacks
Drop attacks
Hydrocephalic attacks
Tonic attacks (MS)
Hypoglycaemia
Hypocalcaemia
Severe vertigo
Cataplexy, narcolepsy, sleep paralysis
Night terrors in children
Drug reactions (e.g. oculogyric crises)
Paroxysmal dyskinesias
Carcinoid syndrome, scombroid poisoning
Phaeochromocytoma

Drop attacks are instant, unexpected episodes of lower limb weakness with falling, largely in women over 60 years. Awareness is preserved. They are due to sudden change in lower limb tone, presumably of brainstem origin, rather than thromboembolism. Sudden attacks of leg weakness also occur in hydrocephalus.

Transient cerebral (posterior circulation) ischaemia rarely leads to loss of consciousness; patients sometimes faint during a severe basilar migraine.

Syncope: investigation and management

Syncope and related conditions can usually be distinguished from epilepsy on clinical grounds. Witnessed accounts are invaluable: persistent jerking movements and post-episode confusion with amnesia are suggestive of a fit, and unusual in a faint. Cardiac monitoring is used to detect dysrhythmia. Tilt testing (p. 697) is sometimes diagnostic, but has low sensitivity.

The immediate management of syncope, or impending syncope, is to lay the patient down, lift the legs and record the pulse. In rare circumstances where brain blood flow cannot be restored (e.g. propped upright in a dentist's chair), brain infarction can follow syncope.

Other conditions

Non-epileptic attacks (pseudoseizures) regularly cause difficulty in diagnosis. Attacks are often labelled grand mal fits. Usually there are bizarre limb movements, but there can be extreme difficulty in separating these attacks from seizures. EEG videotelemetry is valuable. Apparent status epilepticus

can occur. The serum prolactin level is of some value: this rises during a grand mal seizure but not during a pseudoseizure (or a partial seizure).

Panic attacks are usually associated with autonomic disturbances such as tachycardia, sweating and piloerection. Consciousness is usually preserved and attacks easily recognized.

Hyperventilation is common (see Box 22.12, p. 1208) and overbreathing causes alkalosis. This leads to dizziness, anxiety and sometimes circumoral/peripheral tingling and tetany, e.g. carpopedal spasm (p. 1022). Occasionally there is loss of consciousness.

Breath-holding attacks occur in children.

Hypoglycaemia (p. 1043) causes attacks of loss of consciousness, sometimes with a convulsion. There is often warning, with hunger, malaise, shaking and sweating. Prompt recovery occurs with i.v. (or oral) glucose, or household sugar. Prolonged hypoglycaemia causes widespread cerebral damage. Hypoglycaemic attacks unrelated to diabetes are rare (p. 1059). Feeling faint after fasting does not indicate anything serious.

Hypocalcaemia (see also p. 1021) may be accompanied by a grand mal fit as seizure threshold is lowered.

Vertigo. When acute, vertigo can be sufficiently severe as to cause prostration: a few seconds' unconsciousness sometimes follows.

Choking causes intense coughing and laryngeal spasm when obstruction is partial. When a large food bolus (e.g. meat) blocks the larynx, the person becomes blue and silent. Death may follow if the blockage is not relieved promptly, e.g. by the Heimlich manoeuvre (p. 834) or gravity. Apnoeic episodes (p. 842).

Drug reactions. Acute dystonic reactions (oculogyric crises) are sometimes mistaken for epilepsy. Consciousness is preserved.

Carcinoid syndrome, phaeochromocytoma and scombroid poisoning (p. 284, p. 1025 and p. 950). Flushing and palpitation is sometimes mistaken for anxiety, allergy or possibly a partial seizure.

Paroxysmal dyskinesias. These are rare causes of attacks of abnormal limb movement.

SLEEP AND ITS DISTURBANCES

Sleep is required to preserve recent memory, refresh emotional equilibrium, and avoid neurotransmitter depletion. Poorly understood pathways – e.g. hypothalamic hypocretin (orexin) neuropeptides – between cortex and reticular formation are involved in sleep–waking cycles.

In *insomnia*, sleep is fitful (p. 1197). Less time than usual is spent in REM sleep. In old age, sleep requirement falls, sometimes to 4 hours a night. Insomnia is rarely a feature of neurological disease.

Seizures may occur predominantly or solely during sleep. Sleepwalking, jerking episodes and movements in sleep are events seen in the normal population and are not suggestive of brain pathology.

Narcolepsy and cataplexy

Narcoleptic attacks are periods of irresistible sleep, i.e. excessive daytime drowsiness, in inappropriate circumstances. Episodes occur when there is little distraction, after meals, while travelling in a vehicle, or sometimes without obvious cause (see Table 20.5, p. 1084). *Narcolepsy* is strongly associated with HLA-DR2 and HLA-DQBl*0602 antigens. Narcolepsy patients positive for these antigens may have subnormal CSF hypocretin 1 (orexin) levels, possibly on an autoimmune basis.

Cataplexy is sudden loss of lower limb tone – falling with intact awareness. Attacks are often set off by sudden surprise or emotion.

Narcolepsy and cataplexy sometimes coexist and are accompanied by vivid hypnagogic hallucinations (on falling asleep), hypnopompic hallucinations (on waking), and sleep paralysis – a frightening inability to move.

Treatment

Methylphenidate, dexamfetamine, modafinil, or small doses of tricyclic antidepressants, particularly clomipramine, are used, rarely with great success. Siberian ginseng (*Elenthrococcus serticosus*) and St John's Wort (*Hypericum perforatum*) are also used.

Central sleep apnoea (p. 842)

This rarity is due to impairment of central ventilatory control due to brainstem pathology. Apnoea occurs typically at the onset of sleep. It is seen in non-obese people without the typical history of snoring seen in the more usual obstructive sleep apnoea (p. 842).

Familial fatal insomnia

This is a very rare, prion-related disease (p. 122).

FURTHER READING

Dauvilliers Y, Arnulf I, Mignot E. Narcolepsy with cataplexy. *Lancet* 2007; **369**: 499–511.

Zeman A, Reading P. The science of sleep. *Clinical Medicine* 2005; **5**: 97–101.

MOVEMENT DISORDERS

Disorders of movement divide broadly into two:

- akinetic–rigid syndromes, i.e. slowed movement with increased tone
- dyskinesias – added, uncontrollable movements.

Both akinesia/rigidity and dyskinesia may coexist, leading to complex classifications. Table 21.38 outlines the principal clinical varieties: idiopathic Parkinson's disease and essential tremor are much the commonest.

AKINETIC–RIGID SYNDROMES

Idiopathic Parkinson's disease

In 1817, James Parkinson, a physician in Hoxton, London, published *The Shaking Palsy*, describing this common worldwide condition that has a prevalence of 150/100 000 increasing sharply in those over 70 years. Parkinson's disease is clinically and pathologically distinct from other parkinsonian syndromes.

There are few real clues to the cause of idiopathic Parkinson's disease (PD). The relatively uniform world-wide

Table 21.38 Movement disorders

Akinetic–rigid syndromes
Idiopathic Parkinson's disease
Drug-induced parkinsonism (e.g. phenothiazines)
MPTP-induced parkinsonism
Postencephalitic parkinsonism
Parkinsonism-plus
Wilson's disease
Childhood akinetic–rigid syndromes

Dyskinesias
Essential tremor
Chorea
Hemiballismus
Myoclonus
Tics
Dystonias
Paroxysmal dyskinesias

prevalence suggests that an environmental agent is not responsible. Some factors possibly involved are:

Nicotine. Several epidemiological studies confirm the curious fact that PD is less prevalent in tobacco smokers than in lifelong abstainers.

MPTP. Minute doses of a pyridine compound, methylphenyltetrahydropyridine (MPTP), a byproduct of illicit stimulant drug production, cause severe parkinsonism. Any link with idiopathic PD is tenuous. Suggestions that environmental MPTP-like herbicides are implicated are unsubstantiated.

Encephalitis lethargica. Survivors of an epidemic of encephalitis lethargica in the early 20th century sometimes developed severe parkinsonism. It is not thought that idiopathic PD is related to any infective agent.

Genetic factors. Idiopathic PD is not usually familial, but there is clustering of early-onset PD in some families. Mutations in the *parkin* gene on chromosome 6 have been found in families with autosomal recessive cases of PD and some young apparently sporadic cases. *Parkin* mutations probably account for most PD cases with onset below the age of 40. Mutations of the α-synuclein gene and ubiquitin carboxyl-terminal hydrolase L1 (UCHL1), on chromosomes 2p13 and 4p14–16.3 respectively, also account for some of these cases. Relevance to the common sporadic older PD cases remains unclear.

Pathology

In idiopathic PD, the pars compacta of the substantia nigra undergoes progressive neuronal degeneration; eosinophilic inclusion bodies (Lewy bodies) develop. These contain protein filaments of ubiquitin and α-synuclein. Degeneration also occurs in other basal ganglia nuclei. There is loss of dopamine (and melanin) in the striatum. This correlates with cell loss and the degree of akinesia.

Symptoms and signs

The combination of tremor, rigidity and akinesia develops slowly, over months or several years, together with changes in posture. Common initial symptoms are tremor and slowness. Limbs and joints feel stiff; they ache. Fine movements become difficult. Slowness causes difficulty rising from a chair or getting into or out of bed. Writing becomes small (micrographia) and spidery, tending to tail off. Relatives often notice other features – slowness and an impassive face. Idiopathic PD is almost always initially more prominent on one side. The diagnosis is usually evident from the overall appearance.

Tremor

The characteristic 4–7 Hz pill-rolling *tremor at rest* (movements between thumb and forefinger) typically decreases with action. Tremor is often asymptomatic at first.

Rigidity

Stiffness develops throughout movements and is equal in opposing muscle groups, in contrast to the selective increase in limb tone found in spasticity. This lead pipe-like increase in tone (plastic rigidity) is usually more marked on one side and present in the neck and axial muscles.

Plastic rigidity is more easily felt when a joint is moved slowly and gently; tone increases when the opposite arm moves actively. When stiffness occurs with tremor, smooth lead pipe rigidity is broken up into a jerky resistance to passive movement – cogwheeling (cogging).

Akinesia

Poverty/slowing of movement (bradykinesia) is an additional handicap, distinct from rigidity. There is difficulty initiating movement. Rapid fine finger movements, such as piano-playing, become indistinct, slow and tremulous. Facial immobility gives a mask-like semblance of depression. Frequency of spontaneous blinking diminishes, producing a *serpentine stare*.

Postural and gait changes

Stooping is characteristic. Gait becomes hurrying (festinant) and shuffling with poor arm swinging. The posture is sometimes called *simian* to describe the ape-like forward flexion, immobility and lack of animation. Balance deteriorates, but despite this the gait retains a narrow base. Falls, toppling like a falling tree, are common in later stages.

Speech

Pronunciation is initially a monotone and progresses to tremulous slurring dysarthria, the result of akinesia, tremor and rigidity. Eventually, speech may be lost (anarthria).

Cognitive changes

Cognitive decline may occur early in the condition and is rarely absent in advanced disease. Depression is common.

Gastrointestinal and other symptoms

These include constipation, sometimes an early symptom, heartburn, dribbling, dysphagia and weight loss. Urinary difficulties are common, especially in men. Skin is greasy and sweating excessive.

Natural history and other features

PD worsens over the years, beginning as a mild inconvenience but slowly progressing. Remissions are unknown except for rare, remarkable short-lived periods of release. These can occur at times of emotion, fear or excitement, when the sufferer is released for seconds or minutes and able to move quickly.

While bradykinesia and tremor worsen, power remains normal until immobility makes its assessment difficult. Patients often complain bitterly of limb and joint discomfort. There is no sensory loss. The reflexes are brisk; their asymmetry follows the increase in tone. The plantar responses remain flexor. The rate of progression is very variable, with a benign form running over several decades. Usually the course is over 10–15 years, with death resulting from bronchopneumonia with immobility and cognitive impairment.

Diagnosis

There is no laboratory test; diagnosis is made by recognizing physical signs and distinguishing idiopathic PD from other parkinsonian syndromes. Conventional imaging (MR) is normal initially; atrophy develops. Dopamine transporter (DAT) imaging is abnormal but discriminates poorly between PD and other akinetic–rigid syndromes.

Other diffuse and multifocal brain diseases can cause features of *parkinsonism*, i.e. slowing, rigidity and tremor seen in idiopathic PD. Examples are Alzheimer's disease, multi-infarct dementia, sequelae of repeated head injury (p. 1163), and late effects of severe hypoxia or CO poisoning. Slowing also occurs in hypothyroidism, and in depression.

Treatment

While drugs alter little, if at all, the natural history of PD, levodopa and/or dopaminergic agonists produce striking initial improvement. Drugs are avoided until clinically necessary because of delayed unwanted effects. Older treatments such as antimuscarinics, e.g. trihexyphenidyl (benzhexol), helped little and frequently caused confusion; they remain of some value in severe tremor. The mechanism of action of drugs in PD is shown in Figure 21.33.

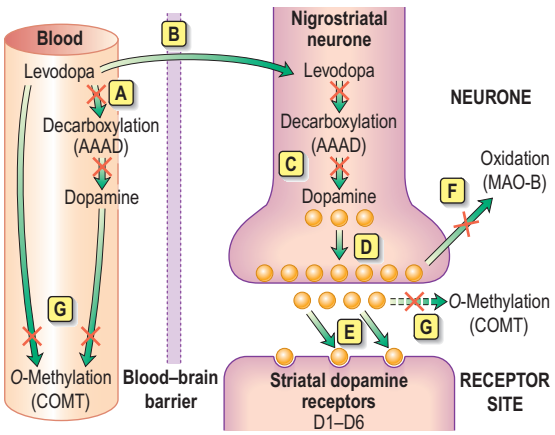


Fig. 21.33 **Drugs in Parkinson's disease.** Levodopa crosses the blood–brain barrier, enters the nigrostriatal neurone and is converted to dopamine. **A.** Carbidopa and benserazide reduce peripheral conversion of levodopa to dopamine, thus reducing side-effects of circulating dopamine. **B.** Dietary amino acids from high-protein meals can inhibit active transport across blood–brain barrier by competing with levodopa. **C.** Levodopa is converted (AAAD) to dopamine. **D.** Amantadine enhances dopamine release. **E.** Dopamine agonists react with dopamine receptors. **F.** Monoamine oxidase B inhibitors block dopamine breakdown. **G.** COMT inhibitors prolong dopamine activity by blocking breakdown. AAAD, aromatic amino acid decarboxylase; COMT, catechol-*O*-methyl transferase.

Levodopa

Levodopa is combined with an aromatic amino acid decarboxylase inhibitor – benserazide (co-beneldopa, as Madopar) or carbidopa (co-careldopa, as Sinemet). The decarboxylase inhibitor reduces peripheral side-effects, principally nausea, of levodopa and its metabolites. Levodopa treatment is commenced (co-beneldopa 62.5 mg or co-careldopa 110 mg, one tablet three times daily) and gradually increased.

The majority of idiopathic PD cases (but not other parkinsonian syndromes) improve initially with levodopa. Response in a severe, previously untreated case is sometimes dramatic.

Unwanted effects of levodopa therapy

Nausea and vomiting are the most common immediate symptoms of excess levodopa. Confusion, formed visual pseudo-hallucinations and chorea also occur. There are difficult issues with long-term therapy (e.g. levodopa-induced involuntary movements). After several years levodopa gradually becomes ineffective, even with increasing doses. As treatment continues, episodes of immobility develop (freezing). Falls are common. Fluctuation in response to levodopa also emerges, its effect turning on and off, to cause freezing alternating with dopa-induced dyskinesias, chorea and dystonic movements. Levodopa's duration of action shrinks, with dyskinesias becoming prominent several hours after a dose (end-of-dose dyskinesia). The patient begins to suffer from a *chronic levodopa-induced movement disorder*.

Levodopa does not alter the natural progression of PD. After 5 years' treatment, around half of PD cases suffer from minor or major unwanted effects of levodopa. The more distressing problems are often largely insoluble. Approaches to treatment of these complications include:

- Shortening the interval between levodopa doses and increasing each dose.
- Monoamine oxidase B inhibitors inhibit catabolism of dopamine in the brain. They are used to smooth out the end-of-dose fluctuations with levodopa.
- Dopaminergic agonists (see below) are added, or replace levodopa.
- Catechol-*O*-methyl transferase (COMT) inhibitors (entacapone and tolcapone) prevent the peripheral breakdown of levodopa, allowing lower doses to be used.
- Apomorphine by subcutaneous pump.

Dopamine receptor agonists

Bromocriptine, pergolide, cabergoline (all ergot derivatives), pramipexole, ropinirole and rotigotine are oral, directly acting dopamine receptor agonists, acting principally on D_1 and D_2 receptors, and also on receptors $D_{3–5}$. Of these, ropinirole is the usual drug of choice, as an alternative or an addition to levodopa therapy. The ergot derivatives are associated with pulmonary, retroperitoneal and pericardial fibrotic reactions and are little used.

Dopamine receptor agonists are in general less effective than levodopa in treating symptoms, but cause fewer late unwanted dyskinesias.

There is variation in clinical practice between the use of levodopa and dopamine receptor agonists. The trend is towards the use of receptor agonists as primary treatment (before levodopa) in cases below 65, and levodopa initially for older cases.

Apomorphine, a potent D_1 and D_2 agonist, given by subcutaneous metered infusion, is one method of attempting to smooth out fluctuations in response to levodopa. Skilled nursing is required to train patients and relatives to deal with the pump. Vomiting is common. Haemolytic anaemia is an unusual side-effect.

Other agents

Antioxidant compounds such as vitamins C and E as possible neuroprotective agents are sometimes suggested; their role is uncertain and evidence of benefit is poor. Amantadine, originally marketed as an antiviral drug, occasionally has a modest effect in PD. Rasagiline and rivastigmine (p. 1168) may help cognitive changes and possibly the movement disorder.

Stereotactic neurosurgery

Stereotactic lesions, usually unilateral in the ventrolateral nucleus of the thalamus or in the globus pallidus (pallidotomy), were used widely before levodopa. Surgery still provides effective, if temporary improvement in tremor and dyskinesia with minor relief of bradykinesia. Thalamic stimulation is used increasingly, sometimes with dramatic improvement.

Tissue transplantation

Despite early promise and suggestive laboratory studies in rats with MPTP-induced parkinsonism, tissue transplantation is little used. Experimental transplantation of fetal or autologous dopamine-containing adrenal medulla and glial cell-line neurotrophic releasing factor (GDNF) into the cerebral ventricles or basal ganglia does not produce any sustained improvement. Stem cell research has produced no promising results in PD to date.

Physiotherapy and physical aids

Skilled therapy can help overcome particular problems. Practical guidance can be given on:

- clothing – avoid zips, fiddly buttons and lace-up shoes
- cutlery – use built-up handles
- chairs – high upright rather than deep, low chairs
- rails – near lavatory and bath
- shoes – easy to wear, with smooth soles
- flooring – vinyl is safer than loose rugs.

Walking aids are often a hindrance early on, but later a frame or a tripod may help. Falls must be avoided.

Neuropsychiatric aspects

Cognitive impairment and depression are common as PD progresses. SSRIs are the drugs of choice for depression. Tricyclic antidepressants (e.g. amitriptyline) have extrapyramidal side-effects. Type A MAO inhibitors (e.g. phenelzine) are contraindicated with levodopa.

All antiparkinsonian drugs can provoke visual hallucinations, especially at night, and exacerbate cognitive impairment.

OTHER AKINETIC–RIGID SYNDROMES

Drug-induced parkinsonism

Phenothiazines and butyrophenones (also reserpine and methyldopa, once used to treat hypertension) induce parkinsonism, usually with little tremor. Tricyclic antidepressants also cause some slowing. These unwanted effects tend not to progress and settle when drugs are stopped.

Neuroleptics and movement disorders

Neuroleptics (i.e. phenothiazines and butyrophenones) also produce other movement disorders.

- **Akathisia**. This is a restless, repetitive and irresistible need to move.
- **Acute dystonic reactions**. Spasmodic torticollis, trismus and oculogyric crises (episodes of sustained upward gaze) develop, dramatically and unpredictably, after single doses of neuroleptics and related drugs used widely as antiemetics and vestibular sedatives (e.g. metoclopramide and prochlorperazine). These acute dystonias respond promptly to i.v. antimuscarinics, e.g. benzatropine 1–2 mg or procyclidine 5–10 mg. The offending drug (and all similar drugs) should be avoided, forever.
- **Chronic tardive dyskinesias**. These mouthing and lip-smacking grimaces occur several years after commencing neuroleptics. They often become temporarily worse when the drug is stopped or the dose reduced. Even if treatment ceases, resolution seldom follows.

Parkinsonism-plus

This term describes disorders in which there is parkinsonism with additional features and specific pathology. *Progressive supranuclear palsy* (Steele–Richardson–Olzewski syndrome) is the commonest. It consists of parkinsonism, axial rigidity, falls, dementia and inability to move the eyes vertically or laterally. Other examples are multiple system atrophies (MSA), such as olivo-ponto-cerebellar degeneration, striatonigral degeneration and primary autonomic failure (Shy–Drager syndrome). Pathologically MSA shows α-synuclein positive glial cytoplasmic inclusions.

These disorders are progressive, although they sometimes respond to levodopa, and usually cause death within a decade.

Post-encephalitic parkinsonism and MPTP

See page 1146.

Akinetic–rigid syndromes in children

Some rare disorders cause akinetic–rigid syndromes usually with onset under 20 years; features may arise during adult life.

Wilson's disease

This rare and treatable disorder of copper metabolism is inherited as an autosomal recessive. Copper deposition occurs in the basal ganglia, the cornea and liver (p. 357), where it can cause cirrhosis. All young patients with an akinetic–rigid syndrome with or without dyskinesia, or with liver cirrhosis should be screened for Wilson's disease. Intellectual impairment develops. Neurological damage is reversible with early treatment. Diagnosis, and treatment with the chelating agent penicillamine, is discussed on page 357.

Athetoid cerebral palsy

Writhing movements, sometimes with progressive dystonia, occur in cerebral palsy following kernicterus. This is now rare following prophylactic eradication of rhesus haemolytic disease with anti-D immunoglobulin.

DYSKINESIAS

Benign essential tremor

This common condition, often inherited as an autosomal dominant trait, causes 5–8 Hz tremor, usually worse in the upper limbs. The head is often tremulous (titubation) and sometimes the trunk. Pathologically there is patchy neuronal loss in the cerebellum and its connections. Tremor develops when the hands adopt a posture, such as holding a glass or a spoon. Essential tremor occurs at any age but occurs most frequently in later life. Tremor is slowly progressive but rarely produces severe disability. Writing is shaky and untidy: micrographia is absent. Anxiety exacerbates the tremor, sometimes dramatically. In essential tremor, shaking occasionally occurs at rest, as in PD, or on action, as in cerebellar disease.

Treatment is often unnecessary, and unsatisfactory. Many patients are reassured to find they do not have PD, with which essential tremor is often confused. Small amounts of alcohol, beta-blockers (propranolol), primidone or the antidepressant mirtazapine may help. Sympathomimetics (e.g. salbutamol) make all tremors worse. Stereotactic thalamotomy and thalamic stimulation are used in severe cases.

Chorea

Chorea means jerky, quasi-purposive, explosive, fidgety movements, flitting around the body. See Table 21.39 for causes.

Huntington's disease

Relentlessly progressive chorea and dementia, usually in middle life but sometimes in childhood, are hallmarks of this inherited disease. Prevalence world-wide is about 5 in 100 000. Inheritance is as an autosomal dominant trait with full penetrance: children of an affected parent have a 50% chance of developing Huntington's disease. Previous family history is often obscure. A mutation occurs in the distal short arm of chromosome 4 (4p16.3) with a variable expansion of a CAG-repeat sequence located in exon 1 of a large gene containing 67 exons. This results in translation of an extended glutamine sequence in *huntingtin*, the protein gene product; its function is unclear. Huntingtin is expressed throughout the body. Most adult-onset HD cases have CAG expansions of 40–55 repeats, while greater expansions (>70 repeats) are seen in childhood-onset HD.

Table 21.39 Causes of chorea

Huntington's disease
Dentato-rubro-pallido-luysian atrophy
Sydenham's chorea
Benign hereditary chorea
Abetalipoproteinaemia with chorea
Chorea associated with:
Drugs – phenytoin, levodopa, alcohol
Thyrotoxicosis, pregnancy and the oral contraceptive pill
Hypoparathyroidism
Systemic lupus erythematosus
Polycythaemia vera
Encephalitis lethargica
Stroke (basal ganglia)
Other, e.g. tumour, trauma, subdural haematoma, following carbon monoxide poisoning, paroxysmal choreoathetosis, Wilson's disease

Pathology

Cerebral atrophy progresses, with marked loss of neurones in the caudate nucleus and putamen.

Changes in neurotransmitters (Fig. 21.9) occur:

- reduced acetylcholine synthesis (due to reduced choline acetyl transferase) and GABA in the striatum
- increased transglutaminase (catalyses aggregates of *huntingtin*) in cortex, cerebellum and corpus striatum
- depleted GABA, angiotensin-converting enzyme and met-enkephalin in substantia nigra
- high somatostatin levels in the corpus striatum.

In contrast to PD, dopamine and tyrosine hydroxylase remain normal.

Management and course

Other causes of chorea should be looked for. Imaging in Huntington's, if practical with the chorea, shows caudate nucleus atrophy. There is steady progression of both dementia and chorea. While nothing arrests this, phenothiazines (e.g. sulpiride) and tetrabenazine reduce chorea. Death usually occurs 10–20 years from the onset.

Mutation analysis is used for presymptomatic testing. Test centres have protocols for counselling families and addressing ethical issues.

Sydenham's chorea (St Vitus's dance)

This transient postinfective chorea occurs largely in children and young adults. Streptococcal infection is one cause: about half the cases follow rheumatic fever, typically within 3 months (p. 136). Sydenham's may recur, or present in adult life, during pregnancy as chorea gravidarum, and with hormonal contraceptives. There is a diffuse mild encephalitis.

The onset is gradual over a few weeks. There is irritability, emotional lability and inattentiveness with fidgetiness, sometimes mainly unilateral. A minority of patients become confused. Rheumatic heart disease is sometimes found. Fever is unusual. Antistreptolysin-O (ASO) titre and ESR are typically normal. Sedation may be needed. Recovery occurs spontaneously within weeks or months. Phenoxymethylpenicillin should continue to the age of 20 to prevent recurrence.

Hemiballismus (see p. 1113 and Fig. 21.9)

Hemiballismus describes violent swinging movements of one side caused usually by infarction or haemorrhage in the contralateral subthalamic nucleus.

Myoclonus

Myoclonus is sudden, involuntary jerking of a single muscle or a group of muscles. This occurs in a wide range of

disorders and is sometimes provoked by sudden stimuli such as loud noise.

Benign essential myoclonus

Nocturnal myoclonus – sudden jerking (often with a feeling of falling) on dropping off to sleep – is common and not pathological. Periodic limb jerks (repetitive dorsiflexion of the great toe or plantar flexion of the foot during Stage I or II sleep) are brief and also occur in restless leg syndrome (p. 626).

Paramyoclonus multiplex describes widespread, random muscle jerking usually occurring in adolescence. Fits do not occur.

Myoclonus in epilepsy

Muscle jerking occurs in many different forms of epilepsy (p. 1140). An antiepileptic drug, e.g. valproate, may be helpful.

Progressive myoclonic epilepsies

These rare conditions include familial and metabolic disorders where myoclonus accompanies progressive encephalopathy. *Lafora body disease* is one example, consisting of myoclonus, epilepsy and dementia, with mucopolysaccharide inclusions in neurones, liver cells and intestinal mucosa.

Static myoclonic encephalopathy

Non-progressive myoclonus sometimes can follow recovery from severe cerebral anoxia.

Non-organic muscle jerking

Sometimes volitional, this can be seen in somatization, non-epileptic attacks and malingering.

Tics

Some idiosyncratic movement of the face, neck or hand is part of our normal motor gestures. Patients or relatives seek advice when movements become frequent or irritating. Simple transient tics (e.g. sniffing or facial grimacing) are common in childhood, but may persist. The borderland between normal and pathological is vague.

Gilles de la Tourette syndrome

This describes multiple tics (motor and speech) with behavioural problems including attention deficit hyperactivity disorder (ADHD) and obsessive-compulsive disorder (OCD). It develops in childhood or adolescence, more commonly in males, and is lifelong. There is sometimes explosive barking and grunting of obscenities and gestures. The cause is believed to be an inherited disorder of synaptic transmission. Haloperidol is sometimes helpful.

Dystonias

Dystonia means movement caused by prolonged muscular contraction – part of the body is thrown into spasm. A brief explanatory classification of these unusual basal ganglia conditions is given in Table 21.40. Their causes are largely unknown.

Table 21.40 A classification of dystonias

Generalized dystonia
Primary torsion dystonia (PTD)
Dopamine-responsive dystonia (DRD)
Drug-induced dystonia (e.g. metoclopramide)
Symptomatic dystonia (e.g. after encephalitis lethargica or in Wilson's disease)
Paroxysmal dystonia (very rare, familial, with marked fluctuation)
Focal dystonia
Spasmodic torticollis
Writer's cramp
Oromandibular dystonia
Blepharospasm
Hemiplegic dystonia, e.g. following stroke
Multiple sclerosis – rare

Primary torsion dystonia (PTD)

Dystonia affecting gait and posture in childhood progresses, spreading to all parts of the body over one to four decades. Cognitive function remains normal. Spontaneous remissions occur occasionally. This rare disease is usually inherited as an autosomal dominant. A PTD gene is on chromosome 9 (9q34), a deletion of three base pairs encoding an ATP-binding protein, torsin A.

Dopamine-responsive dystonia (DRD)

This lower limb dystonia is almost completely abolished by small doses of levodopa. Typically dystonic walking begins in childhood and may resemble a spastic paraparesis. The usual form is autosomal dominant DRD with a point mutation of the GTP cyclohydrolase 1 gene on chromosome 14q 22.3. Patients with dystonic gaits are sometimes given test doses of levodopa.

Spasmodic torticollis

Dystonic spasms gradually develop around the neck, usually in the third to fifth decade. These cause the head to turn (torticollis) or to be drawn backwards (retrocollis) or forwards (antecollis). Minor dystonic movements often also affect the trunk and limbs. A curious feature in some patients is a single trigger area, often on the jaw. A gentle touch with a finger tip at a specific site relieves the spasm temporarily. Torticollis may remit but often persists indefinitely. Similar features can be seen when no organic disease is present, for example people malingering following minor neck trauma.

Writer's cramp

This is a specific inability to perform a previously highly developed repetitive skilled movement, e.g. writing. The movement provokes dystonic posturing. Writer's cramp occurs in those who spend many hours each day writing, and is thus seen less frequently than in years past. Other functions of the hand remain normal. There are no other neurological signs. Prolonged rest sometimes helps but the dystonia can cause substantial disability. Similar dystonias occur in other occupations.

Blepharospasm and oromandibular dystonia

These consist of spasms of forced blinking or involuntary movement of the mouth and tongue (e.g. lip-smacking and protrusion of the tongue and jaw). Speech may be affected.

Treatment

All dystonic movement disorders are particularly difficult to alleviate. Butyrophenones (e.g. haloperidol and sulpiride) and

antimuscarinics (e.g. trihexyphenidyl (benzhexol)) are sometimes helpful. Carefully sited botulinum toxin injections can help, temporarily, blepharospasm, torticollis and writer's cramp. Neurosurgical treatment, principally stereotactic thalamotomy for torticollis, or neurostimulation offers some respite in selected cases.

FURTHER READING

De Lau LML, Breteler MMB. Epidemiology of Parkinson's disease. *Lancet Neurology* 2006; **5**: 525–535.

Lorenz D, Deuschl G. Update on the pathogenesis and treatment of essential tremor. *Current Opinion in Neurology* 2007; **20**: 447–452.

Singer HS. Tourette's syndrome: from biology to behaviour. *Lancet Neurology* 2005; **4**: 149–159.

Tarsy D, Simon DK. Dystonia. *New England Journal of Medicine* 2006; **355**: 818–829.

Walker FO. Huntington's disease. *Lancet* 2007; **369**: 218–228.

MULTIPLE SCLEROSIS (MS)

MS is a chronic inflammatory disorder of the CNS. There are multiple plaques of demyelination within the brain and spinal cord. Plaques are disseminated both in time and place, hence the earlier name 'disseminated sclerosis'.

Prevalence

MS prevalence varies widely, being directly proportional to distance of residence from the equator. At latitudes of 50–65° N (roughly from southern England to Iceland) prevalence is 60–100 per 100 000; at latitudes less than 30° N prevalence is below 10 per 100 000. At the equator MS is a rarity. Dietary factors, e.g. high animal fat consumption, have been suggested for this geographical distribution but there is little evidence for this. In the southern hemisphere this trend is similar – increasing prevalence with distance from the equator. In Europe and North America MS is a common neurological disease of young adults.

Aetiology and pathogenesis

Although the precise mechanism is unknown, there is an inflammatory process in the white matter of the brain and cord mediated by CD4 T cells. In active lesions (plaques) there is an increase in inflammatory cells, active myelin degradation and phagocytosis. Lymphocytes and monocytes gain access to the brain parenchyma from the circulation by adhering to vascular endothelial cells via the glycoprotein $\alpha_4\beta_1$ integrin expressed on their surface. This integrin (very late antigen VLA_4) is also a regulator of immune cell activation.

An initial inflammatory demyelinating event may prime autoreactive cellular and humoral immune responses against myelin. Antibody-mediated demyelination, e.g. against myelin basic protein, probably develops early in MS. Antibodies against myelin oligodendrocyte glycoprotein (MOG), a protein specific to the CNS, have also been found in vitro.

Familial incidence, HLA linkage and migration

First-degree relatives of a patient have an increased chance of developing MS, without a clear-cut pattern of inheritance. The concordance rate is 31% in monozygotic twins.

In Caucasians in northern Europe and the USA, there is a weak association between MS and antigens HLA-A3, B7, D2 and DR2.

Immigrants from low to high prevalence zones (e.g. from the equator to northern Europe) acquire the prevalence of their destination, provided they arrive before the age of 10, irrespective of racial origin.

Infection

Efforts to transmit MS experimentally have been uniformly unsuccessful. However, an abnormal immune response in many MS patients produces increased titres of serum and CSF antibodies to many common viruses, particularly measles. Some epidemic transmissible zoonoses, such as scrapie (demyelinating disease in sheep), have some similarities to MS. In man, human T-cell leukaemia virus 1 (HTLV-1) causes tropical spastic paraparesis, but no other features seen in MS. Chlamydia has been questioned as a cause. No known links exist between MS and any infection.

Pathology (Fig. 21.34)

Plaques of demyelination, initially 2–10 mm in size, are the cardinal features. Plaques are perivenular with a predilection for distinct CNS sites: optic nerves, the periventricular region, the brainstem and its cerebellar connections and the cervical cord (corticospinal tracts and posterior columns).

Acute relapses are caused by focal inflammatory demyelination, which causes conduction block. Remission follows as inflammation subsides and remyelination occurs, helping recovery. When damage is severe, secondary permanent axonal destruction occurs. In the cord, plaques rarely destroy large groups of anterior horn cells, thus focal muscle wasting (e.g. small hand muscles) is unusual. Plaques are not seen in myelin sheaths of peripheral nerves.

Clinical features

The commonest age of onset is between 20 and 45 years; diagnosis before puberty or after 60 is rare. MS is more common in women. No single group of signs or symptoms is diagnostic. The Macdonald criteria (revised in 2005) provide

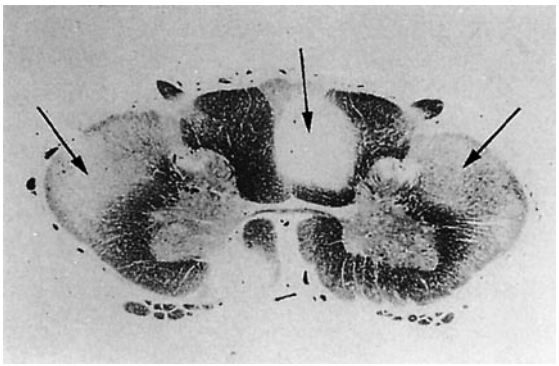


Fig. 21.34 **Multiple sclerosis.** Cross-section of spinal cord showing MS plaques in posterior column and lateral corticospinal tracts. (Courtesy of the late Prof. Ian Macdonald.)

a fairly sensitive method which can be used for the diagnosis and use the clinical presentation along with radiological findings.

MS is often recognizable clinically by different patterns:

- relapsing and remitting MS (80–90%)
- primary progressive MS (10–20%)
- secondary progressive MS – this follows on from relapsing/remitting disease
- occasionally (<10%) MS runs a fulminating course over some months (fulminant MS).

Three characteristic common presentations of relapsing and remitting MS are optic neuropathy, brainstem demyelination and spinal cord lesions, described below.

Optic neuropathy (ON)

Symptoms

Blurring of vision in one eye develops over hours or days, varying from a frosted glass effect to severe unilateral visual loss, but rarely complete blindness. Mild ocular pain is usual. Recovery occurs, typically within 2 months. Bilateral ON can occur.

Signs

The optic disc appearance depends upon the site of the plaque within the optic nerve. When the lesion is in the nerve head there is disc swelling (optic neuritis, p. 1102). If the lesion is several millimetres behind the disc there are no ophthalmoscopic abnormalities – the doctor sees nothing and the patient sees nothing. This is *retrobulbar neuritis* (p. 1102).

Worsening of vision in ON during a fever, hot weather or after exercise is known as Uthoff's phenomenon – central conduction is slowed by an increase in body temperature.

Disc swelling from optic neuritis causes early visual acuity loss, thus distinguishing it from disc swelling from raised intracranial pressure.

A relative afferent pupillary defect (p. 1103) is often found, and persists after recovery.

Late sequelae of optic neuropathy

There is often no residual loss of vision, but small scotomata and defects in colour vision can be found. Following optic neuropathy, disc pallor can develop (optic atrophy), first on the temporal side. Visual evoked responses (VER) remain abnormal (see below and p. 1121).

Brainstem demyelination

Acute MS in the brainstem causes combinations of diplopia, vertigo, facial numbness/weakness, dysarthria or dysphagia. Pyramidal signs in the limbs occur when the corticospinal tracts are involved. A typical picture is sudden diplopia, and vertigo with nystagmus, but without tinnitus or deafness. This lasts for some weeks before recovery. Diplopia is produced by many lesions – a VIth nerve lesion and internuclear ophthalmoplegia (INO, p. 1105) are examples.

Spinal cord lesions

Spastic paraparesis developing over days or weeks (p. 1164) is a typical result of a plaque in the cervical or thoracic cord, causing difficulty in walking and lower limb numbness. Lhermitte's sign may be present (p. 1116). Urinary symptoms are common.

Unusual presentations

Epilepsy and trigeminal neuralgia (p. 1106) occur more commonly in MS patients than in the general population. Tonic spasms (brief spasms of one limb) are other unusual presentations. Organic psychosis is occasionally seen in early MS.

End-stage multiple sclerosis

Late MS causes severe disability with spastic tetraparesis, ataxia, optic atrophy, nystagmus, brainstem signs (e.g. bilateral INO), pseudobulbar palsy and urinary incontinence. Dementia is common. Death follows from uraemia and/or bronchopneumonia.

Differential diagnosis and course

Few other neurological diseases have a similar relapsing and remitting course. Thromboembolism typically causes events with more sudden onset. Other degenerative conditions, such as Friedreich's ataxia, are gradually progressive.

Following an isolated neurological event that is *possible* MS, it is often unclear, even with MR imaging, whether or not MS is the cause. The pattern and imaging appearance of subsequent lesions establishes the diagnosis. Remissions may last for several or more years; their length is unpredictable and the mechanism of relapse and remission unknown. Infection sometimes appears to provoke a relapse. Claims are made for trauma to cause relapses: evidence for this is negligible.

Individual plaques (e.g. in optic nerve, brainstem or cord) must be distinguished from mass, vascular or other inflammatory lesions. Of the latter, CNS sarcoidosis, SLE and Behçet's syndrome may mimic relapsing MS. Adrenoleucodystrophy can cause a progressive paraparesis identical to chronic progressive MS (p. 1122).

Investigations

- *MRI* of brain and cord is the definitive structural investigation. Multiple plaques are seen (Fig. 21.35), principally in the periventricular region, corpus callosum, cerebellar peduncles, juxtacortical posterior fossa, brainstem and cervical cord. Head MRI scans show lesions (elliptical in shape with discrete borders and lack of a mass effect) in 85% of patients with clinical MS. Typical lesions are multifocal (2 mm to 2 cm), with 10 or more lesions. Gadolinium can enhance active lesions. Plaques are rarely visible on CT.
- *CSF examination* is often unnecessary with suggestive MR imaging and a compatible clinical picture. CSF analysis shows oligoclonal IgG bands in 80% of cases but these are not specific. The CSF cell count may be raised (5–60 mononuclear cells/mm^3).
- *Evoked responses*. Delay in visual evoked responses (VER) is seen in optic neuropathy. Some ON attacks are

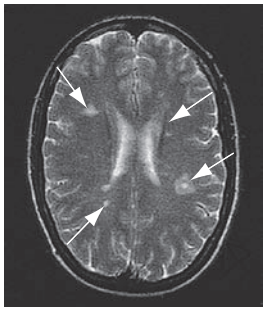


Fig. 21.35 **Multiple sclerosis: MR T2.** Arrows indicate brain plaques.

subclinical: a delayed VER provides evidence of a previous, silent optic nerve lesion. This is helpful for diagnosis if there had been, for example, a previous apparently solitary cord lesion. Brainstem and somatosensory evoked potentials become delayed when these pathways have been damaged.

Peripheral nerve studies are normal and EEGs unhelpful.

- *Blood/urine tests* are unhelpful.

Management and prognosis

Once diagnosed, practical decisions need to be taken about employment, home and plans for the future in the face of a potentially disabling disease for which there is no curative treatment.

The course of MS is unpredictable: a florid MR lesion load at initial presentation is a strong predictor towards serious disability. There is wide variation in severity. Many patients continue to live self-sufficient, productive lives; others become gravely disabled.

Straightforward information and advice, with emphasis on the benign course of many MS cases, is required. The MS Society and other charities have helpful literature.

MS therapies

Many forms of treatment have been marketed, among them cryotherapy, pyrotherapy, radiotherapy, various vaccines, purified TB protein derivative (PPD), transfer factor, electrical stimulation, gluten-free diets, sunflower seed oil, arsenicals, vitamins and hyperbaric oxygen. None has been shown to improve outcome.

- *Acute relapses*. Short courses of steroids, such as i.v. methylprednisolone 1 g/day for 3 days or high-dose oral steroids, are used widely in relapses and do sometimes reduce severity. They do not influence long-term outcome.
- *Preventing relapse and disability*. Beta-interferon (both INF-β1b and 1a) by self-administered injection is used in relapsing and remitting disease. This is defined as at least two attacks of neurological dysfunction over the previous 2 or 3 years followed by a reasonable recovery. IFN-β1b is also now licensed for secondary progressive MS. Interferon certainly reduces relapse rate in some patients and prevents an increase in lesions seen on MRI. Unwanted effects are flu-like symptoms and irritation at injection sites. Cost-benefit analyses are serious issues: beta-interferons are expensive.
- *Immunosuppressants* and antineoplastic drugs (azathioprine, cyclophosphamide, mitoxantrone and others) are also used; trials have sometimes shown slight benefits.
- *Glatiramer acetate*, an immunomodulator, has been shown to reduce relapse frequency in ambulatory patients with relapsing remitting MS – similar to beta-interferon.
- *Natalizumab* is a monoclonal antibody which inhibits migration of leucocytes into the central nervous system by inhibitory α-4 integrins found on the surface of lymphocytes and monocytes. It is useful in severe, relapsing remitting MS that is unresponsive to other treatments. It is associated with a risk of progressive multifocal leucoencephalopathy (PML) and all patients need close surveillance for this and hypersensitivity reactions.

The rehabilitative approach in MS

Much can be done for patients with chronic disabling diseases. Practical advice at work, on walking aids, wheelchairs, car conversions and alterations to houses and gardens is needed, from professionals with experience. Wide-ranging support – for fear, reactive depression and sexual difficulties – is also helpful. Multidisciplinary team liaison between patient, carers, doctors and therapists is essential.

All infections should be treated. Urinary infection frequently exacerbates symptoms. Urinary incontinence may be helped by oxybutinin and/or intermittent self-catheterization.

Physiotherapy is of particular value in reducing pain and discomfort of spasticity, particularly lower limb flexor spasms. Muscle relaxants (e.g. baclofen, benzodiazepines, dantrolene and tizanidine) are sometimes helpful. Injected botulinum toxin is used for severe spasticity. Cannabis is used by many patients for painful spasms. Cannabis extracts and synthetic cannabinoids are also used. Prevention of pressure sores is vital. Amantadine is sometimes used for general fatigue.

FURTHER READING

Association of British Neurologists (ABN). *ABN Guidelines for Treatment of Multiple Sclerosis with β-Interferon and Glatiramer Acetate*. London: ABN, 2007.

Compston A, Coles A. Multiple sclerosis. *Lancet* 2008; **371**: 1502–1517.

Lassmann H, Bruck W, Lucchinetti CF. The immunopathology of multiple sclerosis: an overview. *Brain Pathology* 2007; **17**: 210–218.

Peltonen L. Old suspects found guilty – the first genome profile of multiple sclerosis. *New England Journal of Medicine* 2007; **357**: 927–929.

Polman CH, Reingold SC, Edan G et al. Diagnostic criteria for Multiple Sclerosis: 2005 revisions to the Macdonald criteria. *Annals of Neurology* 2005; **58**: 840–846.

Swanton JK, Rovira A, Tintore M et al. MRI criteria for multiple sclerosis in patients presenting with clinically isolated syndromes: a multicentre retrospective study. *Lancet Neurology* 2007; **6**: 664–665.

NERVOUS SYSTEM INFECTION AND INFLAMMATION

MENINGITIS

Meningitis usually implies serious *infection* of the meninges (Table 21.41). Bacterial meningitis is fatal unless treated. Microorganisms reach the meninges either by direct extension from the ears, nasopharynx, cranial injury or congenital meningeal defect, or by bloodstream spread. Immunocompromised patients (e.g. HIV, cytotoxic drug therapy) are at risk of infection by unusual organisms. Non-infective causes of meningeal inflammation include malignant cells, intrathecal drugs and blood following subarachnoid haemorrhage.

Pathology

In acute bacterial meningitis, the pia–arachnoid is congested with polymorphs. A layer of pus forms. This may organize to form adhesions, causing cranial nerve palsies and hydrocephalus.

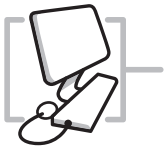
In chronic infection (e.g. TB), the brain is covered in a viscous grey-green exudate with numerous meningeal tubercles. Adhesions are invariable. Cerebral oedema occurs in any bacterial meningitis.

Table 21.41 Infective causes of meningitis in the UK

Bacteria
*Neisseria meningitidis**
*Streptococcus pneumoniae**
Staphylococcus aureus
Streptococcus Group B
Listeria monocytogenes
Gram-negative bacilli, e.g. *E. coli*
Mycobacterium tuberculosis
Treponema pallidum
Viruses
Enteroviruses:
ECHO
Coxsackie
Poliomyelitis
Mumps
Herpes simplex
HIV
Epstein–Barr virus
Fungi
Cryptococcus neoformans
Candida albicans
Coccidioides immitis, Histoplasma capsulatum, Blastomyces dermatitidis (USA)

*These organisms account for 70% of acute bacterial meningitis outside the neonatal period. A wide variety of infective agents are responsible for the remaining 30% of cases. *Haemophilus influenzae* b (Hib) has been eliminated as a cause in many countries by immunization. Malaria often presents with cerebral symptoms and a fever.

www.studentconsult.com
More online



In viral meningitis there is a predominantly lymphocytic inflammatory CSF reaction without pus formation, polymorphs or adhesions; there is little or no cerebral oedema unless encephalitis develops.

Clinical features

The meningitic syndrome

This is a simple triad: headache, neck stiffness and fever. Photophobia and vomiting are often present. In acute bacterial infection there is usually intense malaise, fever, rigors, severe headache, photophobia and vomiting, developing within hours or minutes. The patient is irritable and often prefers to lie still. Neck stiffness and positive Kernig's sign usually appear within hours.

In less severe cases (e.g. many viral meningitides) there are less prominent meningitic signs. However, bacterial infection may also be indolent, with a deceptively mild onset. Clinical pictures of apparently mild meningitis are highly unreliable.

In uncomplicated meningitis, consciousness remains intact, although anyone with high fever may be delirious. Progressive drowsiness, lateralizing signs and cranial nerve lesions indicate complications such as venous sinus thrombosis (p. 1136), severe cerebral oedema, hydrocephalus, or an alternative diagnosis such as cerebral abscess (p. 1159) or encephalitis (p. 1156). Papilloedema may develop.

Specific varieties of meningitis

Clinical clues point to the diagnosis (Table 21.42). If there is access to the subarachnoid space via skull fracture (recent or old) or occult *spina bifida*, bacterial meningitis can be recurrent, and the infecting organism is usually pneumococcus.

Table 21.42 Clinical clues in meningitis

Clinical feature	Possible cause
Petechial rash	Meningococcal infection
Skull fracture Ear disease Congenital CNS lesion	Pneumococcal infection
Immunocompromised patients	HIV opportunistic infection
Rash or pleuritic pain	Enterovirus infection
International travel	Malaria Poliomyelitis
Occupational: (work in drains, canals, polluted water, recreational swimming) Clinical: myalgia, conjunctivitis, jaundice	Leptospirosis

Emergency Box 21.1

Meningococcal meningitis and meningococcaemia: emergency treatment

Suspicion of meningococcal infection is a medical emergency requiring treatment immediately.

Clinical features:

- petechial or non-specific blotchy red rash
- fever, headache, neck stiffness.

All these features may not be present – and meningococcal infection may sometimes begin like any apparently non-serious infection.

Immediate treatment for suspected meningococcal meningitis at first contact before transfer to hospital or investigation:

- benzylpenicillin 1200 mg (adult dose) slow i.v. injection or intramuscularly
- alternative if penicillin allergy – cefotaxime 1 g i.v.

In meningitis, minutes count: delay is unacceptable.
On arrival in hospital:

- routine tests including blood cultures immediately
- watch out for septicaemic shock.

For further management and prophylaxis, see text.

Acute bacterial meningitis

Onset is typically sudden, with rigors and high fever. Meningococcal meningitis is often heralded by a petechial or other rash, sometimes sparse (see Emergency box 21.1). The meningitis may be part of a generalized meningococcal septicaemia (p. 136). Acute septicaemic shock may develop in any bacterial meningitis.

Viral meningitis

This is almost always a benign, self-limiting condition lasting 4–10 days. Headache may follow for some months. There are no serious sequelae, unless an encephalitis is present (p. 1156).

Chronic meningitis (see below)

Differential diagnosis

It may be difficult to distinguish between the sudden headache of subarachnoid haemorrhage, migraine and acute meningitis. Meningitis should be considered seriously in

Table 21.43 Typical CSF changes in viral, pyogenic and TB meningitis

	Normal	Viral	Pyogenic	Tuberculosis
Appearance	Crystal clear	Clear/turbid	Turbid/purulent	Turbid/viscous
Mononuclear cells	$<5/mm^3$	$10–100/mm^3$	$<50/mm^3$	$100–300/mm^3$
Polymorph cells	Nil	Nil*	$200–300/mm^3$	$0–200/mm^3$
Protein	0.2–0.4 g/L	0.4–0.8 g/L	0.5–2.0 g/L	0.5–3.0 g/L
Glucose	⅔–½ blood glucose	> ½ blood glucose	< ⅓ blood glucose	< ½ blood glucose

*Some CSF polymorphs may be seen in the early stages of viral meningitis and encephalitis.

anyone with headache and fever and in any sudden headache. Neck stiffness should be assessed carefully – it may not be obvious. Chronic meningitis sometimes resembles an intracranial mass lesion, with headache, epilepsy and focal signs. Cerebral malaria can mimic bacterial meningitis.

Management (Emergency Box 21.1)

Recognition and immediate treatment of acute bacterial meningitis is vital. Minutes save lives. Bacterial meningitis is lethal. Even with optimal care, mortality is around 15%. The following applies to adult patients; management is similar in chidren.

When meningococcal meningitis is diagnosed clinically by the petechial rash, immediate i.v. antibiotics should be given and blood cultures taken; lumbar puncture is unnecessary. In other causes of meningitis a lumbar puncture is performed if there is no clinical suspicion of a mass lesion (p. 1121). If the latter is suspected an immediate CT scan must be performed because coning of the cerebellar tonsils may follow. Typical CSF changes are shown in Table 21.43. CSF pressure is characteristically elevated. If a presumptive diagnosis of the organism can be made (e.g. pneumococcus is likely with skull fracture or sinus infection), targeted treatment should be started immediately. Immediate antibiotic treatment in acute bacterial meningitis is shown in Table 21.44.

Blood should be taken for cultures, glucose and routine tests. Chest and skull films should be obtained if appropriate.

CSF stains demonstrate organisms (e.g. Gram-positive intracellular diplococci – pneumococcus; Gram-negative cocci – meningococcus). Ziehl–Neelsen stain demonstrates acid-fast bacilli (TB), though TB organisms are rarely numerous. Indian ink stains fungi.

It cannot be emphasized enough that meticulous attention should focus on microbiological studies in suspected CNS infection with close liaison between clinician and microbiologist. Specific techniques (e.g. polymerase chain reaction for meningococci and other bacteria) are invaluable. Syphilitic serology should always be carried out.

The clinical picture and CSF examination should thus yield a presumptive cause for acute meningitis within hours. Antibiotics, however, must be started before the actual organism is identified.

If bacterial meningitis is diagnosed, further discussion with the microbiologist should include antibiotics, drug resistance, recent infections in the locality, barrier nursing and prophylaxis.

In adults with pneumococcal meningitis, there is some evidence to support the use of dexamethasone, given first with the initial antibiotics.

Table 21.44 Antibiotics in acute bacterial meningitis

Organism	Antibiotic	Alternative (e.g. allergy)
Unknown pyogenic	Cefotaxime	Benzylpenicillin and chloramphenicol
Meningococcus	Benzylpenicillin	Cefotaxime
Pneumococcus	Cefotaxime	Penicillin
Haemophilus	Cefotaxime	Chloramphenicol

Intrathecal antibiotics are no longer used.

Local infection (e.g. paranasal sinus) should be treated surgically if necessary. Repair of a depressed skull fracture or meningeal tear may be required.

Prophylaxis

Meningococcal infection should be notified to public health authorities, and advice sought about immunization and prophylaxis of contacts, e.g. with rifampicin or ciprofloxacin. MenC, a meningococcal C conjugate vaccine, is part of childhood UK immunization and often given to case contacts. A combined A and C meningococcal vaccine is sometimes used prior to travel to endemic regions, e.g. Africa, Asia; and a quadrivalent ACWY vaccine for specific events, e.g. the Hajj and Umrah in Mecca. There is no vaccine for Group B.

A polyvalent pneumococcal vaccine is used after recurrent meningitis, e.g. after a CSF leak following skull fracture.

Hib (Haemophilus influenzae) vaccine is given routinely in childhood in the UK and other countries, e.g. Gambia, virtually eliminating a common cause of fatal meningitis.

Chronic meningitis

Tuberculous meningitis (TBM) and cryptococcal meningitis commence typically with vague headache, lassitude, anorexia and vomiting. Acute meningitis can occur but is unusual. Meningitic signs often take some weeks to develop. Drowsiness, focal signs (e.g. diplopia, papilloedema, hemiparesis) and seizures are common. Syphilis, sarcoidosis and Behçet's also cause chronic meningitis. In some cases of chronic meningitis an organism is never identified.

Management of tuberculous meningitis

TBM is a common and serious disease world-wide. Brain imaging, usually with MRI, may show meningeal enhancement, hydrocephalus and tuberculomas (p. 1159), though it may remain normal (see Table 21.43 for CSF changes). In many cases the sparse TB organisms cannot be seen on

Ask the authors
www.studentconsult.com

staining and PCR testing should be performed, although results may be negative. Repeated CSF examination is often necessary and it will be some weeks before cultures are confirmatory. Treatment with antituberculosis drugs (p. 866) – rifampicin, isoniazid and pyrazinamide – must commence on a presumptive basis and continue for at least 9 months. Ethambutol should be avoided because of its eye complications. Adjuvent corticosteroids, e.g. prednisolone 60 mg for 3 weeks, are now recommended (often tapered off). Relapses and complications (e.g. seizures, hydrocephalus) are common in TBM. The mortality remains over 60% even with early treatment.

Malignant meningitis

Malignant cells can cause a subacute or chronic non-infective meningitic process. A meningitic syndrome, cranial nerve lesions, paraparesis and root lesions are seen, often in confusing and fluctuating patterns. The CSF cell count is raised, with high protein and low glucose. Treatment with intrathecal cytotoxic agents is rarely helpful.

Cells in a sterile CSF (pleocytosis)

A raised CSF cell count is present without an evident infecting organism. CSF pleocytosis, i.e. a mixture of lymphocytes and polymorphs, is the usual situation (Table 21.45).

Table 21.45 Causes of sterile CSF pleocytosis

Partially treated bacterial meningitis
Viral meningitis
TB or fungal meningitis
Intracranial abscess
Neoplastic meningitis
Parameningeal foci, e.g. paranasal sinus
Syphilis
Cerebral venous thrombosis
Cerebral malaria
Cerebral infarction
Following subarachnoid haemorrhage
Encephalitis, including HIV
Rarities, e.g. cerebral malaria, sarcoidosis, Behçet's syndrome, Lyme disease, endocarditis, cerebral vasculitis

FURTHER READING

Department of Health. *Immunisation Against Infectious Disease*, 1999. [PL CMO (99) 4, PL CNO (99) 8, PL CPHO (99) 3].

Hall AJ, Wild CP. Management of bacterial meningitis in adults. *British Medical Journal* 2003; **326**: 996–997.

Tunbridge A, Read RC. Management of meningitis. *Clinical Medicine* 2004; **4**: 499–505.

van de Beek D, de Gans J, Tunkel AR et al. Community-acquired bacterial meningitis in adults. *New England Journal of Medicine* 2006; **354**: 44–53.

ENCEPHALITIS

Encephalitis means inflammation of brain parenchyma, usually viral.

Acute viral encephalitis

The viruses cultured from adult UK cases are usually herpes simplex, ECHO, Coxsackie, mumps and Epstein–Barr viruses. Adenovirus, varicella zoster, influenza, measles and other viral encephalitides are less common. Often, the virus is never identified.

Epidemic and endemic viral encephalitides occur world-wide:

- Japanese encephalitis in South East Asia
- Tick-borne flavivirus encephalitis over a large area from Western Europe to Japan (p. 116)
- Ross River fever in Australia
- California encephalitis in the USA
- Omsk haemorrhagic fever in Russia
- West Nile encephalitis in Egypt and Sudan.

Local epidemics occur. For example, encephalitis caused by Nipah virus, a virus hitherto not known to cause human disease, occurred in Malaysian abattoir workers in the 1990s. In New York in the 1990s, West Nile virus caused an epidemic. Chandipura virus caused an outbreak in children in Southern India in 2003. In 2007/8, St Louis encephalitis was epidemic in Florida, and a new epidemic/epizootic Venezuelan equine virus was isolated from encephalitis cases in South America.

Clinical features

Many encephalitides are mild; recovery occurs. In the minority, serious illness develops with high fever, headache, mood change and drowsiness over hours or days. Focal signs, seizures and coma ensue. Death, or brain injury follows; herpes simplex (HSV-1) accounts for many of these in the UK. In South East Asia, Japanese encephalitis (arbovirus) is a more usual cause.

Differential diagnosis

This includes:

- bacterial meningitis with cerebral oedema
- cerebral venous thrombosis
- brain abscess
- endocarditis
- acute disseminated encephalomyelitis
- cerebral malaria
- septicaemia, vasculitis
- toxic confusional states.

Investigations

CT and MR imaging show areas of oedema, often in the temporal lobes. EEG shows slow wave changes; a normal EEG is exceptional. CSF cell count is raised. Viral serology (blood and CSF) is helpful. Brain biopsy is occasionally performed.

Treatment

Suspected HSV encephalitis is treated immediately with i.v. aciclovir (10 mg/kg 3 times a day for 14–21 days), the active form of which inhibits DNA synthesis. Phosphorylation of aciclovir is dependent upon viral thymidine kinase; the drug is specific for herpesvirus infections. Coma is usually followed by a poor outlook, with or without aciclovir. Seizures are treated with anticonvulsants.

Prophylactic immunization against Japanese encephalitis is available for travellers to endemic areas in Asia. Tick-borne

encephalitis vaccine is also available. Both vaccines are given routinely to children in endemic areas.

Acute disseminated encephalomyelitis (ADEM)

ADEM follows many viral infections (e.g. measles, varicella zoster, mumps and rubella) and rarely immunization against rabies, influenza or pertussis. The syndrome is often similar to acute viral encephalitis, with additional focal brainstem and/or spinal cord lesions due to demyelination (see also MS, p. 1152). Viral particles are not present in the lesions, which are well seen on T2 weighted MRI. Prognosis is variable. Mild cases recover completely. When there is coma, mortality remains approximately 25%. Survivors often have permanent brain damage. Treatment is supportive, with steroids and anticonvulsants.

MYELITIS

Myelitis means spinal cord inflammation causing paraparesis or tetraparesis. This occurs with varicella zoster or ADEM. Poliomyelitis is a specific spinal cord anterior horn cell enterovirus infection (p. 109). Transverse myelitis is mentioned on page 1158.

HERPES ZOSTER (SHINGLES)

This is recrudescence of varicella zoster infection within dorsal root ganglia, the original infection having been chickenpox many years previously. Chickenpox and shingles viruses are identical.

Clinical features

For shingles rashes, see page 106.

In the cranial nerves, herpes zoster has a predilection for the Vth and VIIth nerves. Ophthalmic herpes is infection of V_1. This can lead to corneal scarring and secondary panophthalmitis. Geniculate herpes (geniculate ganglion of VIIth nerve) is also called Ramsay Hunt syndrome (p. 1108). Local complications are secondary bacterial infection, very rarely local purpura with necrosis (purpura fulminans), generalized zoster, and postherpetic neuralgia. Myelitis, meningoencephalitis and motor radiculopathy (usually lumbar or brachial) can follow varicella zoster. Treatment is with aciclovir or its pro-drug, famiciclovir or valaciclovir. Bromovinyl deoxyuridine, an oral thymidine analogue, has recently been introduced. It is given as early as possible for one week. It should only be used in immunocompetent patients and is contraindicated for use with all forms of 5-fluorouracil.

Postherpetic neuralgia

Postherpetic neuralgia is pain in a previous shingles zone; this occurs in some 10% of patients (often elderly). Burning, continuous pain responds poorly to analgesics. Treatment is unsatisfactory but there is a trend towards gradual recovery over 2 years. Amitryptyline is commonly used. Methylprednisolone intrathecally is sometimes helpful.

Vaccination. The Centers for Disease Control (CDC) in America have suggested that all adults over 60 years old should be vaccinated against herpes zoster (even those who have had shingles previously) as it reduces the incidence of shingles by about 50%.

NEUROSYPHILIS

Many neurological symptoms occur, sometimes mixed. (See also syphilis, p. 177.)

Asymptomatic neurosyphilis

This means positive CSF serology without signs.

Meningovascular syphilis

This causes:

- subacute meningitis with cranial nerve palsies and papilloedema
- a gumma – a chronic expanding intracranial mass
- paraparesis – a spinal meningovasculitis.

Tabes dorsalis

Demyelination in dorsal roots causes a complex deafferentation syndrome. The elements of tabes are:

- lightning pains (p. 1116)
- ataxia, stamping gait, reflex/sensory loss, wasting
- neuropathic (Charcot) joints
- Argyll Robertson pupils (p. 1104)
- ptosis and optic atrophy.

General paralysis of the insane (GPI)

The grandiose title describes dementia and weakness. GPI dementia is typically similar to Alzheimer's (p. 1167). Progressive cognitive decline, seizures, brisk reflexes, extensor plantar reflexes and tremor develop. Death follows within 3 years. Argyll Robertson pupils are usual. GPI and tabes are rarities in the UK.

Other forms of neurosyphilis

In *congenital* neurosyphilis (acquired in utero), features of combined tabes and GPI develop in childhood – taboparesis.

S*econdary* syphilis can be symptomless or cause a self-limiting subacute meningitis.

Treatment

Benzylpenicillin 1 g daily i.m. for 10 days in primary infection eliminates any risk of neurosyphilis. Allergic (Jarisch–Herxheimer) reactions can occur; steroid cover is usually given with penicillin (p. 179). Established neurological disease is arrested but not reversed by penicillin.

NEUROCYSTICERCOSIS

The adult pork tapeworm, *Taenia solium*, is endemic in Latin America, non-Muslim African countries, India and much of South East Asia. This is a major global health challenge (p. 170, p. 1140). Epilepsy is a feature of neurocysticercosis. Other neurological patterns include brainstem dysfunction, cerebellar ataxia, hydrocephalus, and rarely dementia, arachnoiditis and vasculitis. Most infected people remain symptomless.

Brain CT and MRI are helpful, but sometimes not diagnostic. Serological tests indicate infection but not activity. Biopsy of a lesion may be necessary.

Management is primarily the control of seizures with anticonvulsants. The case for taenicides (e.g. praziquantel, albendazole) is debated, but evidence supports their use. Surgery is sometimes needed for giant brain cysts or hydrocephalus.

Nervous system infection and inflammation
Encephalitis
Myelitis
Herpes zoster (shingles)
Neurosyphilis
Neurocysticercosis

FURTHER READING

Garcia HH, Del Brutto OH. Neurocysticercosis: updated concepts about an old disease. *Lancet Neurology* 2005; **4**: 653–661.

HIV AND NEUROLOGY (p. 188)

HIV-infected individuals frequently present with or develop neurological conditions. Immunosuppression leads to indolent, atypical clinical patterns. HIV patients also have a high incidence of stroke. The pattern of disease is changing where antiretroviral therapy is available.

Brain and meningeal disease in HIV

Acute aseptic meningitis is a primary HIV infection, with spontaneous recovery.

Chronic meningitis occurs with fungi (e.g. *Cryptococcus neoformans* or *Aspergillus*), TB, listeria, coliforms or other organisms. Aggressive treatment is essential.

AIDS-dementia complex (ADC). This diffuse, progressive, HIV-related dementia, sometimes with cerebellar signs, is still seen where antiretroviral therapy is unavailable.

Encephalitis and brain abscess. Toxoplasma, cytomegalovirus, herpes simplex and other organisms cause severe encephalitis. Multiple brain abscesses develop in HIV infection.

CNS lymphoma: this is typically fatal (p. 483).

Progressive multifocal leucoencephalopathy (PML) is due to papovavirus (JC virus, p. 108).

Spinal cord disease

Paraparesis with HIV occurs in several settings:

- acute transverse myelitis – primary HIV myelitis with spontaneous recovery
- myelopathy due to infection, e.g. herpes simplex, zoster, cytomegalovirus or TB
- CNS lymphoma.

Peripheral nerve disease

Many HIV neuropathies occur:

- mononeuropathy (e.g. common peroneal, facial)
- mononeuritis multiplex (p. 1172)
- polyneuropathy (p. 1172)
- autonomic neuropathy (p. 1176)
- neuropathies caused by treatment (Table 4.53).

Management of HIV

This is described on page 191.

Other infections

Many other infections involve the CNS and are discussed in Chapter 4, e.g. rabies, tetanus, botulism, Lyme disease and leprosy.

Transmissible spongiform encephalopathies

Creutzfeldt–Jakob diseases (CJDs)

CJDs are rare, progressive dementias recognized pathologically by spongiform changes in the brain (p. 121) caused by prions. Prions (proteinaceous infectious particle, PrP) are transmissible and resistant to many sterilization processes. All PrP diseases are associated with accumulation of a disease-related protein isoform, PrPSc, derived from a normal cellular precursor, PrPC.

Sporadic CJD cases develop progressive dementia, usually over the age of 50 years.

Iatrogenic CJD is transmitted from surgical specimens, autopsy, transplant material (e.g. corneal grafts) and the old pituitary derived growth hormone from cadavers. Iatrogenic CJD has a long incubation period, up to 5 years. Death is usual within approximately 6 months of clinical onset in both sporadic and iatrogenic forms. CJD pathology is very similar to that of bovine spongiform encephalopathy (BSE, mad cow disease), recognized first in the UK in the 1980s and in 2003 in the USA.

Incidence. *Sporadic CJD* occurs world-wide with an annual incidence of less than 1 per million; some 30–55 cases occur annually in the UK. Numbers have not been rising. *Iatrogenic* CJD is rarer: 0–6 UK cases annually. *Familial CJD* cases (0–4 annually in the UK) are associated with *PrP* gene mutations. Another rare form, *Gerstmann–Straüssler–Scheinker syndrome*, is an inherited autosomal recessive condition typified by chronic progressive ataxia and terminal dementia, with a duration of 2–10 years. *Familial fatal insomnia* is a rare familial prion disease presenting with prolonged progressive insomnia.

Variant CJD (vCJD). This condition was first seen in the UK in 1995. The incidence is not increasing. vCJD patients are younger than sporadic cases with a mean age of 29. Early symptoms are neuropsychiatric, followed by ataxia and dementia with myoclonus or chorea. The diagnosis can be confirmed by tonsillar biopsy and CSF gel electrophoresis. vCJD has a longer course than sporadic CJD – up to several years. vCJD and BSE are caused by the same prion strain, giving rise to speculation that transmission from animal to human food chain occurred, i.e. infection from BSE-infected cattle to humans (p. 122). Transmission via blood transfusion may occur.

Kuru. This prion dementia and cerebellar ataxia of New Guinea transmitted by cannibalism is described on page 122 and is now virtually extinct.

OTHER INFLAMMATORY CONDITIONS

Subacute sclerosing panencephalitis (SSPE)

Persistence of measles antigen in the CNS is associated with this rare late sequel of measles. Progressive mental deterioration, fits, myoclonus and pyramidal signs develop, typically in a child. Diagnosis is made by high measles antibody titre in blood and CSF. Measles immunization protects against SSPE, which has now been almost eliminated in the UK.

Progressive rubella encephalitis

Some 10 years after primary rubella infection, this causes progressive mental impairment, fits, optic atrophy, cerebellar and pyramidal signs. Antibody to rubella viral antigen is produced locally within the CNS. It is even rarer than SSPE and is not a complication of rubella immunization.

FURTHER READING

Bird S. Recipients of blood or blood products 'at vCJD risk'. *British Medical Journal* 2004; **328:** 118–119.

Collins SJ. Transmissible spongiform encephalitis. *Lancet* 2004; **363**: 51–61.

FURTHER READING

Manji H, Miller R. The neurology of HIV infection. *Journal of Neurology, Neurosurgery and Psychiatry* 2004; **75**: i29.

Reye's syndrome
This severe encephalitic illness of children (p. 364) is accompanied by fatty infiltration of liver with hypoglycaemia.

Mollaret's meningitis
This is recurrent self-limiting episodes of aseptic meningitis (i.e. no bacterial cause is found) over many years. Viral (possibly herpes simplex) infection is postulated.

Vogt–Koyanagi–Harada syndrome
This obscure recurrent inflammation of cells of neural crest origin causes uveitis, meningoencephalitis, vitiligo, deafness and alopecia.

Progressive multi-focal leucoencephalopathy (PML)
PML, seen in HIV and in immunosuppression, is due to JC virus and is described on page 108.

Whipple's disease
CNS Whipple's disease is characterized by myoclonus, dementia and supranuclear ophthalmoplegia (p. 281).

Neurosarcoidosis
Neurosarcoid with or without systemic sarcoid causes chronic meningoencephalitis, cord lesions, cranial nerve palsies, particularly bilateral VIIth nerve lesions, polyneuropathy and myopathy (p. 869).

Behçet's syndrome (see also p. 552)
Behçet's principal features are recurrent oral and/or genital ulceration, inflammatory ocular disease (uveitis, p. 1092) and neurological syndromes. Brainstem and cord lesions, aseptic meningitis (p. 1156), encephalitis and cerebral venous thrombosis occur.

Granulomatous (primary cerebral) angiitis
In this rare condition there is necrotizing inflammation in the brain and meningeal vessels. Stroke, seizures and confusional states develop. There is some response to steroid and immunosuppressive therapy.

BRAIN AND SPINAL ABSCESSES

Brain abscess (see Fig. 21.36)

Focal bacterial infection behaves as any expanding mass. Typical bacteria found are *Streptococcus anginosus* and *Bacteroides* species (paranasal sinuses and teeth) and staphylococci (penetrating trauma). Mixed infections are common.

Fig. 21.36 **Pyogenic cerebral abscess: CT.** There is a ring enhancing lesion with adjacent oedema.

Multiple abscesses develop, particularly in HIV. Fungi also cause brain abscesses. A parameningeal infective focus (e.g. ear, nose, paranasal sinus, skull fracture) or a distant source of infection (e.g. lung, heart, abdomen) may be present. Frequently no underlying cause is found. An abscess is more than 10 times rarer than a brain tumour in the UK.

Clinical features
Headache, focal signs (e.g. hemiparesis, aphasia, hemianopia), epilepsy and raised intracranial pressure develop. Fever, leucocytosis and raised ESR are usual though not invariable. Cerebral abscesses may be indolent, developing over weeks. Cerebellar abscesses tend to develop rapidly over days or hours, producing hydrocephalus.

Management
Urgent imaging is essential. Contrast enhanced CT shows a thin high-attenuation rim, virtually diagnostic of an abscess. The search for a focus of infection should include a detailed examination of the skull, ears, paranasal sinuses and teeth, and distant sites such as heart and abdomen. Lumbar puncture is dangerous and usually unhelpful. Aspiration with stereotactic guidance allows the infective organism to be identified.

Treatment requires liaison between neurosurgeon and microbiologist. Streptococcal and anaerobic infections are treated with cefuroxime 1.5 g i.v. plus metronidazole 500 mg i.v. 8 hourly. For staphylococcal infections, flucloxacillin 2–3 g i.v. 6 hourly with cefuroxime is given. Surgical decompression may be necessary if parenteral antibiotics are unsuccessful. Despite treatment, mortality remains high at approximately 25%. Epilepsy is common in survivors.

Brain tuberculoma

TB causes chronic caseating intracranial granulomas – tuberculomas. These are the commonest intracranial masses in countries where TB is common, e.g. India. Brain tuberculomas either present as mass lesions de novo or develop during tuberculous meningitis; they are also found as symptomless intracranial calcification on imaging. Spinal cord tuberculomas also occur. Treatment is described on page 866.

Subdural empyema and intracranial epidural abscess

Intracranial subdural empyema is a collection of subdural pus, usually secondary to local skull or middle ear infection. Features are similar to those of a cerebral abscess. Imaging is diagnostic.

In *intracranial epidural abscess* a layer of pus, 1–3 mm thick, tracks along the epidural space causing sequential cranial nerve palsies, typically without raised intracranial pressure. There is usually local infection, e.g. in the middle ear. MRI outlines the abscess; CT is typically normal. Drainage is required, with antibiotics.

Spinal epidural abscess

Staphylococcus aureus is the usual organism, reaching the spine via the bloodstream, e.g. from a boil. Fever and back

FURTHER READING

Koralnik IJ, Schellingerhout D, Frosch MP. A 66-year-old man with progressive neurologic deficits. *New England Journal of Medicine* 2004; **350**: 1882–1893.

pain are followed by paraparesis and/or root lesions. Emergency imaging and antibiotics are essential and surgical decompression is often necessary.

FURTHER READING

Friedlander RM, Gonzalez RG, Afridi NA et al. A 58-year-old woman with left-sided weakness and a right frontal brain mass. *New England Journal of Medicine* 2003; **348**: 2125–2132.

BRAIN TUMOURS

Primary intracranial tumours account for some 10% of neoplasms. The most common tumours are outlined in Table 21.46. *Metastases* are the commonest intracranial tumours (Fig. 21.37). Symptomless meningiomas (benign) are also found on imaging or at autopsy.

Gliomas (Fig. 21.38)

These malignant tumours of neuroepithelial origin are usually seen within the hemispheres, but occasionally in the cerebellum, brainstem or cord. Their cause is unknown. Glioma is occasionally associated with neurofibromatosis. They tend to spread by direct extension, virtually never metastasizing outside the CNS.

- **Astrocytomas** are gliomas that arise from astrocytes. They are classified histologically into grades I–IV. Grade I astrocytomas grow slowly over many years, while grade IV tumours (*glioblastoma multiforme*) cause death within several months. Cystic astrocytomas of childhood are relatively benign, and usually cerebellar.
- **Oligodendrogliomas** arise from oligodendrocytes. They grow slowly, usually over several decades. Calcification is common.

Table 21.46 Common brain tumours

Tumour	Approximate frequency
Metastases	**50%**
Bronchus	
Breast	
Stomach	
Prostate	
Thyroid	
Kidney	
Primary malignant tumours of neuroepithelial tissues	**35%**
Astrocytoma	
Oligodendroglioma	
Mixed (oligoastrocytomas) gliomas	
Ependymoma	
Primary cerebral lymphoma	
Medulloblastoma	
Benign	**15%**
Meningioma	
Neurofibroma	

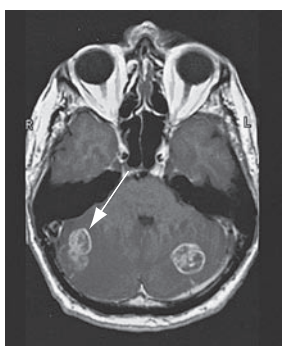


Fig. 21.37 **Bilateral cerebellar metastases: MR T1.**

Meningiomas

These benign tumours (Figs 21.39–21.41) arise from the arachnoid membrane and may grow to a large size, usually over years. Those close to the skull erode bone. They often occur along the intracranial venous sinuses, which they may invade. They are unusual below the tentorium. Common sites are the parasagittal region, sphenoidal ridge, subfrontal region, pituitary fossa and skull base.

Neurofibromas (Schwannomas)

These solid benign tumours arise from Schwann cells and occur principally in the cerebellopontine angle, where they arise from the VIIIth nerve sheath (acoustic neuroma, p. 1109).

Other neoplasms

Other less common neoplasms include cerebellar haemangioblastoma, ependymoma of the IVth ventricle, colloid cyst of the IIIrd ventricle, pinealoma, chordoma of the skull base, glomus tumour of the jugular bulb, medulloblastoma (a cerebellar childhood tumour), craniopharyngioma (p. 972) and primary CNS lymphoma (p. 483). For pituitary tumours, see page 972.

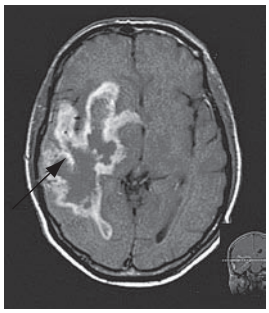


Fig. 21.38 **Cerebral glioblastoma multiforme: MR T1.**

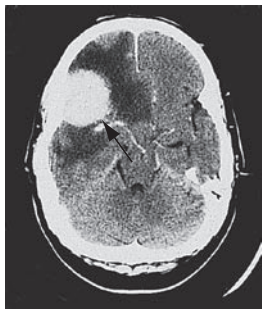


Fig. 21.39 **Frontal meningioma: CT.**

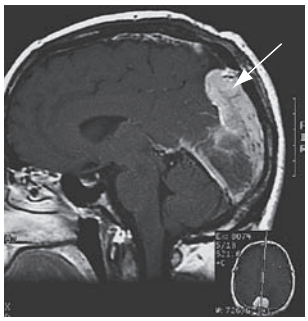


Fig. 21.40 **Falx meningioma (occipital): MR T1.**

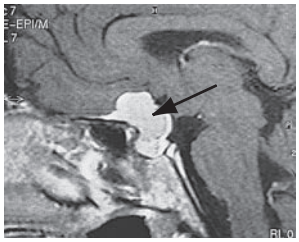


Fig. 21.41 **Meningioma within sella and suprasellar extension: MR T1.**

Clinical features

Mass lesions within the brain produce symptoms and signs by three mechanisms:

- by direct effect – brain is destroyed and local function impaired
- by secondary effects of raised intracranial pressure and shift of intracranial contents (e.g. papilloedema, vomiting, headache)
- by provoking generalized and/or partial seizures.

Although neoplasms, either secondary or primary, are the commonest mass lesions in the UK, cerebral abscess, tuberculoma, neurocysticercosis, subdural and intracranial haematoma can also produce features that are clinically similar.

Direct effects of mass lesions

The hallmark of a direct effect of a mass is local progressive deterioration of function. Tumours can occur anywhere within the brain. Three examples are given:

- *A left frontal meningioma* caused a frontal lobe syndrome over several years with vague disturbance of personality, apathy and impaired intellect. Expressive aphasia developed, followed by progressive right hemiparesis as the corticospinal pathways became involved. As the mass enlarged further, pressure headaches and papilloedema developed.
- *A right parietal lobe glioma* caused a left homonymous field defect (optic radiation). Cortical sensory loss in the left limbs and left hemiparesis followed over 3 months. Partial seizures (episodes of tingling of the left limbs) developed.
- *A left VIIIth nerve sheath neurofibroma (an acoustic neuroma, Schwannoma*) in the cerebellopontine angle caused, over 3 years, progressive perceptive deafness (VIII), vertigo (VIII), left facial numbness (V) and weakness (VII), followed by cerebellar ataxia on the same side. Papilloedema was a late sequel.

With a hemisphere tumour, epilepsy and the direct effects commonly draw attention to the problem. The rate of progression varies greatly, from a few days or weeks in a highly malignant glioma, to several years with a slowly enlarging mass such as a meningioma. Cerebral oedema surrounds mass lesions: its effect is difficult to distinguish from that of the tumour itself.

Raised intracranial pressure

Raised intracranial pressure causing headache, vomiting and papilloedema is a relatively unusual presentation of a mass lesion in the brain. These symptoms often imply hydrocephalus – obstruction to CSF pathways. Typically this is produced early by posterior fossa masses that obstruct the aqueduct and IVth ventricle but only later with lesions above the tentorium. Shift of the intracranial contents produces features that coexist with the direct effects of any expanding mass:

- *Distortion of the upper brainstem*, as midline structures are displaced either caudally or laterally by a hemisphere mass (Fig. 21.38). This causes impairment of consciousness.
- *Compression of the medulla*, by herniation of the cerebellar tonsils caudally through the foramen magnum – an example of coning. This causes impaired consciousness, respiratory depression, bradycardia, decerebrate posturing and death.
- *False localizing signs* – false only because they do not point *directly* to the site of the mass.

Three examples of false localizing signs are:

- *A VIth nerve lesion*, first on the side of a mass and later bilaterally as the VIth nerve is compressed during its long intracranial course.
- *A IIIrd nerve lesion* develops as the temporal lobe uncus herniates caudally, compressing III against the petroclinoid ligament. The first sign is ipsilateral pupil dilatation as parasympathetic fibres are compressed.
- *Hemiparesis* on the same side as a hemisphere tumour, i.e. the side you might not expect, from compression of the contralateral cerebral peduncle on the edge of the tentorium.

These false localizing signs, though unusual, indicate that shift of the brain has occurred.

Seizures

Partial seizures, simple or complex, that may evolve into generalized tonic–clonic seizures, are characteristic of many hemisphere masses, whether malignant or benign. The pattern of partial seizure is of localizing value (p. 1140).

Investigations

Imaging

Both CT and MRI are useful in detecting brain tumours; MRI is much more accurate for posterior fossa lesions. Benign and malignant tumours, brain abscess, TB, neurocysticercosis and infarction have characteristic, but not entirely reliable, appearances and refined imaging techniques and biopsy are often necessary. MR angiography is used occasionally to define blood supply and MR spectroscopy to study patterns typical of certain gliomas. PET is sometimes helpful to locate an occult primary tumour with metastases.

Plain skull X-rays have no value in brain tumour diagnosis.

Routine tests

Since metastases are common, routine tests, e.g. chest X-ray, should be performed.

Lumbar puncture

This is contraindicated when there is any possibility of a mass lesion as withdrawing CSF may provoke immediate herniation of the cerebellar tonsils. CSF examination is rarely helpful and has been superseded by imaging.

Biopsy and tumour removal

Stereotactic biopsy via a skull burr-hole is carried out to ascertain the histology of most suspected malignancies. With a benign tumour, e.g. a symptomatic, accessible meningioma, craniotomy is performed.

Management

Cerebral oedema surrounding a tumour is reduced rapidly by steroids: i.v. or oral dexamethasone. Mannitol, an osmotic diuretic, is also used intravenously to reduce oedema. Epilepsy is treated with antiepileptic drugs.

Whilst complete surgical removal of a tumour is an objective, it is often not possible, nor is surgery always necessary. Follow-up with serial imaging is sometimes preferable. At exploration, some benign tumours can be entirely removed

(e.g. acoustic neuromas, some parasagittal meningiomas). With a brain malignancy it is not possible to remove an infiltrating mass. Biopsy and debulking are performed.

Within the posterior fossa, tumour removal is often necessary because of raised pressure and danger of coning. Overall mortality for posterior fossa exploration remains around 10%.

For gliomas and metastases, radiotherapy is usually given and improves survival, if only slightly. An isolated posterior fossa metastasis can sometimes be excised successfully. Chemotherapy has little real value in the majority of primary or secondary brain tumours. Vincristine, procarbazine and temozolomide have been used. Temozolomide has an active metabolite that interrupts DNA replication by methylation of the 06 position of guanine, and is used for gliomas that have relapsed following radiotherapy. Most brain malignancies have a poor prognosis despite advances in imaging, surgery, chemotherapy and radiotherapy – less than 50% survival for high-grade gliomas at 2 years.

Idiopathic intracranial hypertension (IIH)

This syndrome, benign intracranial hypertension or pseudotumor cerebri, is included with tumours because marked papilloedema develops. There is neither a mass nor an increase in ventricular size. IIH develops mainly in obese young women with vague menstrual irregularities. *Headaches* and visual blurring with papilloedema are common features. A VIth nerve lesion may develop – a false localizing sign (p. 1106). CSF pressure is elevated, with normal constituents. Imaging is normal. Steroid therapy is sometimes thought to be a cause: many other drugs have occasionally been implicated. Other causes of papilloedema should be excluded. Sagittal sinus thrombosis can cause a similar picture.

IIH is benign in the sense that it is not fatal: it is usually self-limiting. However, optic nerve infarction can follow, with consequent blindness when papilloedema is severe and long-standing. Repeated lumbar puncture, acetazolamide, and thiazide diuretics are used in an attempt to reduce the intracranial pressure. Weight reduction is helpful. Surgical decompression or shunting is sometimes necessary.

FURTHER READING

Wen PY, Kesari S. Malignant gliomas in adults. *New England Journal of Medicine* 2008; **359**: 492–507.

Whittle IR, Smith C, Parthiban N et al. Meningiomas. *Lancet* 2004; **363**: 1535–1543.

HYDROCEPHALUS

Hydrocephalus means excessive water (CSF) within the head. In practice, hydrocephalus usually describes situations where there is, or has been, obstruction to CSF outflow, consequent high pressure and dilatation of the lateral and/or IIIrd and IVth ventricles (Fig. 21.42). Also, rarely, excessive CSF production occurs.

Infantile hydrocephalus

Head enlargement in infancy occurs in 1 in 2000 live births. There are several causes:

- *Arnold–Chiari malformation*. There is elongation of the medulla. Cerebellar tonsils descend into the cervical canal. Associated spina bifida is common. *Syringomyelia* may develop (p. 1164).

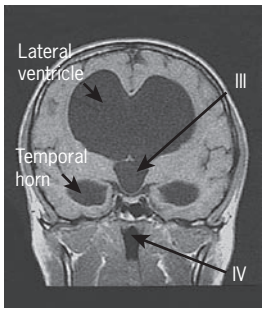


Fig. 21.42 **Hydrocephalus with gross IIIrd, IVth and lateral ventricle enlargement: MR T1.**

- *Stenosis of the aqueduct of Sylvius.* Aqueduct stenosis is either congenital or acquired following neonatal meningitis/haemorrhage.
- *Dandy–Walker syndrome*. There is cerebellar hypoplasia and obstruction to IVth ventricle outflow foramina.

Hydrocephalus in adult life

Hydrocephalus is sometimes an unsuspected finding on imaging. However, infantile hydrocephalus can become apparent in adult life (e.g. aqueduct stenosis). Combinations of headache, cognitive impairment, vomiting, papilloedema, ataxia and bilateral pyramidal signs develop. Hydrocephalus may develop in other circumstances:

- *Posterior fossa and brainstem tumours* obstruct the aqueduct or IVth ventricular outflow.
- *Following subarachnoid haemorrhage*, head injury or meningitis (particularly tuberculous).
- *A IIIrd ventricle colloid cyst* causes lateral ventricle enlargement, headache and papilloedema. These rare intraventricular tumours also sometimes produce intermittent hydrocephalus – recurrent prostrating headaches with episodes of lower limb weakness.
- *Choroid plexus papilloma* (rare) secretes CSF.

Frequently the underlying cause for hydrocephalus remains obscure.

Treatment

Ventriculo-atrial or ventriculo-peritoneal shunting is necessary when progressive hydrocephalus causes symptoms. Neurosurgical removal of tumours is carried out when appropriate, sometimes urgently.

Normal pressure hydrocephalus (NPH)

NPH describes a syndrome of enlarged lateral ventricles, dementia, urinary incontinence and gait apraxia, usually in the elderly. CSF pressure and constituents are typically normal. There is some debate about the reality of NPH as a true entity.Ventriculo-peritoneal shunting is occasionally carried out.

TRAUMATIC BRAIN INJURY

In most western countries head injury accounts for about 250 hospital admissions per 100 000 population annually. Trau-

matic brain injury (TBI) describes injuries with potentially permanent consequences. For each 100 000 people, 10 die annually following TBI; 10–15 are transferred to a neurosurgical unit – the majority of these require rehabilitation for a prolonged period of 1–9 months. The prevalence of survivors with a major persisting handicap is around 100 per 100 000. Road traffic accidents and alcohol abuse are the principal aetiological factors in this major cause of morbidity and mortality, in many countries.

Skull fractures

Linear skull fracture of the vault or base is one indication of the severity of a blow, but is itself not necessarily associated with any brain injury. Healing of linear fractures takes place spontaneously. *Depressed* skull fracture is followed by a high incidence of post-traumatic epilepsy. Surgical elevation and debridement are usually necessary.

Principal local complications of skull fracture are:

- *meningeal artery rupture* – causing extradural haematoma (p. 1135)
- *dural vein tears* – causing subdural haematoma (p. 1135)
- *CSF rhinorrhoea/otorrhoea* and consequent meningitis.

Mechanisms of brain damage

Older classifications attempted to separate *concussion,* transient coma for hours followed by apparent complete clinical recovery, from brain *contusion*, i.e. bruising, with prolonged coma, focal signs and lasting damage. Pathological support for this division is poor. Mechanisms of TBI are complex and interrelated:

- diffuse axonal injury – shearing and rotational stresses on decelerating brain, sometimes at the site opposite impact (the *contrecoup* effect)
- neuronal and axonal damage from direct trauma
- brain oedema and raised intracranial pressure
- brain hypoxia
- brain ischaemia.

Clinical course

In a mild TBI a patient is stunned or dazed for a few seconds or minutes. Following this the patient remains alert without post-traumatic amnesia. Headache can follow; complete recovery is usual. In more serious injuries duration of unconsciousness and particularly of post-traumatic amnesia (PTA) helps grade severity. PTA of more than 24 hours defines severe TBI. The Glasgow Coma Scale (GCS, p. 1123) is used to record the degree of coma; this has prognostic value. A GCS below 5/15 at 24 hours implies a serious injury; 50% of such patients die or remain in a vegetative or minimal conscious state (p. 1125). However, prolonged coma of up to some weeks is occasionally followed by good recovery.

Recovery after severe TBI takes many weeks or months. During the first few weeks, patients are often intermittently restless or lethargic and have focal deficits such as hemiparesis or aphasia. Gradually they become more aware, though they may remain in post-traumatic amnesia, being unable to lay down any continuous memory despite being awake. This amnesia may last some weeks or more, and may not be obvious clinically. PTA is one predictor of outcome. PTA over a week implies that persistent organic cognitive deficit is almost inevitable, although return to unsupported paid work may be possible.

Late sequelae

Sequelae of TBI are major causes of morbidity and can have serious social and medicolegal consequences. They include:

- *Incomplete recovery,* e.g. cognitive impairment, hemiparesis.
- *Post-traumatic epilepsy* (p. 1141).
- *The post-traumatic (post-concussional) syndrome*. This describes the vague complaints of headache, dizziness and malaise that follow even minor head injuries. Litigation is frequently an issue. Depression is prominent. Symptoms may be prolonged.
- *Benign paroxysmal positional vertigo* (BPPV, p. 1079).
- *Chronic subdural haematoma* (p. 1135).
- *Hydrocephalus* (p. 1162).
- *Chronic traumatic encephalopathy.* This follows repeated and often minor injuries. It is known as the punch-drunk syndrome and consists of cognitive impairment, extrapyramidal and pyramidal signs, seen typically in professional boxers.

Immediate management

Attention to the airway is vital. If there is coma, depressed fracture or suspicion of intracranial haematoma, CT imaging and discussion with a neurosurgical unit are essential. Indications for CT imaging vary from imaging all minor head injuries in some US centres to more stringent criteria elsewhere.

In many severe TBI cases, assisted ventilation will be needed. Intracranial pressure monitoring is valuable. Hypothermia lowers intracranial pressure when used early after a TBI; an effect on outcome has only been seen in specialized neurotrauma centres. Care of the unconscious patient is described on page 1124. Prophylactic antiepileptic drugs have been shown to be of no value in prevention of late post-traumatic epilepsy.

Rehabilitation

TBI cases require skilled, prolonged and energetic support. Survivors with severe physical and cognitive deficits require rehabilitation in specialized units. Rehabilitation includes care from a multidisciplinary team with physiotherapeutical, psychological and practical skills. Many survivors are left with cognitive problems (amnesia, neglect, disordered attention and motivation) and behavioural/emotional problems (temper dyscontrol, depression and grief reactions). Long-term support for both patients and families is necessary.

FURTHER READING

Polderman KH. Induced hypothermia and fever control for prevention and treatment of neurological injuries. *Lancet* 2008; **371**: 1955–1969.

Ropper AH, Gorson KC. Concussion. *New England Journal of Medicine* 2007; **356**: 166–172.

SPINAL CORD DISEASE

The cord extends from C1, the junction with the medulla, to the vertebral body of L1 where it becomes the conus medullaris. Blood supply is from the anterior spinal artery and a plexus on the posterior cord. This network is supplied by vertebral arteries, the thyrocervical trunk and several branches from lumbar and intercostal vessels.

SPINAL CORD COMPRESSION

Principal features of chronic and subacute cord compression are spastic paraparesis or tetraparesis, radicular pain at the

Table 21.47 Causes of spinal cord compression

Spinal cord neoplasms
Disc and vertebral lesions:
- Chronic degenerative
- Trauma

Inflammatory:
- Epidural abscess
- Tuberculosis
- Granuloma

Vertebral neoplasms:
- Metastases
- Myeloma

Epidural haemorrhage/haematoma
Rarities
Paget's disease, scoliosis and vertebral anomalies
Epithelial, endothelial and parasitic cysts
Aneurysmal bone cyst
Vertebral angioma
Haematomyelia, arachnoiditis
Osteoporosis with fracture
Cord arteriovenous malformation

Table 21.48 Principal spinal cord neoplasms

Extradural
Metastases:
- Bronchus
- Breast
- Prostate
- Lymphoma
- Thyroid
- Melanoma

Extramedullary
Meningioma
Neurofibroma
Ependymoma

Intramedullary
Glioma
Ependymoma
Haemangioblastoma
Lipoma
Arteriovenous malformation
Teratoma

level of compression, and sensory loss below the compression (Table 21.47).

For example, in compression at T4 (Fig. 21.14) a band of pain radiates around the thorax, characteristically worse on coughing or straining. Spastic paraparesis develops over months, days or hours, depending upon underlying pathology. Numbness commencing in the feet rises to the level of compression. This is called the *sensory level*. Retention of urine and constipation develop.

Cord compression is a medical emergency. It is sometimes difficult on clinical grounds alone to distinguish chronic progressive cord compression from other, non-surgical causes of worsening paraparesis and tetraparesis. One reason for this is that pain at the level of compression can be absent.

Causes

Spinal cord neoplasms. Extramedullary tumours, both extradural and intradural, cause cord compression (Table 21.48 and Fig. 21.43) gradually over weeks to months, often with root pain and a sensory level (p. 1197).

The rarer intramedullary tumours (e.g. glioma) typically progress slowly, sometimes over many years. Sensory disturbances similar to syringomyelia may develop (p. 1164).

Tuberculosis. Spinal TB is the commonest cause of cord compression in countries where TB is common. There is destruction of vertebral bodies and disc spaces, with local spread of infection. Cord compression and paraparesis follow, culminating in paralysis – Pott's paraplegia.

Disc and vertebral lesions. Central cervical disc and thoracic disc protrusion cause cord compression (p. 1177).

Spinal epidural abscess. This is described on page 1159.

Epidural haemorrhage and haematoma. These are rare sequelae of anticoagulant therapy, bleeding disorders or trauma and can follow lumbar puncture when clotting is

Fig. 21.43 **Thoracic meningioma compressing cord: MR T2.**

abnormal. A rapidly progressive cord or cauda equina lesion develops.

Management

Early recognition of cord compression is vital. MRI shows spinal cord pathology extremely well and is the imaging technique of choice. Plain spinal films show degenerative bone disease and vertebral destruction by infection or neoplasm. Routine tests (e.g. chest X-ray) may indicate a primary neoplasm or infection. Surgical exploration is frequently necessary; if decompression is not performed promptly, irreversible cord damage tends to follow. Results are excellent if benign tumours and haematomas are removed early. Radiotherapy is used to treat cord malignancies, with largely poor outcome.

FURTHER READING

Prasad D, Schiff D. Malignant spinal-cord compression. *Lancet Oncology* 2005; **6**: 15–24.

OTHER CAUSES OF PARAPARESIS

Paraparesis (and tetraparesis) occurs in many conditions recognizable by their clinical patterns. The elucidation of the cause of paraparesis has a pivotal place in neurology (see Table 21.13).

Syringomyelia and syringobulbia

The syrinx, a fluid-filled cavity within the cervical or thoracic spinal cord, is the essential feature. Syringobulbia means a cavity in the brainstem.

Aetiology and mechanism

Syringomyelia is frequently associated with the Arnold–Chiari malformation (p. 1169). The abnormality at the foramen magnum probably allows normal pulsatile CSF pressure waves to be transmitted to fragile tissues of the cervical cord and brainstem, causing secondary cavity formation. The syrinx is in continuity with the central canal of the cord.

Pathological anatomy

The expanding cavity in the cord, usually between C2 and T9, gradually destroys spinothalamic neurones, anterior horn cells and lateral corticospinal tracts. In the medulla (syringobulbia), trigeminal nuclei, sympathetic trunk, IXth, Xth, XIth and XIIth nerve nuclei and the vestibular connections are destroyed by the expanding syrinx.

Clinical features

Syringomyelia cases with the Arnold–Chiari malformation usually develop symptoms around the age of 20–30. Upper limb pain exacerbated by exertion or coughing is typical. Spinothalamic sensory loss – pain and temperature – leads to painless upper limb burns and trophic changes. Paraparesis develops. The following are typical signs of a substantial cervical syrinx (Fig. 21.44):

- *Areas of dissociated sensory loss*, i.e. spinothalamic loss without loss of light touch. Bizarre patterns are seen.
- *Loss of upper limb reflexes*.
- *Muscle wasting* in the hand and forearm.
- *Spastic paraparesis* – initially mild.
- *Neuropathic joints*, trophic skin changes (scars, nail dystrophy) and ulcers.
- *Brainstem signs* – as the syrinx extends into the brainstem (syringobulbia) there is tongue atrophy and fasciculation, bulbar palsy, nystagmus, Horner's syndrome, hearing loss and impairment of facial sensation.

Course, investigation and management

MRI demonstrates the cavity and herniation of cerebellar tonsils. Syringomyelia is gradually progressive over several decades. Sudden deterioration sometimes follows minor trauma, or occurs spontaneously.

There is no curative treatment. Surgical decompression of the foramen magnum sometimes slows deterioration.

Other causes of syrinx formation are bony anomalies at the foramen magnum, spina bifida (p. 1169), arachnoiditis, hydrocephalus (p. 1162), intrinsic cord tumours (e.g. glioma and ependymoma) and trauma.

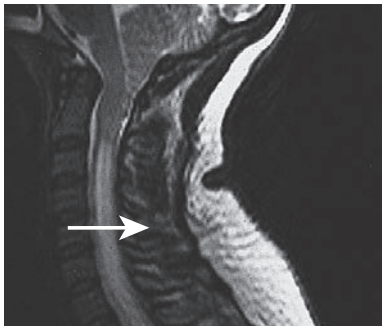


Fig. 21.44 **Cervical syrinx – cavity within cord: MR T2.**

Metabolic, toxic, inflammatory and vascular cord disease

Vitamin B_{12} deficiency

Subacute combined degeneration of the cord resulting from vitamin B_{12} deficiency (p. 399) is the most common example of metabolic disease causing spinal cord damage. Copper deficiency is an exceedingly rare cause of a similar picture. Cord lesions probably due to multiple B vitamin deficiencies are also seen in severe malnutrition.

Lathyrism

This is an endemic spastic paraparesis of central India caused by the toxin β-(*N*)-oxalylamino-L-alanine. It occurs when excessive quantities of a drought-resistant pulse, *Lathyrus sativa*, are consumed.

Konzo (tropical ataxic neuropathy)

In West Africa and the Caribbean inadequate preparation of cassava root leads to ingestion of cyanogenic glycosides. A subacute spastic paraparesis or quadraparesis, called konzo, follows. Tropical ataxic neuropathy consists of sensory ataxia, loss of reflexes, deafness and optic atrophy, probably due to chronic low-level exposure to these glycosides.

Acute transverse myelopathy (transverse myelitis)

This term is used to describe a cord lesion and paraparesis (or paraplegia) occurring with viral infections, MS, mixed connective tissue disease and other inflammatory and vascular disorders, e.g. HIV, sarcoid, syphilis, radiation myelopathy and anterior spinal artery occlusion. MRI is usually required to exclude cord compression.

Anterior spinal artery occlusion

Cord infarction, causing an acute paraplegia or tetraplegia (or paresis), occurs in many thrombotic or embolic vascular diseases, e.g. endocarditis, severe hypotension, atheroma, diabetes mellitus, polycythaemia, syphilis and polyarteritis. Cord infarction sometimes occurs during surgery to the posterior mediastinum, or follows aortic dissection and trauma. It also occurs as an isolated event.

Radiation myelopathy

Paraparesis with sensory loss sometimes develops within several weeks to a year following radiotherapy; precautions should be taken to shield the cord during radiotherapy.

MANAGEMENT OF PARAPLEGIA

General considerations

General health and morale should be reviewed carefully and regularly. Any intercurrent infection is potentially dangerous and should be treated early. Chronic renal failure is a common cause of death. The paraplegic patient needs skilled and prolonged nursing care and training to be aware of problems. Particular issues are discussed below.

Bladder. Catheterization is usually necessary initially. Many patients self-catheterize, or develop reflex bladder emptying, helped by abdominal pressure. Free urinary drainage is essential to avoid stasis, subsequent infection and calculi.

Bowel. Constipation and impaction must be avoided. Following acute paraplegia, manual evacuation is necessary; reflex emptying develops later.

Skin care. Risks of pressure sores and their sequelae are serious. Meticulous attention must be paid to cleanliness and to turning every 2 hours. The sacrum, iliac crests, greater trochanters, heels and malleoli should be inspected frequently (p. 1262). Ripple mattresses/water beds are useful. If pressure sores develop, plastic surgical repair may be required. Pressure palsies, e.g. of ulnar nerves, must be avoided.

Lower limbs. Passive physiotherapy helps to prevent contractures. Severe spasticity, with flexor or extensor spasms, may be helped by baclofen, diazepam, dantrolene, tizanidine or botulinum toxin injections.

Rehabilitation

Many patients with traumatic paraplegia or tetraplegia return to self-sufficiency. Specialist advice from a skilled rehabilitation unit is necessary. Lightweight, specially adapted wheelchairs are available. Patients with paraplegia have substantial practical, psychological and sexual needs.

DEGENERATIVE NEURONAL DISEASES

FURTHER READING

Mitchell JD, Borasio GD. Amyotrophic lateral sclerosis. *Lancet* 2007; **369**: 2031–2041.

Degenerative underlines our incomplete understanding of these progressive CNS diseases. Their molecular and genetic basis is the subject of intense study.

MOTOR NEURONE DISEASE (MND)

In this common disease, seen world-wide, there is progressive degeneration of lower and upper motor neurones in the spinal cord, in cranial nerve motor nuclei and within the cortex. Most MND is sporadic and of unknown cause. It is thought that relentless degeneration of motor neurones may be programmed genetically. Abnormal mRNA splicing of the excitatory amino-acid transporter (*EAAT2*) gene, a major glial transporter, was once thought to be specific for MND, but is sometimes found in normal people. Oxidative neuronal damage, aggregation of abnormal neuronal proteins, glutamate mishandling and abnormalities of neurofilament-mediated axonal transport are also involved. Apoptosis and prolonged neuronal caspase activity (p. 27) are the final pathways of cell destruction.

There is a rare familial form of MND in which there are mutations in the gene (chromosome 21q) encoding the free radical scavenging enzyme copper/zinc superoxide dismutase (SOD-1).

Incidence of the sporadic form is around 2 per 100 000 per year, with a slight male predominance. Onset occurs typically in the over 50s.

Clinical features

The sensory system is not involved. In the common, sporadic form of MND, four broad patterns are seen:

- amyotrophic lateral sclerosis (ALS)
- progressive muscular atrophy
- progressive bulbar and pseudobulbar palsy
- primary lateral sclerosis (rare).

Although useful as a means of recognizing MND, these are not distinct aetiological or pathological variants and usually merge as MND progresses. Awareness is preserved; dementia sometimes develops.

Amyotrophic lateral sclerosis (ALS)

Lateral sclerosis means disease of the lateral corticospinal tracts (i.e. one cause of spastic paraparesis). Amyotrophy means muscle atrophy, i.e. wasting, unusual in most other forms of spastic tetraparesis or paraparesis. The course is progressive spastic tetraparesis or paraparesis with added lower motor neurone signs and fasciculation. ALS is the term used for MND in the USA.

Progressive muscular atrophy

Wasting beginning in the small muscles of one hand (or both) spreads inexorably. Although it may commence unilaterally, wasting soon follows on the opposite side. Fasciculation is common. It is due to spontaneous firing of abnormally large motor units formed by branching fibres of surviving axons that are striving to innervate muscle fibres that have lost their nerve supply. Cramps may occur.

Physical signs are of wasting and weakness, with fasciculation – often widespread. Tendon reflexes are lost when the reflex arc is interrupted by anterior horn cell loss, but are preserved or exaggerated when there is loss of corticospinal motor neurones. Sphincter disturbance occurs late, if at all.

Progressive bulbar and pseudobulbar palsy

The lower cranial nerve nuclei and their supranuclear connections are initially involved. Dysarthria, dysphagia, nasal regurgitation of fluids and choking are common symptoms. This form of MND is more common in women. Characteristic features are bulbar and pseudobulbar palsy (p. 1110), with mixed upper and lower motor neurone signs in lower cranial nerves – for example, a wasted fibrillating tongue with a spastic weak palate. Eye movements remain unaffected in all forms of MND. There are neither cerebellar nor extrapyramidal signs.

Primary lateral sclerosis

This describes the least common form of MND confined to upper motor neurones. There is progressive tetraparesis with terminal pseudobulbar palsy.

Diagnosis

Diagnosis is largely clinical. There are no specific tests. Usually MND is easily identifiable clinically. Denervation, a feature of all forms of MND except primary lateral sclerosis, is confirmed by EMG. Chronic partial denervation with preserved motor conduction velocity is characteristic. CSF constituents are normal.

Cervical radiculopathy with myelopathy, syringomyelia and the rare, almost extinct, syphilitic cervical pachymeningitis can cause diagnostic difficulty. Motor neuropathies and spinal muscular atrophies sometimes resemble the progressive muscular atrophy form of MND – their course is more prolonged. Kennedy's syndrome, an X-linked bulbar and spinal muscular atrophy, sometimes causes confusion: upper motor neurone signs are not seen in this condition.

Bulbar myasthenia gravis (p. 1180) may sometimes appear similar in the early stages.

Multifocal motor neuropathy can appear like motor neurone disease (see p. 1172).

Course and management

Remission is unknown. MND progresses, spreading gradually, and causes death, often from bronchopneumonia. Survival for more than 3 years is unusual, although there are rare MND cases who survive for a decade or longer.

No treatment has been shown to influence outcome substantially. Riluzole, a sodium-channel blocker that inhibits glutamate release, slows progression slightly, particularly with disease of bulbar onset. Spasticity may be helped by baclofen, and drooling by propantheline or amitriptyline. Ventilatory support and feeding via gastrostomy helps prolong survival. Giving accurate prognostic advice to MND patients is particularly difficult. Shared-care protocols between patient, primary care physician, carers and support groups are helpful.

Spinal muscular atrophies

These rare genetic disorders of motor neurones give rise to slowly progressive, usually symmetrical, muscle wasting and weakness. An acute infantile type (Werdnig–Hoffmann disease), a chronic childhood type (Kugelberg–Welander disease) and adult forms are recognized. Clinically these conditions may be confused initially with muscular dystrophies (p. 1181), hereditary motor neuropathies or MND.

Dementia

Dementia describes progressive decline of cognitive function, usually affecting the cortex as a whole, though sometimes patchily. Memory is especially affected; intellect gradually fails. There is loss of emotional control, deterioration of social behaviour and loss of motivation. There are many causes of this syndrome, common in later life (Table 21.49). Dementia is a substantial cause of morbidity in any ageing population, with profound social and economic effects. Dementia affects some 10% of any population over 65, and 20% over 80. The commonest causes are Alzheimer's disease, fronto-temporal dementia, vascular dementia and dementia with Lewy bodies (DLB).

In practice it is often difficult to recognize each type of dementia, and to define landmarks that separate dementia from the mild decline in memory and cognitive function seen during ageing.

Table 21.49 Causes of dementia

Degenerative
Alzheimer's disease
Dementia with Lewy bodies
Frontotemporal dementia
Huntington's disease
Parkinson's disease
Hydrocephalus
Primary progressive aphasia
Vascular
Vascular dementia
Cerebral vasculitis, cranial arteritis (very rare)
Metabolic
Uraemia
Liver failure
Paraneoplastic syndromes
Toxic
Alcohol
Solvent abuse
Lead, mercury poisoning
Vitamin deficiency
B_{12}
Thiamin
Traumatic
Punch-drunk syndrome
Following traumatic brain injury
Intracranial lesions
Subdural haematoma
Tumours
Infections
Prion diseases, e.g. Creutzfeldt–Jakob
HIV
Neurosyphilis
Whipple's disease
Endocrine
Hypothyroidism
Hypoparathyroidism
Psychiatric
Pseudodementia

Differential diagnosis

This includes depression with pseudodementia, in which many of the features of an early dementia (especially memory impairment, slowed thinking and lack of spontaneity) are apparent. Patients with pseudodementia often have a previous history of depression or a family history of mood disorder. Successful treatment of the mood disorder results in restoration of intellectual function. Delirium and mild or moderate learning difficulties are also differential diagnoses.

Alzheimer's disease

This is the commonest dementia, a degenerative disease of the cortex, accounting for over 65% of dementia in any age group. The cardinal clinical features are:

- progressive loss of ability to learn, retain and process new information (i.e. memory loss)
- decline in language – difficulty in naming and in understanding what is being said (various aphasias, p. 1099)
- apraxia – impaired ability to carry out skilled motor activities (p. 1097)
- agnosia – failure to recognize objects, e.g. clothing, places or people
- progressive loss of executive function – organizing, planning and sequencing
- behavioural change – agitation, aggression, wandering and persecutory delusions
- loss of insight, relative or complete
- depression – though severe depression is unusual, perhaps because of loss of insight.

The onset is gradual. Alzheimer's disease (AD) is rare below the age of 50 but increasingly common thereafter. Disturbance of gait, motor and occasionally sensory abnormalities and seizures occur late. The course is progressive over

several to 10 or more years, with death in a state of extreme cognitive decline.

Neuropathology and neurochemical changes

Loss of neurones, neurofibrillary tangles, senile plaques and amyloid angiopathy are seen, especially within the frontal, temporal and parietal cortex, hippocampi, substantia innominata and locus caeruleus.

Neurofibrillary tangles in neuronal cell bodies are made up of the microtubular-associated tau protein in paired helical filaments. Ubiquitin is also seen in association with tangles. Aggregation of beta-amyloid (Aβ) appears to play a central role, developing progressively in extracellular senile plaques. These plaques contain deposits of Aβ-protein, produced by enzymes, β and γ secretase. Apoptosis follows, leading to cell death with a marked reduction in choline acetyltransferase, acetylcholine itself, noradrenaline and serotonin.

Aetiology

The cause of AD is unknown. No environmental toxin has been found. Incidence is increased in Down's syndrome. Increase in free radical formation and failure of antioxidant defences may contribute to neuronal degeneration. Superoxide dismutase (SOD), the free radical defence enzyme, is reduced by some 25% in AD frontal cortex and hippocampi.

AD is occasionally familial. Genetic studies have shown linkage between familial (presenile) AD and loci on chromosome 1 (presenilin, *PS1* gene), chromosome 14 (presenilin, *PS2* gene) and chromosome 21 (amyloid β-protein precursor, *APP* gene).

Late-onset AD (the great majority) is probably a heterogeneous disorder. Evidence suggests a link between AD and atherosclerosis, inflammation and cholesterol. Linkage has been found in some cases to a gene locus on chromosome 19q. The gene encoding apolipoprotein E is a possible candidate. This has three alleles, e2, e3 and e4. The e4 allele is a major risk factor for AD and the e2 allele is under-represented. It is possible that Aβ accumulation is more extensive in those with the e4 allele than in those lacking it.

Dementia with Lewy bodies (DLB)

This accounts for some 25% of dementias and is characterized by fluctuating cognition with pronounced variation in attention and alertness. Prominent or persistent memory loss may not occur in the early stages. Impairment in attention, frontal, subcortical and visuospatial ability is often prominent. Depression and sleep disorders occur. Recurrent formed visual hallucinations, e.g. strange faces, frightening creatures, are a common feature. Parkinsonism, i.e. slowing and rigidity, is common, with repeated falls. Delusions and transient loss of consciousness occur. Cortical Lewy bodies are prominent at autopsy. These inclusions, first described in idiopathic Parkinson's, are a hallmark of DLB.

Vascular dementia (multi-infarct dementia)

This common cause of dementia is distinguished from AD by its clinical features and imaging. Dementia can be progressive and similar to AD. There is sometimes a history of TIAs or the dementia follows a succession of cerebrovascular events. Widespread vessel disease seen on MRI is the typical finding and may produce a variety of cognitive deficits reflecting the site of ischaemic damage.

Frontotemporal dementia

Progressive deterioration of social behaviour, disinhibition and personality develops in middle life, followed by decline in memory, intellect and language. This accounts for some 3% of dementias. Imaging shows selective atrophy of frontal and temporal lobes. There is no excess of either senile plaques or neurofibrillary tangles. Cytoplasmic inclusion bodies – silver-staining, Pick's bodies – are seen. These changes separate this condition pathologically from AD.

Other dementias

Subcortical encephalopathies comprise disorders that affect subcortical structures such as the basal ganglia. They include dementia occurring in idiopathic Parkinson's, Huntington's, progressive supranuclear palsy (p. 1148) and dementia occurring in rarities such as cerebral vasculitis.

Rarer conditions such as Creutzfeldt–Jakob disease (p. 121), Wilson's disease (p. 357) and Whipple's (p. 281) cause progressive dementia. Dementia also occurs commonly in the late stages of MS (p. 1152). Dementia is usually one feature of these conditions, i.e. they are suspected or recognized by their clinical patterns.

Primary progressive aphasia – a language-based dementia – is recognized increasingly; there is a relentless deterioration of language skills with relative initial preservation of memory.

Investigations

Dementia is diagnosed from the history and basic examination, especially cognitive testing (p. 1189), and confirmed by psychometric testing. The history should be taken from someone who has known the patient for a long time. On MRI, bilateral atrophy in the hippocampi is seen in AD, with increased *tau* protein in the CSF and a reduction of glucose metabolism in bilateral temporo-parietal regions on PET scanning.

In later life, few cases of dementia are investigated in great detail. Defined medical conditions should be excluded (Table 21.50). Investigations in younger patients with dementia include detailed imaging, psychometry, EEG, CSF examination and brain biopsy, the latter only usually carried out in units specializing in dementia.

Management and social consequences

It is rare that a treatable cause for dementia is found, for example hypothyroidism. Management is supportive, to preserve dignity and to provide care for as long as possible in the familiar home environment. The burden of illness falls frequently on relatives. In an ageing population, medical problems of the carers themselves deserve particular attention. Some evidence suggests that participation in cognitively demanding activities in later life may protect against or delay the onset of dementia.

Treatment with antioxidants, mainly vitamin E, has shown some possible, slight benefit. Anticholinesterase inhibitors (donepezil, rivastigmine and galantamine) and memantine, an *N*-methyl D-aspartate (NMDA) receptor antagonist that affects glutamate transmission, have been shown to slow, if slightly, the rate of cognitive decline in AD. Whilst there is dispute about the place for these drugs – all are costly – they contribute in some cases to dementia patients being able to prolong independence and remain at home for longer than might otherwise be the case. There are complex cost-benefit debates.

Table 21.50 Tests in dementia

Blood tests
Full blood count
ESR, C-reactive protein
Urea and electrolytes
Blood glucose
Liver biochemistry
Serum calcium
Vitamin B_{12}, folate
TSH, T_3, T_4
Syphilis and HIV serology
Imaging
Chest X-ray
CT and MR brain
Other
Serum copper, genetic studies
EEG
CSF
Brain biopsy

FURTHER READING

Blennow K, de Leon MJ, Zetterberg H. Alzheimer's disease. *Lancet* 2006; **368**: 387–403.

Dubois B, Feldman HH, Jacova C. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS–ADRDA criteria. *Lancet Neurology* 2007; **6**: 734–746.

Rodriguez JJL, Ferri CP, Acosta D. Prevalence of dementia in Latin America, India, and China: a population-based cross-sectional survey. *Lancet* 2008; **372**: 464–474.

CONGENITAL AND INHERITED DISEASES

CEREBRAL PALSY

Cerebral palsy encompasses disorders apparent at birth or in childhood due to intra-uterine or neonatal brain damage; deficits are non-progressive. Learning problems, mild to severe, are frequent, though not exclusive – physical disability is independent of cognitive impairment.

The precise cause of damage in an individual child may be difficult to determine. The following are largely responsible:

- hypoxia in utero and/or during parturition
- neonatal cerebral haemorrhage and/or infarction
- trauma, neonatal or during parturition
- prolonged seizures – status epilepticus
- hypoglycaemia
- kernicterus with athetoid movement disorder – now rare with maternal Rh immunization.

Clinical features

Failure to achieve normal milestones is usually the earliest feature. Specific motor syndromes become apparent later in childhood or, rarely, in adult life.

- **Spastic diplegia** is spasticity, predominantly of lower limbs, with scissoring of gait.
- **Athetoid cerebral palsy** (p. 1149).
- **Infantile hemiparesis**. Hemiparesis may be noted at birth or later. One hemisphere is hypotrophic and the contralateral, hemiparetic limbs small (hemiatrophy).
- **Congenital ataxia** is incoordination and hypotonia of limbs and trunk.

DYSRAPHISM

Failure of normal fusion of the fetal neural tube leads to a group of congenital anomalies. Folate deficiency during pregnancy is contributory: supplements should always be given (p. 223). Antiepileptic drugs, e.g. valproate, are also implicated (p. 1143). If there is access from the skin, e.g. a sinus connecting to the subarachnoid space, bacterial meningitis may follow.

- **Anencephaly** is absence of brain and skull – incompatible with life.
- **Meningoencephalocele**. Brain and meninges extrude through a midline skull defect – protrusion can be minor or massive.
- **Spina bifida** is failure of lumbosacral neural tube fusion. Several varieties occur.
- **Spina bifida occulta** is isolated failure of vertebral arch fusion (usually lumbar), often seen incidentally on X-rays (3% of the population). A dimple or a tuft of hair may overlie the anomaly; clinical abnormalities are unusual.
- **Meningomyelocele with spina bifida**. Meningomyelocele consists of elements of spinal cord and lumbosacral roots within a meningeal sac. This herniates through a vertebral defect. In severe cases both lower limbs and sphincters are paralysed. The lumbosacral defect is visible at birth. Meningocele is a meningeal defect alone. The defect should be closed in the first 24 hours after birth. Subsequent abscess formation can develop.

BASILAR IMPRESSION OF SKULL (PLATYBASIA)

Platybasia is usually a congenital anomaly. There is upward invagination of the foramen magnum and skull base. Lower cranial nerves, medulla, upper cervical cord and roots are affected. A syrinx may be present. The Arnold–Chiari malformation may coexist, with aberrant cerebellar tonsils extending through the foramen magnum. Basilar invagination also develops in Paget's disease and rarely in osteomalacia (p. 567).

Clinical features are spastic tetraparesis, cerebellar and lower brainstem signs.

NEUROECTODERMAL SYNDROMES

Tissue derived from ectoderm forms tumours and hamartomas, with lesions in the skin, eye and nervous system.

Neurofibromatosis type 1 (von Recklinghausen's disease, NF-1)

This is characterized by multiple skin neurofibromas and pigmentation (café-au-lait patches). The neurofibromas arise from the neurilemmal sheath. One new case occurs in every 4000 live births. The mode of inheritance is autosomal dominant with complete penetrance.

NF-1 is autosomal with an abnormal gene on chromosome 17 (q11.2) which encodes for a protein, neurofibromin. This protein plays a role in cell growth.

FURTHER READING

Mitchell LE, Adzick NS, Melchionne J et al. Spina bifida. *Lancet* 20024; **364**: 1884–1895.

Skin neurofibromas present as soft subcutaneous, sometimes pedunculated, lumps (p. 1253). They increase in number throughout life. Multiple café-au-lait patches are present (p. 1265). Plexiform neurofibromas appear on major nerves and proximal nerve roots, sometimes involving the spinal cord. Treatment is surgical removal if pressure symptoms develop.

Neurofibromatosis type 2 (NF-2)

This is much less common than NF-1. It is also autosomal dominant; the gene for NF-2 is located on chromosome 22 (q12.2). The gene product (Merlin or Schwannomin) is a cytoskeletal protein. Many neural tumours occur:

- meningioma
- acoustic neuroma (often bilateral)
- glioma (including optic nerve glioma)
- plexiform neuroma (massive cutaneous overgrowth)
- cutaneous neurofibroma (30%).

Rarely, the benign tumours undergo sarcomatous change.

Associated abnormalities

- Scoliosis
- Orbital haemangioma
- Local gigantism of a limb
- Phaeochromocytoma and ganglioneuroma
- Renal artery stenosis
- Pulmonary fibrosis
- Obstructive cardiomyopathy
- Fibrous dysplasia of bone.

Treatment

Tumours causing pressure symptoms within the nervous system require excision, if feasible (e.g. VIIIth nerve schwannomas).

Tuberous sclerosis (epiloia)

This rare autosomal dominant condition comprises adenoma sebaceum (see p. 1253 for other skin lesions), epilepsy and cognitive impairment. Retinal phakomas (glial masses), renal tumours and glial overgrowth in the brain develop.

Sturge–Weber syndrome (encephalofacial angiomatosis)

There is an extensive port-wine naevus on one side of the face, usually in the distribution of a division of the Vth nerve, with a meningeal angioma. Epilepsy is common. Familial occurrence is exceptional.

von Hippel–Lindau syndrome (retinocerebellar angiomatosis)

This is dominantly inherited. Retinal and cerebellar haemangioblastomas develop or, less commonly, haemangioblastomas of the cord and cerebrum. Renal, adrenal and pancreatic tumours (and haemangioblastomas) may also be found. Polycythaemia sometimes develops.

There are numerous other rare disorders.

SPINOCEREBELLAR DEGENERATIONS (SCAs)

The classification of this group of inherited disorders is complex. SCAs are of interest; over 20 genetically defined SCAs have been recognized.

Friedreich's ataxia

This is an autosomal recessive progressive degeneration of dorsal root ganglia, spinocerebellar tracts, corticospinal tracts and cerebellar Purkinje cells. Most patients are homozygous for the GAA triplet expansion in the Friedreich's ataxia gene. This gene, mapped to chromosome 9q13, encodes a mitochondrial protein (fraxatin) of unknown function. There is abnormal function of mitochondrial ATP. The expression of fraxatin is decreased in Friedreich's patients. Progressive difficulty in walking occurs around the age of 12. Death is usual before 40. With the identification of the gene, mildly affected cases have been diagnosed in middle age.

Clinical findings are:

- ataxia of gait and trunk
- nystagmus (25%)
- dysarthria
- absent lower limb joint position and vibration sense
- absent lower limb reflexes
- optic atrophy (30%)
- pes cavus
- cardiomyopathy.

Ataxia telangiectasia

This rare, autosomal recessive condition is a progressive ataxic syndrome in childhood and early adult life (p. 70). There is striking telangiectasia of conjunctiva, nose, ears and skin creases. There are also defects in cell-mediated immunity and antibody production. A defect in DNA repair has been demonstrated. Death is usual by the third decade, either from infection or due to the development of lymphoma.

Hereditary spastic paraparesis with ataxia

Progressive paraparesis runs in some families. Additional features include prominent cerebellar signs, sometimes with pes cavus, wasted hands and optic atrophy. Various genetic linkages have been found. Paraparesis is usually mild, and with ataxia progresses slowly over years. Some cases have dystonic features and respond to levodopa.

PERIPHERAL NERVE DISEASE

The various nerve fibre types are shown in Table 21.51. All are myelinated except C fibres that carry impulses from pain receptors.

Mechanisms of damage to peripheral nerves

Peripheral nerves consist of two principal cellular structures – the nerve nucleus with its axon and the myelin sheath, which is produced by Schwann cells between each node

Ask the authors
www.studentconsult.com

of Ranvier (Fig. 21.1). Blood supply is via vasa nervorum. Six principal mechanisms, some coexisting, cause nerve malfunction.

Demyelination
Schwann cell damage leads to myelin sheath disruption. This causes marked slowing of conduction, seen for example in Guillain–Barré syndrome, post-diphtheritic neuropathy and many hereditary sensorimotor neuropathies.

Axonal degeneration
Axon damage leads to the nerve fibre dying back from the periphery. Conduction velocity initially remains normal (cf. demyelination) because axonal continuity is maintained in surviving fibres. Axonal degeneration occurs typically in toxic neuropathies.

Wallerian degeneration
This describes changes following nerve section. Both axon and distal myelin sheath degenerate, over several weeks.

Compression
Focal demyelination at the point of compression causes disruption of the myelin sheath. This occurs typically in entrapment neuropathies, e.g. carpal tunnel syndrome (p. 508).

Infarction
Microinfarction of vasa nervorum occurs in diabetes and arteritis, e.g. polyarteritis nodosa, Churg–Strauss syndrome (p. 870). Wallerian degeneration occurs distal to the ischaemic zone.

Infiltration
Infiltration of peripheral nerves by inflammatory cells occurs in leprosy and granulomas, e.g. sarcoid, and by neoplastic cells, causing neural metastases.

Nerve regeneration

Regeneration occurs either by remyelination – recovering Schwann cells spin new myelin sheaths around an axon – or by axonal growth down the nerve sheath with sprouting from the axonal stump. Axonal growth takes place at up to 1 mm/day.

Definitions (Fig. 21.45)

- **Neuropathy** simply means a pathological process affecting a peripheral nerve or nerves.
- **Mononeuropathy** means a process affecting a single nerve.
- **Mononeuritis multiplex** (*multiple mononeuropathy* and or *multifocal neuropathy*) affects several or multiple nerves.
- **Polyneuropathy** describes diffuse, symmetrical disease, usually commencing peripherally. The course may be acute, chronic, static, progressive, relapsing or towards recovery. Polyneuropathies are motor, sensory, sensorimotor and autonomic. They are classified broadly into demyelinating and axonal types, depending upon which principal pathological process predominates. It is often impossible to separate these clinically. Many systemic diseases cause neuropathies. Widespread loss of tendon reflexes is typical, with distal weakness and distal sensory loss.
- **Radiculopathy** means disease affecting nerve roots and *plexopathy*, the brachial or lumbosacral plexus.
- **Myelopathy** means disease of the cord.

Terms such as radiculomyelopathy and radiculoneuropathy describe various combinations of these processes.

Diagnosis is made by clinical pattern, nerve conduction/EMG, nerve biopsy, usually sural or radial, and identification of systemic or genetic disease.

Table 21.51 A, B and C fibres in peripheral nerves

Diameter (μm)	Conduction velocity m/s	Function
Aα (1–20)	c. 70	Motor; proprioception
Aβ (5–10)	30–60	Touch
Aγ (3–6)	20–30	Fusimotor, to spindles
Aδ (2–5)	20–30	Sharp pain
B (<3)	5–15	Autonomic, preganglionic
C (<1.3)	0.5–2	Slow pain

MONONEUROPATHIES

Peripheral nerve compression and entrapment (Table 21.52)

Nerve damage by compression is either acute, e.g. due to a tourniquet, or chronic, such as in entrapment neuropathies. In both, focal demyelination predominates at the compression site, and some distal axonal degeneration occurs. Acute compression usually affects nerves exposed anatomically, e.g. the common peroneal nerve at the head of the fibula, or the ulnar nerve at the elbow. Entrapment develops in relatively tight anatomical passages, e.g. the carpal tunnel.

These neuropathies are recognized largely by clinical features. Diagnosis is confirmed by nerve conduction studies.

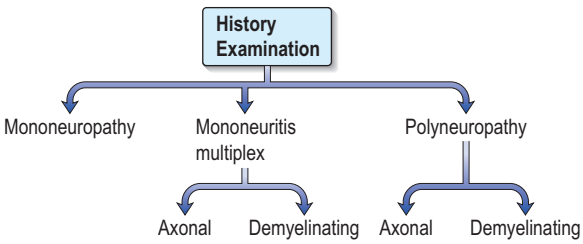


Fig. 21.45 **Peripheral neuropathies.** The type of neuropathy (axonal or demyelinating) can be assessed by electrical nerve studies (p. 1121).

Table 21.52 Nerve compression and entrapment

Nerve	Entrapment/compression site
Median	Carpal tunnel (wrist)
Ulnar	Cubital tunnel (elbow)
Radial	Spiral groove (of humerus)
Posterior interosseous	Supinator muscle (forearm)
Lateral cutaneous of thigh	Inguinal ligament
Common peroneal	Neck of fibula
Posterior tibial	Tarsal tunnel (flexor retinaculum – foot)

The commonest are mentioned here. All are seen more frequently in people with diabetes.

Carpal tunnel syndrome (CTS) (p. 508)

This common mononeuropathy, median nerve entrapment at the wrist, is usually known as carpal tunnel syndrome (CTS). CTS is typically not associated with any underlying disease. CTS is, however, seen in:

- hypothyroidism
- diabetes mellitus
- pregnancy (third trimester)
- obesity
- rheumatoid disease
- acromegaly
- amyloid
- renal dialysis patients.

CTS also occurs following wrist trauma.

There is nocturnal tingling and pain in the hand and/or forearm, followed by weakness of thenar muscles. Sensory loss in the palm and radial three-and-a-half fingers develops, followed by wasting of abductor pollicis brevis. *Tinel's sign* is often present and *Phalen's test* positive. Tinel's is elicited by tapping the flexor aspect of the wrist: this causes tingling and pain. In Phalen's, symptoms are reproduced on passive maximal wrist flexion.

A wrist splint at night or a local steroid injection (p. 509) in the wrist gives relief in mild cases. In pregnancy CTS is often self-limiting as fluid retention subsides postpartum. Reduction in obesity usually alleviates CTS. Surgical decompression of the carpal tunnel is the definitive treatment.

Ulnar nerve compression

The nerve is compressed in the cubital tunnel at the elbow. This follows prolonged or recurrent pressure and elbow fracture.

Weakness and wasting of ulnar innervated muscles leads to clawing of the hand – hypothenar muscles, interossei and medial two lumbricals – with sensory loss in the ulnar one-and-a-half fingers. Decompression and transposition of the nerve at the elbow is sometimes helpful.

The deep, solely motor branch of the ulnar nerve can be damaged in the palm by repeated trauma, e.g. from a crutch, screwdriver handle, or cycle handlebars.

Radial nerve compression

The radial nerve is compressed acutely against the humerus, e.g. when the arm is draped over a hard chair for several hours, known as Saturday night palsy. Wrist drop and weakness of brachioradialis and finger extension follow. Recovery is usual, though not invariable, within 1–3 months. Posterior interosseous nerve compression in the forearm also leads to wrist drop, without weakness of brachioradialis.

Lateral cutaneous nerve of the thigh compression

This is also known as meralgia paraesthetica and is described on page 512.

Common peroneal nerve palsy

The common peroneal nerve (lateral popliteal nerve) is compressed against the head of the fibula following prolonged squatting, yoga, pressure from a cast, prolonged bed rest or coma, or for no apparent reason. There is foot drop and weakness of ankle eversion. The ankle jerk (S1) is preserved. A patch of numbness develops on the anterolateral border of the dorsum of the foot and/or lower shin – the nerve has a small sensory twig. Recovery is usual, though not invariable, within several months.

Multifocal neuropathy (mononeuritis multiplex, multiple mononeuropathy)

This occurs in:

- diabetes mellitus
- leprosy
- vasculitis
- sarcoidosis
- amyloidosis
- malignancy
- neurofibromatosis
- HIV infection
- Guillain–Barré syndrome – usually polyneuropathy, see page 1172
- idiopathic multifocal motor neuropathy.

Diagnosis is largely clinical, supported by electrical studies. Several nerves become affected sequentially or simultaneously, e.g. ulnar, median, radial and lateral popliteal nerves. When multifocal neuropathy is symmetrical, there is difficulty distinguishing it from polyneuropathy. Where leprosy is prevalent, e.g. in India, a single nerve lesion can be the presenting feature.

Idiopathic multifocal motor neuropathy

A distal motor neuropathy (often asymmetrical and predominantly in the hands) of unknown cause develops gradually over months with profuse fasciculation, hence confusion with motor neurone disease (p. 1166). Conduction block and denervation are seen electrically. Antibodies to the ganglioside GM_1 are found in over 50% of cases; this is non-specific – antibodies are sometimes seen in other neuropathies, e.g. Guillain–Barré syndrome.

Steroids and/or cyclophosphamide and intravenous immunoglobulin slow the condition in some cases.

POLYNEUROPATHIES

Many diseases cause polyneuropathy, though the aetiology sometimes remains obscure. The most common presentations are acute, chronic or subacute sensorimotor neuropathies. A classification is given in Table 21.53. Antibodies to different gangliosides are emerging as diagnostic tools.

Guillain–Barré syndrome (GBS)

Clinical features

GBS is the most common acute polyneuropathy (3/100 000/year); it is usually demyelinating but occasionally axonal and probably has an autoallergic basis. GBS is monophasic – it does not recur. GBS is also known as acute inflammatory or postinfective neuropathy, acute inflammatory demyelinating polyradiculoneuropathy and Landry-GBS. The clinical spec-

FURTHER READING

England JD, Ashbury AK. Peripheral neuropathy. *Lancet* 2004; **363**: 2151–2161.

Table 21.53 Varieties of polyneuropathy

- Guillain–Barré syndrome
- Chronic inflammatory demyelinating polyradiculoneuropathy
- Diphtheritic polyneuropathy
- Idiopathic sensorimotor neuropathy
- Metabolic, toxic and vitamin deficiency neuropathies (Table 21.54)
- Drug-related neuropathies (Table 21.55)
- Hereditary sensorimotor neuropathies, e.g. Charcot–Marie–Tooth
- Other polyneuropathies:
  - Neuropathy in cancer
  - Neuropathies in systemic diseases
  - Autonomic neuropathy
  - HIV-associated neuropathy
  - Critical illness neuropathy

trum of GBS extends to an acute motor axonal neuropathy and the Miller Fisher syndrome – a rare proximal form causing ocular muscle palsies and ataxia.

Paralysis follows 1–3 weeks after an infection that is often trivial and seldom identified. *Campylobacter jejuni* and cytomegalovirus infections are well-recognized causes of severe GBS. Infecting organisms induce antibody responses against peripheral nerves. Molecular mimicry, i.e. sharing of homologous epitopes between microorganism liposaccharides and nerve gangliosides (e.g. GM_1), is the possible mechanism.

The patient complains of weakness of distal limb muscles and/or distal numbness. These symptoms progress proximally, over several days to 6 weeks. Loss of tendon reflexes is almost invariable. In mild cases there is mild disability before spontaneous recovery begins, but in some 20% respiratory and facial muscles become weak, sometimes progressing to complete paralysis. Autonomic features sometimes develop.

Diagnosis

This is established on clinical grounds and confirmed by nerve conduction studies; these show slowing of conduction in the common demyelinating form, prolonged distal motor latency and/or conduction block. CSF protein is often raised to 1–3 g/L; cell count and glucose level remain normal.

In the Miller Fisher syndrome antibodies against GQ1b (ganglioside) have a sensitivity of 90%.

Differential diagnosis includes other acute paralytic illnesses, e.g. botulism, cord compression, muscle disease and myasthenia.

Course and management

Paralysis may progress rapidly (hours/days) to require ventilatory support. It is essential that ventilation (vital capacity, blood gases) is monitored repeatedly to recognize emerging respiratory muscle weakness. Subcutaneous heparin (p. 444) should be given to reduce the risk of venous thrombosis.

Immunoglobulin given intravenously within the first 2 weeks reduces duration and severity of paralysis. There is concern about using this expensive pooled blood product. Patients should be screened for IgA deficiency before immunoglobulin is given – severe allergic reactions due to IgG antibodies may occur when congenital IgA deficiency is present. Angina or myocardial infarction can be precipitated by i.v. immunoglobulin. Plasma exchange is also of proven benefit in shortening disability, though seldom used. Steroids used to be given, but are valueless. Recovery begins, with or without treatment, between several days and 6 weeks from the outset. Prolonged ventilation may be necessary. Improvement towards independent mobility is gradual over many months but may be incomplete. Some 15% either die or are left disabled.

Chronic inflammatory demyelinating polyradiculoneuropathy (CIDP)

CIDP develops over weeks or months, usually relapsing and remitting, but generally persists long term. There are several CIDP varieties: the usual demyelinating form, an axonal form and multifocal neuropathy. In some cases plaques resembling MS lesions are seen on MRI, both in the brain and cord, but typically without the clinical features of MS. CSF protein is raised and segmental demyelination usually seen in nerve biopsy specimens. CIDP responds to long-term, low-dose steroids and to i.v. immunoglobulin, which is used for exacerbations.

Progress is variable; with drug therapy most CIDP cases run a mild course over many years. Recovery occasionally occurs.

Chronic sensorimotor neuropathy: no cause found

This situation is not uncommon – progressive symmetrical numbness and tingling occurs in hands and feet, spreading proximally in a glove and stocking distribution. Distal weakness also ascends. Tendon reflexes are lost. Symptoms may progress, remain static or occasionally remit. Autonomic features are sometimes seen.

Nerve conduction studies show either axonal degeneration or demyelination, or features of both, and no underlying cause is found. Nerve biopsy helps to classify some cases, for example diagnosing CIDP, unsuspected vasculitis or hereditary neuropathy.

Diphtheritic neuropathy

Palatal weakness followed by pupillary paralysis and a sensorimotor neuropathy occur several weeks after the throat infection (p. 126).

Cranial polyneuropathy

This means multiple cranial nerve lesions. These can develop in diabetes, malignant infiltration, particularly with nasopharyngeal and breast carcinoma, lymphomas, sarcoidosis, as part of a paraneoplastic syndrome and occasionally with giant cell arteritis. Relapsing and remitting cranial polyneuropathy of unknown cause is sometimes seen in South East Asia.

Metabolic, toxic and vitamin deficiency neuropathies

Causes of the most common neuropathies are shown in Table 21.54.

Table 21.54 Metabolic, toxic and vitamin deficiency neuropathies

Metabolic
Diabetes mellitus
Uraemia
Hepatic disease
Thyroid disease
Porphyria
Amyloid disease
Malignancy
Refsum's disease
Critical illness

Toxic
Drugs (Table 21.55)
Alcohol
Industrial toxins, e.g. lead, organophosphates

Vitamin deficiency
B_1 (thiamin)
B_6 (pyridoxine)
Nicotinic acid
B_{12}

Metabolic neuropathies

Diabetes mellitus. Several varieties of neuropathy occur (p. 1055):

- symmetrical sensory polyneuropathy
- acute painful neuropathy
- mononeuropathy and multiple mononeuropathy:
  - cranial nerve lesions
  - isolated peripheral nerve lesions (e.g. carpal tunnel syndrome)
- diabetic amyotrophy
- autonomic neuropathy.

Uraemia. Progressive sensorimotor neuropathy develops in chronic uraemia. Response to dialysis is variable; the neuropathy usually improves after transplantation.

Thyroid disease. A mild chronic sensorimotor neuropathy is sometimes seen in both hyperthyroidism and hypothyroidism (pp. 988 and 986). Myopathy also occurs in hyperthyroidism (p. 986).

Porphyria. In acute intermittent porphyria (p. 1073) there are episodes of a severe, mainly proximal neuropathy in the limbs, sometimes with abdominal pain, confusion and coma. Alcohol, barbiturates and intercurrent infection can precipitate attacks.

Amyloidosis. Polyneuropathy or multifocal neuropathy develops (p. 1072).

Refsum's disease

This is an autosomal recessive rarity. It is a treatable sensorimotor polyneuropathy with ataxia, retinal damage and deafness, due to defective phytanic acid metabolism.

Table 21.55 Drug-related neuropathies

Drug	Neuropathy	Mode/site of action
Phenytoin	M	A
Chloramphenicol	S, M	A
Metronidazole	S, S/M	A
Isoniazid	S, S/M	A
Dapsone	M	A
Antiretroviral drugs	S > M	A
Nitrofurantoin	S/M	A
Vincristine	S > M	A
Paclitaxel	S > M	A
Disulfiram	S, M	A
Cisplatin	S	A
Amiodarone	S, M	D, A
Chloroquine	S, M	A, D
Suramin	M > S	D, A

A, axonal; D, demyelinating; M, motor; S, sensory.

Toxic neuropathies

Alcohol

Polyneuropathy, mainly in the lower limbs, occurs with chronic alcohol abuse. Calf pain is common. The response to thiamin treatment is variable, even with complete alcohol abstention. Recurrence and progression occur with even small amounts of alcohol.

Drugs and industrial toxins

Many drugs (Table 21.55) and a wide variety of industrial toxins cause polyneuropathy. Toxins include:

- lead – motor neuropathy
- acrylamide (plastics industry), trichlorethylene, hexane, fat-soluble hydrocarbons, e.g. glue-sniffing, page 1214
- arsenic and thallium – polyneuropathy, initially sensory.

Vitamin deficiencies

Vitamin deficiencies cause nervous system damage that is potentially reversible if treated early, and inexorably progressive if not. Deficiencies, often of multiple vitamins, develop in malnutrition.

Thiamin (vitamin B_1)

Dietary deficiency causes *beriberi* (p. 220). Its principal features are polyneuropathy and cardiac failure. Thiamin deficiency also leads to an amnestic syndrome (Wernicke–Korsakoff psychosis, see below). Alcohol is the commonest cause in western countries and, rarely, anorexia nervosa. For other neurological consequences of alcohol, see Table 21.56.

Wernicke–Korsakoff syndrome. This thiamin-responsive encephalopathy is due to ischaemic damage in the brainstem and its connections. It consists of:

- *eye signs* – nystagmus, bilateral lateral rectus palsies, conjugate gaze palsies, fixed pupils and, rarely, papilloedema
- *ataxia* – broad-based gait, cerebellar signs and vestibular paralysis with absent calorics
- *cognitive change* – amnestic syndrome with confabulation, restlessness, stupor and coma

Table 21.56 Neurological effects of ethyl alcohol

Acute intoxication:
Disturbance of balance, gait and speech
Coma
Head injury and sequelae
Alcohol withdrawal:
Morning shakes and 'dry pukes'
Tremor
Delirium tremens
Thiamin deficiency:
Polyneuropathy
Wernicke–Korsakoff
Epilepsy
Acute intoxication
Alcohol withdrawal
Hypoglycaemia
Cerebellar degeneration
Cerebral infarction
Cerebral atrophy, dementia
Central pontine myelinolysis
Marchiafava–Bignami syndrome (corpus callosum degeneration, rare)

- *delirium tremens* (p. 1213)
- *hypothermia and hypotension* – due to hypothalamic involvement (rare).

Wernicke–Korsakoff syndrome is underdiagnosed. Thiamin should be given parenterally if the diagnosis is a possibility. The vitamin is harmless but should be given slowly as anaphylaxis can occur (p. 221). Untreated Wernicke–Korsakoff syndrome commonly leads to an irreversible amnestic state (Table 21.4, p. 1150). Erythrocyte transketolase activity is reduced but the test is rarely available.

Pyridoxine (vitamin B_6)

Deficiency causes a mainly sensory neuropathy. In practical terms this is seen as limb numbness developing during anti-TB therapy in slow isoniazid acetylators (p. 925). Prophylactic pyridoxine 10 mg daily is given with isoniazid.

Vitamin B_{12} (cobalamin)

Deficiency causes damage to the spinal cord, peripheral nerves and brain.

Subacute combined degeneration of the cord (SACD). Combined cord and peripheral nerve damage is a sequel of Addisonian pernicious anaemia and rarely other causes of vitamin B_{12} deficiency (p. 399). Initially there is numbness and tingling of fingers and toes, distal sensory loss, particularly posterior column, absent ankle jerks and, with cord involvement, exaggerated knee jerks and extensor plantars. Optic atrophy and retinal haemorrhage may occur. In later stages sphincter disturbance, severe generalized weakness and dementia develop. Exceptionally, dementia develops in the early stages.

Macrocytosis with megaloblastic marrow is usual though not invariable in SACD. Parenteral B_{12} reverses nerve damage but has little effect on the cord and brain. Without treatment, SACD is fatal within 5 years. Copper deficiency is a very rare cause of a similar picture.

Hereditary sensorimotor neuropathies (HSMN)

These are large and complex groups, phenotypically similar but genetically separate.

Charcot–Marie–Tooth disease

Charcot–Marie–Tooth (CMT) disease is also called peroneal muscular atrophy. CMT describes the common clinical phenotype (1 in 2500 persons), distal limb wasting and weakness that slowly progresses over many years, mostly in the legs, with variable loss of sensation and reflexes. In advanced disease, distal wasting is so marked that the legs are said to resemble inverted champagne bottles. Mild cases have pes cavus and toe clawing that can pass unnoticed. Multiple genetic variants of the CMT phenotype are recognized, e.g.:

- *HSMN Ia (CMT 1A)* – the commonest (70% of CMT; 1:2500 births), autosomal dominant demyelinating neuropathy caused by duplication (or point mutation) of a 1.5 megabase portion p11.2 of chromosome 17 encompassing the peripheral myelin protein 22 gene (*PMP-22*, 17p11.2).
- *HSMN Ib (CMT 1B)* – the second commonest, an autosomal dominant demyelinating neuropathy due to mutations in the myelin protein zero gene (*MPZ*) on chromosome 1 (1q22).
- *HSMN II (CMT 2)* – rare axonal polyneuropathies also caused by *MFN2* or *KIFIB* on chromosome 1p36 and other mutations. There is prominent sensory involvement with pain and parasthesias.
- *Distal spinal muscular atrophy* – a rare cause of CMT phenotype.
- *CMT with optic atrophy*, deafness, retinitis pigmentosa and spastic paraparesis.
- *CMTX* is an X-linked dominant HSMN on chromosome Xq13.1. The gene product is a gap junction B1 protein (GJB1) or connexin 32 (p. 26).

HSMN III

HSMN III is a rare childhood demyelinating sensory neuropathy (Déjérine–Sottas disease) leading to severe incapacity during adolescence. Nerve roots become hypertrophied. CSF protein is greatly elevated to 10 g/L or more. Point mutations either of *PMP-22* gene or of *P0* can generate this phenotype.

Other polyneuropathies

Neuropathy in cancer

Polyneuropathy is seen as a paraneoplastic syndrome (non-metastatic manifestation of malignancy), sometimes with anti-neuronal antibodies (Box 9.1). Polyneuropathy occurs in myeloma and other dysproteinaemic states, probably owing to impaired perfusion of nerve trunks or to demyelination associated with allergic reactions affecting peripheral nerves. Individual nerves are sometimes infiltrated with malignant cells.

Neuropathies in systemic diseases

Vasculitic neuropathy occurs in SLE (p. 541), polyarteritis nodosa (p. 550), Churg–Strauss syndrome (p. 870), rheumatoid disease (p. 528), sarcoidosis and giant cell arteritis (p. 550). Both multifocal neuropathy and symmetrical sensorimotor polyneuropathy occur.

FURTHER READING

Triggs WJ, Brown RH, Menkes DL. Charrot-Marie-Tooth case 18-2006. *New England Journal of Medicine* 2006; **354**: 2584–2592.

POEMS syndrome. This rarity consists of chronic inflammatory demyelinating *P*olyneuropathy, *O*rganomegaly (hepatomegaly 50%), *E*ndocrinopathy (gynaecomastia and atrophic testes), an *M* protein band, and *S*kin hyperpigmentation.

Autonomic neuropathy

Autonomic neuropathy causes postural hypotension, urinary retention, erectile dysfunction, diarrhoea or occasionally constipation, diminished sweating, impaired pupillary responses and cardiac arrhythmias. This can develop in diabetes and amyloidosis and may complicate Guillain–Barré syndrome. Many varieties of neuropathy cause autonomic problems in a mild form. Occasionally, with severe damage to small myelinated and non-myelinated B and C fibres, features of the autonomic neuropathy predominate, e.g. disabling postural hypotension.

HIV-associated polyneuropathies

Several neuropathies occur. Inflammatory demyelinating neuropathies similar to Guillain–Barré syndrome and CIDP are seen early in HIV infection, particularly during seroconversion. The most common neuropathy seen later in HIV is a chronic distal neuropathy that is symmetrical and painful.

Severe lumbosacral polyradiculoneuropathy develops develops, with profound lower limb flaccid paralysis, areflexia and sphincter dysfunction.

All HIV-associated neuropathies respond poorly to retroviral therapy; the drugs themselves cause neuropathies (p. 192).

Critical illness polyneuropathy (p. 916)

Some 50% of critically ill ITU patients with multiple organ failure and/or sepsis develop an axonal polyneuropathy. The severe inflammatory response impairs neural metabolism. Typically distal weakness and absent reflexes are seen during recovery from critical illness. Resolution is usual.

PLEXUS AND NERVE ROOT LESIONS

The common conditions that cause these are summarized in Table 21.57.

Cervical and lumbar degeneration

Spondylosis (Tables 10.3 and 10.6) describes vertebral and ligamentous degenerative changes occurring during ageing or following trauma. Several factors produce neurological signs and symptoms:

- osteophytes – local overgrowth of bony spurs or bars
- thickening of spinal ligaments
- congenital narrowing of the spinal canal
- disc degeneration and protrusion (posterior and lateral protrusion: cord and root compression)
- vertebral collapse (osteoporosis, infection)
- rheumatoid synovitis (p. 525)
- ischaemic changes within cord and nerve roots.

Narrowing of disc spaces, osteophytes, narrowing of exit foramina, and narrowing of the spinal canal are also seen on X-rays and MRI in the *symptomless* population, commonly in the mid and lower cervical and lower lumbar region, and imaging must not be falsely interpreted.

Lateral cervical disc protrusion (Fig. 21.46)

The patient complains of pain in the arm. A C7 protrusion is the most common problem. There is root pain that radiates to the C7 myotome (triceps, deep to scapula and extensor aspect of forearm), with a sensory disturbance, tingling and numbness in the C7 dermatome.

In an established C7 root lesion there is:

- weakness/wasting – triceps, wrist and finger extensors
- loss of the triceps jerk (C7 reflex arc)
- C7 dermatome sensory loss.

Although the initial pain can be very severe, most cases recover with rest and analgesics. It is usual to immobilize the neck. Disc protrusion with root compression is seen on MRI. Root decompression is sometimes helpful.

Lateral lumbar disc protrusion

The L5 and S1 roots are commonly compressed by lateral prolapse of L4–L5 and L5–S1 discs – the root number below a disc interspace is compressed. There is low back pain and *sciatica*, i.e. pain radiating from the back to buttock and leg. Onset is typically acute. This can follow lifting, bending or minor injury. When pain follows such an event, it is tempting to ascribe the disc protrusion to it. However, lateral lumbar disc protrusion is commonly apparently spontaneous – lifting or injury is usually only bringing forward an inevitable disc prolapse.

Straight-leg raising is limited. There is reflex loss, e.g. ankle jerk in an S1 root lesion, weakness of plantar flexion

Table 21.57 Common root and plexus problems

Nerve root
Cervical and lumbar spondylosis
Trauma
Herpes zoster
Tumours e.g. neurofibroma, metastases
Meningeal inflammation, e.g. syphilis, arachnoiditis
Plexus
Trauma
Malignant infiltration
Neuralgic amyotrophy
Thoracic outlet syndrome (cervical rib)

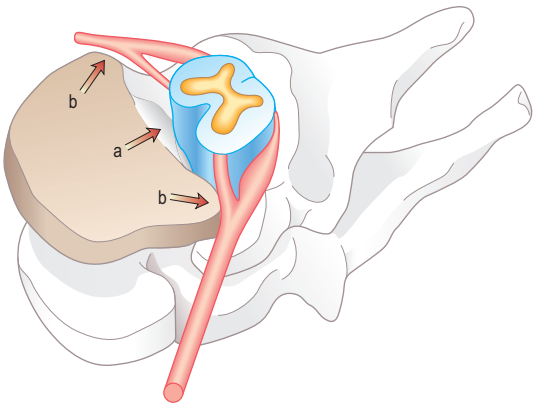


Fig. 21.46 **Central and lateral disc protrusions.** **(a)** Central disc protrusion compressing cord. **(b)** Lateral disc protrusion compressing nerve roots.

(S1) or great toe extension (L5). Sensory loss is found in the affected dermatome.

Most sciatica resolves with initial rest and analgesia followed by early mobilization. MRI is sometimes appropriate: surgery is indicated when a substantial persistent symptomatic disc lesion is shown.

Acute low back pain

Acute low back pain is extremely common. Most pain is of disc or facet joint origin. Significant nerve root compression is unusual. Bed rest on hard boards, corsets or casts have long been advocated: the evidence for these is scanty. Activity and a trial of gentle manipulation is preferred (see also p. 511).

Central disc protrusion (p. 509)

Central cervical disc protrusion (cervical myelopathy)

Posterior disc protrusion (Fig. 21.47), common at C4–5, C5–6 and C6–7 levels, causes spinal cord compression. Congenital spinal canal narrowing, osteophytic bars, ligamentous thickening and ischaemia are contributory. Frequently there are no or few neck symptoms. The patient complains of difficulty walking. Spastic paraparesis or tetraparesis is found, with variable sensory loss. A reflex level in the upper limbs and evidence in the arms of lateral disc protrusion may coexist. MRI demonstrates the level and extent of cord compression. Neck manipulation should be avoided.

Cervical laminectomy or anterior fusion of the vertebral bodies with removal of the disc may be necessary when cord compression is severe or progressive. The results of surgery are variable. Complete recovery of the pyramidal signs is unusual; progression may be halted.

Central thoracic disc protrusion

Central protrusion of a thoracic disc is an unusual cause of paraparesis.

Central lumbosacral disc protrusion

A central disc protrusion causes a *cauda equina syndrome*, i.e. bilateral flaccid (cf. spasticity in higher lesions) lower limb weakness, sacral numbness, retention of urine, erectile dysfunction and areflexia – usually with back pain. Many lumbosacral nerve roots are involved. Onset is either acute – an acute flaccid paraparesis – or chronic, sometimes with intermittent claudication. A central lumbosacral protrusion should be suspected if a patient with back pain develops retention of urine or sacral numbness. Urgent imaging and surgical decompression is indicated for this emergency. Neoplasms in the lumbosacral region can present with similar features.

Fig. 21.47 **C5/6 disc compressing cord: MR T2.**

Spinal stenosis

A narrow spinal canal is developmental and frequently symptomless but a congenital narrowing of the cervical canal predisposes the cervical cord to damage from minor disc protrusion later.

In the lumbosacral region, further narrowing of the canal by disc protrusion causes root pain, and/or buttock and lower limb claudication. As the patient walks, nerve roots become hyperaemic and swell, producing buttock and lower limb pain with numbness. Surgical decompression is required.

Neuralgic amyotrophy

This is a clearly recognizable entity – severe pain in muscles around one shoulder is followed by wasting, usually of infraspinatus, supraspinatus, deltoid and serratus anterior. A demyelinating brachial plexopathy develops over several days. The cause is unknown; an allergic or viral basis is postulated. Rarely, a similar condition develops in distal upper limb muscles or in a lower limb. Recovery of wasted muscles is usual, but not invariable, over some months.

Thoracic outlet syndrome

A fibrous band or cervical rib extending from the tip of the C7 transverse process towards the first rib compresses the lower brachial plexus roots, C8 and T1. There is forearm pain (ulnar border), T1 sensory loss and thenar muscle wasting, principally abductor pollicis brevis. Horner's syndrome may develop. The rib or band can be excised. Frequency of this diagnosis varies widely – thoracic outlet problems are sometimes invoked to explain ill-defined arm symptoms, typically on poor evidence.

A rib or band can also produce subclavian artery or venous occlusion. Neurological and vascular problems rarely occur together.

Malignant infiltration and radiation plexopathy

Metastatic disease of nerve roots, the brachial or lumbosacral plexus causes a painful radiculopathy and/or plexopathy. An example is apical bronchial carcinoma (Pancoast's tumour) causing a T1 and sympathetic outflow lesion – wasting of small hand muscles, pain and T1 sensory loss with ipsilateral Horner's syndrome. This also occurs in apical TB. In the upper limb, radiotherapy following breast cancer can produce a plexopathy.

Paraneoplastic Syndromes

Neurological disease may accompany malignancy in the absence of metastases. These paraneoplastic syndromes are associated with anti-neuronal antibodies, believed to be involved in generation of signs and symptoms. Numerous anti-neuronal antibodies have been described.

Clinical pictures include:

- sensorimotor neuropathy (p. 1173)
- mononeuritis multiplex (p. 1172)
- cranial polyneuropathy (p. 1173)
- Lambert–Eaton myasthenic–myopathic syndrome (LEMS, p. 1181)
- motor neurone disease variants (p. 1166)
- spastic paraparesis (p. 1112)
- cerebellar syndrome (p. 1114)
- limbic encephalitis, dementia and encephalopathy (p. 1167)
- progressive multifocal leucoencephalopathy (p. 108)
- paraneoplastic stiff person syndrome.

The neurological syndrome sometimes precedes evidence of the neoplasm – often a small-cell bronchial carcinoma, breast or ovarian cancer. Diagnosis is based on the clinical pattern and antibody profile. Neuroimaging is typically normal. Treatment is often unsatisfactory.

FURTHER READING

Darnell RB, Posner JB. Paraneoplastic syndromes involving the nervous system. *New England Journal of Medicine* 2003; **349**: 1543–1554.

MUSCLE DISEASES

Definitions

- *Myopathy* means a disease of voluntary muscle.
- *Myositis* indicates inflammation.
- *Muscular dystrophies* are inherited disorders of muscle cells.
- *Myasthenia* means fatiguable (worse on exercise) weakness – seen in neuro-muscular junction diseases.
- *Myotonia* is sustained contraction/slow relaxation.
- *Channelopathies* are ion channel disorders of muscle cells.

Weakness is the predominant feature of muscle disease. A selection of these conditions is mentioned here (Table 21.58).

Table 21.58 Muscle disease: classification

Acquired	Genetic
Inflammatory Polymyositis Dermatomyositis Inclusion body myositis Viral, bacterial and parasitic infection Sarcoidosis	**Dystrophic** Duchenne Facioscapulohumeral Limb girdle, and others
Endocrine and toxic Corticosteroids/Cushing's Thyroid disease Calcium disorders Hypokalaemia Ethanol Drugs	**Myotonic** Myotonic dystrophy Myotonia congenita
Myasthenic Myasthenia gravis Lambert–Eaton myasthenic–myopathic syndrome (LEMS)	**Channelopathies** Hypokalaemic periodic paralysis Hyperkalaemic periodic paralysis Normokalaemic periodic paralysis
	Metabolic (McArdle's syndrome) Myophosphorylase deficiency Other defects of glycogen and fatty acid metabolism Mitochondrial disease Malignant hyperpyrexia

Pathophysiology

Muscle fibres are affected by:

- acute inflammation and fibre necrosis (e.g. polymyositis, infection)
- genetically determined metabolic failure (e.g. Duchenne muscular dystrophy)
- infiltration by inflammatory tissue (e.g. sarcoidosis)
- fibre hypertrophy and regeneration
- mitochondrial diseases
- immunological damage
- ion channel disorders.

Examples of immunological damage, ion channel and mitochondrial disorders are:

- In *myasthenia gravis*, antibodies to the postsynaptic membrane acetylcholine receptor protein block neuromuscular transmission.
- In *Lambert–Eaton myasthenic–myopathic syndrome* (LEMS), antibodies to muscle calcium channel components are found.
- In *myotonias*, defective chloride ion membrane conductance is associated with delayed muscle relaxation.
- In *McArdle's syndrome*, myophosphorylase deficiency produces weakness after exercise.
- In *mitochondrial disease*, various ATP enzyme defects cause weakness.

Recognition

Diagnosis is often made from the distribution of weakness, wasting or hypertrophy, and muscle consistency (e.g. induration). Tendon reflexes are typically preserved.

Serum muscle enzymes

Serum creatine phosphokinase (CPK) is greatly elevated in many dystrophies, e.g. Duchenne, and in inflammatory muscle disorders, e.g. polymyositis. CPK remains normal in myasthenia gravis and LEMS, and usually remains normal in myotonias.

Neurogenetic tests

These are used in muscular dystrophies and mitochondrial disease.

Electromyography

The normal interference pattern is reduced. Changes in interference pattern indicate:

- *Myopathy*. Short-duration spiky polyphasic muscle action potentials are seen. Spontaneous fibrillation is occasionally recorded.
- *Myotonic discharges*. A characteristic high-frequency whine is heard on the loudspeaker.
- *Decrement and increment*. In myasthenia gravis, a characteristic decrement in evoked muscle action potential follows repetitive motor nerve stimulation. The reverse occurs, i.e. increment, following repetitive stimulation in LEMS, page 1181.
- In *denervation*, profuse fibrillation potentials (c. 1 ms duration, 50–200 mV in amplitude) are typical.

Muscle biopsy

Histology and muscle histochemistry of fibre types (type 1, slow; type 2, fast) indicates denervation, inflammation and

dystrophic changes. Electron microscopy is often valuable. In dystrophies and myositis, muscle fibres are diffusely abnormal, nuclei become central, and invasion by inflammatory cells and/or necrosis occurs; hypertrophic fibres are seen. In acute denervation, small angulated muscle fibres are scattered randomly, between normal fibres. In chronic partial denervation, fibre type grouping is seen, i.e. groups of atrophic fibres of the same type. Much experience is needed to interpret these changes.

Imaging

MRI shows *signal changes* within muscles in some cases of myositis.

INFLAMMATORY MYOPATHIES

Polymyositis

Polymyositis is characterized by non-suppurative skeletal muscle inflammation. The aetiology is unknown, though HLA-B8/DR3 phenotypes have a predisposition to develop polymyositis, and Coxsackie (and other) viruses are possibly implicated. It is described on page 547.

Differential diagnosis

Muscular dystrophies rarely progress as rapidly as polymyositis and there is no muscle pain. Pseudohypertrophy, seen in some dystrophies, does not occur in polymyositis. There is no family history. Motor neurone disease is always eventually accompanied by upper motor neurone signs, and prominent fasciculation is common.

Treatment

Steroids with azathioprine or cyclophosphamide reduce symptoms in about 75% of cases. Only rarely does polymyositis cause death.

Dermatomyositis

This is described on page 547.

Inclusion body myositis

This is described on page 548.

Viral, bacterial and parasitic myositis

Muscle tissue is involved in many infections (Table 21.59).

Myopathy in sarcoidosis and rheumatoid disease

In *sarcoidosis*, a subacute myopathy can develop, either with muscle swelling and induration or wasting. This is usually during pulmonary sarcoidosis, but occasionally it is isolated. Sarcoid nodules are seen on muscle biopsy. Treatment is with steroids.

In *rheumatoid disease*, a rare localized nodular myositis causes painful swelling of limb, trunk and facial muscles.

Table 21.59 Muscle involvement in infections

Myalgia	Influenza viruses
Bornholm disease	Coxsackie B5
Acute polymyositis, with myoglobinuria	Coxsackie B6, ECHO 9
Acute suppurative (tropical) myositis	*Staphylococcus* spp.
Gas gangrene (clostridial myositis)	*Clostridium perfringens*
Necrotizing myositis	*Streptococcus* spp.
Tuberculous myositis	*Mycobacterium tuberculosis*
Trichinosis	*Trichinella spiralis*
Cysticercosis	*Taenia solium* (larval form)
Hydatid disease	*Echinococcus granulosus*
Toxoplasma myositis	*Toxoplasma gondii*
Sarcosporidiosis	*Sarcocystis lindemanni*
Chagas' disease	*Trypanosoma cruzi*
Actinomycosis	*Actinomyces* spp.

METABOLIC AND ENDOCRINE MYOPATHIES

Corticosteroids and Cushing's syndrome

Proximal weakness occurs with prolonged high-dose steroid therapy, particularly with 9-α-fluorinated steroids, e.g. dexamethasone and triamcinolone, and in Cushing's syndrome (p. 1012). Selective type-2 fibre atrophy is seen on biopsy.

Thyroid disease (see also p. 988)

Several myopathies occur. Thyrotoxicosis can be accompanied by severe proximal myopathy. There is also an association between thyrotoxicosis and myasthenia gravis, and between thyrotoxicosis and hypokalaemic periodic paralysis (p. 1182). Both associations are seen more frequently in South East Asia. In ophthalmic Graves' disease, there is swelling and lymphocytic infiltration of extraocular muscles (p. 987).

Hypothyroidism is sometimes associated with muscle pain and stiffness, resembling myotonia. A proximal myopathy also occurs.

Disorders of calcium metabolism

Proximal myopathy develops in hypocalcaemia, rickets and osteomalacia (p. 567).

Hypokalaemia

Acute hypokalaemia (e.g. with diuretics) causes flaccid paralysis reversed by potassium, given slowly (p. 665). Chronic hypokalaemia leads to mild, mainly proximal, weakness. See also periodic paralysis (p. 1182).

Alcohol and drugs of abuse

Severe myopathy with muscle pain, necrosis and myoglobinuria occurs in acute excess. A subacute proximal myopathy occurs with chronic alcohol abuse. A similar syndrome occurs in diamorphine and amfetamine addicts.

Drugs

Drug-induced muscle disorders include proximal myopathy (steroids), muscle weakness (lithium), painful muscles (fibrates), rhabdomyolysis (fibrate combined with a statin) and malignant hyperpyrexia (p. 1182). All are rare, and most respond to drug withdrawal.

NEUROMUSCULAR JUNCTION CONDITIONS

Myasthenia gravis (MG)

MG is an acquired and probably heterogeneous condition. It is characterized by weakness and fatiguability of proximal limb, bulbar and ocular muscles, the latter sometimes in isolation. The heart is not affected. The prevalence is about 4 in 100 000. MG is twice as common in women as in men, with a peak age incidence around 30. The underlying cause is unknown.

Antibodies to acetylcholine receptor protein (anti-AChR antibodies) are commonly found. Immune complexes of anti-AChR IgG and complement are deposited at the postsynaptic membranes, causing interference with and later destruction of AChRs.

A second group of antibodies against muscle-specific receptor tyrosine kinase (anti-MuSK antibodies) has more recently been identified in anti-AChR antibody negative cases. Ocular muscle MG is another subgroup.

Thymic hyperplasia is found in 70% of MG patients below the age of 40. In some 10% a thymic tumour is found, the incidence increasing with age; antibodies to striated muscle can be demonstrated in some of these patients. Young patients without a thymoma have an increased association with HLA-B8 and DR3.

There is an association between MG and thyroid disease, rheumatoid disease, pernicious anaemia and SLE. Transient MG is sometimes caused by D-penicillamine treatment.

Clinical features

Fatiguability is typical. Limb muscles (proximal), extraocular, speech, facial expression and mastication muscles are commonly affected. Respiratory difficulties can be prominent. The clinical picture of fluctuating, fatiguable weakness is usually diagnostic. Muscle pain is typically absent. Early complaints of fatigue are frequently dismissed.

Complex extraocular palsies, ptosis and fluctuating proximal weakness are found. The reflexes are initially preserved but may be fatiguable, i.e. disappear following repetitive activity. Wasting is sometimes seen after many years.

Investigations

- *Serum anti-AChR and anti-MuSK antibodies.* Anti-AChR antibodies are present in some 80–90% of cases of generalized MG. These antibodies are not found in healthy controls but are seen rarely in other muscle disorders. In pure ocular MG, anti-AChR antibodies are detectable in less than 30% of cases.

  Anti-MuSK antibodies define a subgroup of MG patients characterized by weakness predominantly in bulbar, facial and neck muscles.
- *Repetitive nerve stimulation.* A characteristic decrement occurs in the evoked muscle action potential during repetitive stimulation. EMG is otherwise normal.
- *Tensilon (edrophonium) test.* This is seldom required and the drug is not available world-wide. Edrophonium 10 mg is given intravenously following a 1–2 mg test dose from the 10 mg vial. When the test is positive, there is substantial improvement in weakness within seconds and this lasts for up to 5 minutes. It is prudent to perform a control using saline and to have an observer. The sensitivity of the test is 80% but there are false negative and positive tests. Occasionally edrophonium (an anticholinesterase) causes bronchospasm and syncope. Resuscitation facilities must be available.
- *Imaging and other tests.* Mediastinal MR provides optimal structural imaging for thymoma. Routine blood studies are normal: the ESR is not raised, and CPK is normal. Antibodies to striated muscle suggest a thymoma; intrinsic factor and thyroid antibodies may be found. Rheumatoid factor and anti-nuclear antibody tests can be positive. Muscle biopsy is usually not performed, though ultrastructural neuromuscular junction abnormalities are well described.

Course and management

Myasthenia gravis fluctuates in severity; most cases have a protracted, lifelong course. Respiratory impairment, nasal regurgitation and dysphagia occur; emergency assisted ventilation may be required. Simple monitoring tests, such as the duration for which an arm can be held outstretched, and the vital capacity are useful.

Exacerbations are usually unpredictable and unprovoked but may be brought on by infections and by aminoglycosides. Magnesium sulphate enemas can provoke severe weakness.

Treatment

Oral anticholinesterases

Pyridostigmine (60 mg tablet) is widely used. The duration of action is 3–4 hours, the dose (usually 4–16 tablets daily) determined by response. Pyridostigmine prolongs acetylcholine action by inhibiting cholinesterase. Overdose of anticholinesterases causes severe weakness (cholinergic crisis). Muscarinic side-effects, e.g. colic and diarrhoea, are common; oral atropine (antimuscarinic) 0.5 mg helps reduce this. Anticholinesterases help weakness but do not alter the natural history of myasthenia.

Immunosuppressant drugs

These drugs are used in patients who do not respond to pyridostigmine or who relapse on treatment. Steroids are often used. There is improvement in 70%, although this may be preceded by an initial relapse. Azathioprine, mycophenolate and other immunosuppressants are also used.

Thymectomy

Thymectomy improves prognosis, more so in women than men below 50 years with positive AchR antibodies, even in patients without a thymoma. Cases positive for anti-MuSK antibodies tend not to improve following thymectomy. When a thymoma is present, the potential for malignancy also makes surgery necessary.

Plasmapheresis and intravenous immunoglobulin

During exacerbations these interventions are of value.

Other rare myasthenic syndromes exist, e.g. congenital myasthenia.

Lambert–Eaton myasthenic–myopathic syndrome (LEMS)

This paraneoplastic manifestation of small-cell bronchial carcinoma is due to defective acetylcholine release at the neuromuscular junction. Proximal limb muscle weakness, sometimes with ocular/bulbar muscles develops, with some absent tendon reflexes, a cardinal sign. Weakness tends to improve after a few minutes of muscular contraction, and absent reflexes return (cf. myasthenia). Diagnosis is confirmed by EMG and repetitive stimulation (increment, see above). Antibodies to P/Q-type voltage-gated calcium channels are found in most cases (90%). 3,4-Diaminopyridine (DAP) is a reasonably safe and sometimes effective treatment.

MUSCULAR DYSTROPHIES

These progressive genetically determined disorders of skeletal and sometimes cardiac muscle have a complex clinical and neurogenetic classification.

Duchenne muscular dystrophy (DMD) and Becker's muscular dystrophy

These are inherited as X-linked recessive disorders, though one-third of cases are spontaneous mutants. DMD occurs in 1 in 3000 male infants. The gene locus has been localized to the Xp21 region (X chromosome); there is absence of the gene product *dystrophin*, a rod-shaped cytoskeletal muscle protein in DMD. In Becker's dystrophy dystrophin levels are present but low. DMD is usually obvious by the fourth year, and often causes death by 20.

Dystrophin is essential for cell membrane stability. Deficiency leads to reduction in three glycoproteins (a-, β- and γ-sarcoglycans) in the dystrophin-associated protein complex (DAP-complex) that link dystrophin to laminin within cell membranes.

Becker's muscular dystrophy is less severe than Duchenne and weakness only becomes apparent in young adults.

Clinical features

A boy with DMD is noticed to have difficulty running and rising to his feet – he uses his hands to climb up his legs (Gowers' sign). There is initially a proximal limb weakness with calf pseudohypertrophy. The myocardium is affected. Severe disability is typical by the age of 10.

Investigations

The diagnosis is often suspected clinically. CPK is grossly elevated (100–200× normal). Biopsy shows variation in muscle fibre size, necrosis, regeneration and replacement by fat, and on immunochemical staining, absence of dystrophin. EMG shows a myopathic pattern.

Management

There is no curative treatment. Passive physiotherapy helps prevent contractures in the later stages. Portable respiratory support improves life expectancy.

Carrier detection. Females with an affected brother have a 50% chance of carrying the DMD gene. In carriers, 70% have a raised CPK, usually EMG abnormalities and/or changes on biopsy. Carrier and prenatal diagnosis (cDNA probes, co-inheritance with DMD locus) are available.

Genetic advice should be given. Many carriers choose not to have offspring. Determination of fetal sex and selective abortion of a male fetus is sometimes carried out.

Limb-girdle and facioscapulohumeral dystrophy

These less severe but disabling dystrophies are summarized in Table 21.60. There are many other varieties of dystrophy. Many are associated with a sarcoglycan deficiency. Genes for various limb-girdle muscular dystrophies (LGMD) have been located; some proteins coded are known, e.g. calpain. Type 1 LGMD is a rare autosomal dominant condition. Type 2 LGMD includes all recessive varieties, with complex subdivisions, e.g. LGMD2A (calpain III deficiency), LGMD2C (γ-sarcoglycan deficiency).

MYOTONIAS

Myotonias are characterized by continued, involuntary muscle contraction after cessation of voluntary effort. EMG is characteristic (p. 1121). The two most common myotonias are mentioned below. Since there is defective skeletal muscle cell chloride ion membrane conductance, myotonias are also classified as channelopathies (see below). Patients with myotonia tolerate general anaesthetics poorly.

Dystrophia myotonica (DM) or myotonic dystrophy (MD)

This autosomal dominant condition is a genetic disorder with two different triple repeat mutations. There is an expanded CTG repeat in the DM protein-kinase (*DMPK*) gene on chromosome 19q13.3 (DM1). The less common variety (DM2) is caused by expanded CCTG repeat in the zinc finger protein 9 (*ZNF9* gene). There is a correlation between disease severity, age at onset and approximate size of triple repeat mutations. There is progressive distal muscle weakness, with ptosis, weakness and thinning of the face and sternomastoids. Myotonia is typically present. Muscle disease is part of a syndrome comprising:

Table 21.60 Limb-girdle and facioscapulohumeral dystrophies

	Limb-girdle	Facioscapulohumeral
Inheritance	Autosomal, various	Dominant, usually
Onset	10–20 years	10–40 years
Muscles affected	Shoulders, pelvic girdle	Face, shoulders, pelvic girdle
Progress	Severe disability <25 years	Life expectancy normal, slow progress
Hypertrophy	Rare	Very rare

- cataracts
- frontal baldness
- cognitive impairment (mild)
- oesophageal dysfunction (and aspiration)
- cardiomyopathy and conduction defects (sudden death can occur in type 1)
- small pituitary fossa and hypogonadism
- glucose intolerance
- low serum IgG.

This gradually progressive condition usually becomes evident between 20 and 50 years. Phenytoin or procainamide sometimes helps the myotonia.

FURTHER READING

Cooper TA. A reversal of misfortune for myotonic dystrophy. *New England Journal of Medicine* 2006; **355**: 1825–1827.

Myotonia congenita (Thomsen's disease)

An isolated autosomal dominant myotonia, usually mild, becomes evident in childhood. The gene, *CLC1*, is on chromosome 7q35. The myotonia, which persists, is accentuated by rest and by cold. Diffuse muscle hypertrophy occurs – the patient has bulky muscles.

CHANNELOPATHIES

Many conditions are now included in this group, e.g. myotonias (see above) and Lambert–Eaton myasthenic–myopathic syndrome (p. 1181).

Hypokalaemic periodic paralysis

This channelopathy is characterized by generalized weakness, including bulbar muscles, that often starts after a heavy carbohydrate meal or following exertion. Attacks last for several hours. It often first comes to light in the teenage years and tends to remit after the age of 35. Serum potassium is usually below 3.0 mmol/L in an attack. The weakness responds to (slow) i.v. potassium chloride. It is usually an autosomal dominant trait caused by mutation in a muscle voltage-gated calcium channel gene (*CACLN1A3*). Other mutations in the sodium channel (*SCNA4* on chromosome 17q13.1–13.3) and potassium channel (*KCNE3*) also occur. Acetazolamide sometimes helps prevent attacks. Weakness can be caused by diuretics. A similar condition can also occur with thyrotoxicosis.

Hyperkalaemic periodic paralysis

This condition, also autosomal dominant, is characterized by attacks of weakness, sometimes with exercise. Attacks start in childhood and tend to remit after the age of 20; they last about 30–120 minutes. Myotonia may occur. Serum potassium is elevated. An attack can be terminated by i.v. calcium gluconate or chloride. There are point mutations in a muscle voltage-gated sodium channel gene (*SCN4A*, chromosome 17q13.1–13.3). Acetazolamide or a thiazide diuretic can be helpful.

A very rare normokalaemic, sodium-responsive periodic paralysis also occurs.

Stiff person syndrome

Stiff person syndrome (SPS) is a rare disease, commoner in females, of varying muscular stiffness with abnormal posturing and falls. Attacks of stiffness are sometimes provoked by noise or emotion, but sometimes occur spontaneously. Between attacks, which last from hours to days or even weeks, the patient may appear normal.

Widespread muscle stiffness is typical during an attack; there are no other neurological signs. SPS has been mistaken for Parkinson's, dystonia and non-organic conditions. Anti-glutamic acid decarboxylase antibodies (anti-GAD) are found in more than 50% of cases and are believed to be involved in the generation of muscle stiffness. Continuous motor activity is seen on EMG.

Features of SPS remain lifelong. Treatment with diazepam, other muscle relaxants and i.v. immunoglobulin can be helpful during attacks.

A form of SPS is also seen occasionally as a paraneoplastic condition associated with antibodies to the 128 kDa synaptic protein amphiphysin, and one case has been described with anti-gephyrin antibodies.

METABOLIC MYOPATHIES

This is a large group of rare, genetically determined conditions.

Myophosphorylase deficiency (McArdle's syndrome)

Lack of skeletal muscle myophosphorylase causes easy fatiguability and severe cramp on exercise, with myoglobinuria; McArdle's is an autosomal recessive disorder. Venous lactate during ischaemic exercise does not increase – the basis for a specific test. Sucrose ingestion may help (p. 1069).

Malignant hyperpyrexia

Widespread skeletal muscle rigidity with hyperpyrexia as a sequel of general anaesthesia or neuroleptic drugs, e.g. haloperidol, is due to a genetic defect in the sarcoplasmic reticulum calcium-release channel of the muscle ryanodine receptor, RyR1. Death during or following anaesthesia can occur in this rarity, sometimes inherited as an autosomal dominant. Dantrolene is of some help for rigidity.

Mitochondrial diseases

These comprise a complex group of disorders involving muscle, peripheral nerves and CNS, characterized by morphological and biochemical abnormalities in mitochondria. Mitochondrial DNA is inherited maternally (p. 37). The spectrum is wide, ranging from optic atrophy (see Leber's, p. 39) to myopathies, neuropathies and encephalopathy.

MELAS (mitochondrial encephalomyopathy, lactic acidosis, stroke-like episodes) is one well-recognized form.

Chronic progressive ophthalmoplegia (CPEO) is another.

MERRF describes myoclonic epilepsy with abnormal muscle histology, the muscle appearance being described as *ragged red* fibres.

Research into mitochondrial disease continues.

FURTHER READING

Ryan A, Matthews E, Hanna MG. Skeletal muscle channelopathies. *Current Opinion in Neurology* 2007; **20**: 558–563.

Shapira AHV. Mitochondrial disease. *Lancet* 2006; **368**: 70–82.

CHAPTER BIBLIOGRAPHY

Clarke C, Howard R, Rossor M, Shorvon S (eds) *Neurology: A Queen Square Textbook*. Oxford: Blackwell, 2009.

SIGNIFICANT WEBSITES

http://www.theabn.org *Association of British Neurologists information service*

http://jnnp.bmjjournals.com *Journal of Neurology, Neurosurgery, and Psychiatry*

http://www.epilepsynse.org.uk *UK National Society for Epilepsy*

UK Patient Support Group

Narcolepsy Association UK, 121 Kingsway, London WC2B 6PA (http://www.narcolepsy.org.uk).

UK National Charities

The MS Society: http://www.mssociety.org.uk

Meningitis Trust: http://www.meningitis-trust.org.uk

INFORMATION FOR PATIENTS

Knowing about meningitis and septicaemia (a leaflet for parents). Department of Health, PO Box 410, Wetherby LS23 7LN.

This page intentionally left blank

22 Psychological medicine

Introduction 1185
The psychiatric history 1186
The mental state examination 1186
Classification of psychiatric disorders 1189
Causes of a psychiatric disorder 1190
Psychiatric aspects of physical disease 1190
Severe behavioural disturbance 1191
The sick role and illness behaviour 1191
Functional or psychosomatic disorders: medically unexplained symptoms 1192
Somatoform disorders 1195
Dissociative/conversion disorders 1195
Sleep difficulties 1197
Mood (affective) disorders 1198
Depressive disorders 1198
Mania and hypomania 1204
Suicide and deliberate self-harm 1206
Anxiety disorders 1206
Obsessive-compulsive disorder 1210
Alcohol misuse and dependence 1211
Drug misuse and dependence 1214
Schizophrenia 1216
Organic mental disorders 1218
Eating disorders 1219
Sexual disorders 1221
Personality disorders 1222
Mental Health Act 1222
Mental Capacity Act 1222

INTRODUCTION

Psychiatry is concerned with the study and management of disorders of mental function: primarily thoughts, perceptions, emotions and purposeful behaviours. Psychological medicine, or liaison psychiatry, is the discipline within psychiatry that is concerned with psychiatric and psychological disorders in patients who have physical complaints or conditions. This chapter will primarily concern itself with this particular branch of psychiatry.

The long-held belief that diseases are either physical or psychological has been broken down by the accumulated evidence that the brain is functionally or anatomically abnormal in most if not all psychiatric disorders. Physical, psychological and social factors, and their interactions must be considered in order to understand psychiatric conditions. This philosophical change of approach rejects the Cartesian dualistic approach of the mind/body *medical model* and replaces it with the more integrated *biopsychosocial model*.

Epidemiology (Box 22.1)

The prevalence of psychiatric disorders in the community in the UK is about 20%, mainly composed of depressive and anxiety disorders and substance misuse (mainly alcohol). The prevalence is about twice as high in patients attending the general hospital, with the highest rates in the accident and emergency department and medical wards. The higher rates in the general hospital are due to several factors, such as admission for deliberate self-harm, a psychiatric disorder or treatment causing physical harm (e.g. alcohol-induced hepatitis), and physical presentations of psychiatric disorders (such as weight loss due to anorexia nervosa). Patients with medical conditions are also more, rather than less, likely to have psychiatric disorders and vice versa.

Culture and ethnicity

These can alter either the presentation or the prevalence of psychiatric ill-health. Biological factors in mental illness are usually similar across cultural boundaries, whereas psychological and social factors will vary. For example, the prevalence and presentation of schizophrenia vary little between countries, suggesting that biological/genetic factors are operating independently of cultural factors. In contrast, conditions in which social factors play a greater role vary between cultures, so that anorexia nervosa is found more often in developed cultures. Culture can also influence the presentation of illnesses, such that physical symptoms are more

Box 22.1 The approximate prevalence of psychiatric disorders in different populations

	% (approx.)
Community	20%
Neuroses	16%
Psychoses	0.5%
Alcohol misuse	5%
Drug misuse (total in community 20% due to comorbidity)	2% (an underestimate)
Primary care	25%
General hospital outpatients	30%
General hospital inpatients	40%

Box 22.2 The essentials of a safe psychiatric interview

- **Beforehand**: Ask someone of experience who knows the patient whether it is safe to interview the patient alone.
- **Access to others**: If in doubt, interview in the view or hearing of others, or accompanied by another member of staff.
- **Setting**: If safe; in a quiet room alone for confidentiality, not by the bed.
- **Seating**: Place yourself between the door and the patient.
- **Alarm**: If available, find out where the alarm is and how to use it.

common presentations of depressive illness in Asia than in Europe. Similarly culture will influence the healthcare sought for the same condition.

THE PSYCHIATRIC HISTORY

The purpose of the history is to help to make a diagnosis, determine possible aetiology, and estimate prognosis. Data may be taken from several sources, including interviewing the patient, a friend or relative (usually with the patient's permission), or the patient's general practitioner. The patient interview enables a doctor to establish a relationship with the patient and is the principal means to make a psychiatric diagnosis. Box 22.2 gives essential guidance on how to safely conduct such an interview, although it is unlikely that a patient will physically harm a healthcare professional. When interviewing a patient for the first time, follow the guidance outlined in Chapter 1 (see p. 11).

The history consists of:

- *Reason for referral* – a brief statement of why and how the patient came to the attention of the doctor.
- *Complaints* – as reported by the patient.
- *Present illness* – a detailed account of the illness from the earliest time at which a change was noted until the patient came to the attention of the doctor.
- *Past psychiatric history* – previous episodes of psychiatric illness and their treatments, including responses and adverse reactions. Always ask about previous episodes of self-harm.
- *Past medical history* – this should include emotional reactions to illness and procedures.
- *Family history* – focusing on the way the parents or carers cared (physically and emotionally) for the patient, and the occurrence of both mental and physical illnesses in first-degree relatives.
- *Personal (biographical) history* – a short biography that covers childhood difficulties including both abuse and neglect, educational problems (e.g. bullying and truanting), qualifications (to judge premorbid intelligence), job problems, sexual relationships, children, present housing, financial situation, bereavements and life stresses.
- *A reproductive history* in women should include menstrual problems, pregnancies, terminations, miscarriages, contraception and the menopause, if relevant.
- *Personality* – this may help to determine prognosis. The doctor should find out how other people would describe the patient. Is the patient generally a worrier, shy, introverted, dependent on others, passive, aggressive, irritable, over-emotional, prone to moodiness, conscientious, or perfectionist? These are all personality traits that may predict a poorer outcome in both medical and psychiatric disorders.
- *Drug history* – both prescribed and over-the-counter medication, the use (units per week) of alcohol, tobacco, caffeine, and illicit drugs.
- *Forensic history* – you should explain that you need to ask about this since ill-health can sometimes lead to problems with the law. Particularly note any violent or sexual offences. This is part of a *risk assessment* and is necessary in order to assess potential risks to those close to the patient as well as staff. Ask the patient what is the worst harm they have ever inflicted on someone else, which will give an indication of the potential for violence.
- *A systematic review* of physical symptoms is particularly necessary in patients complaining of physical symptoms.

THE MENTAL STATE EXAMINATION (MSE)

The history will already have assessed several aspects of the MSE, but the interviewer will need to expand several areas as well as test specific areas, such as cognition.

Appearance and general behaviour

State and colour of clothes, facial appearance, eye contact, posture and movement provide information about a patient's affect. Patients with psychomotor retardation due to a depressive illness sit with shoulders hunched, immobile, tearful, with a downcast gaze. Depressed individuals tend to wear clothes with dark colours. Agitation and anxiety cause an easy startle response, sweating, tremor, restlessness, fidgeting, visual scanning (for danger) and even pacing up and down. Patients with mania are often physically overactive and disinhibited, wearing colourful clothes. Someone who is actively hallucinating will seem distracted and suddenly stop talking or listening and stare intently at a particular place in the room.

Mood and affect

The patient has an *emotion* or *feeling*, tells the doctor about their *mood*, and the doctor observes the patient's *affect*. In psychiatric disorders, mood may be altered in three ways:

A persistent change in mood

- *Depression* is a lowering of mood, such as feeling sad, tearful, melancholic or low in spirits. Some patients report *anhedonia*, which is a lack of positive pleasure or loss of interest. Depression is the cardinal feature of depressive illness. Sometimes the word 'depression' is used as shorthand to describe a depressive illness. *Diurnal variation* in mood, feeling worse on waking,

www.studentconsult.com
Ask the authors

Introduction
The psychiatric history
The mental state examination

suggests a more severe illness, whereas a *reactive mood*, in which the patient can sometimes respond positively, indicates less severity.

- *Anxiety* is a feeling of constant, inappropriate or excessive worry, fear, apprehension, tension or inner restlessness, seen in anxiety and depressive disorders as well as drug withdrawal.
- *Elation* is a feeling of high spirits, exuberant happiness, vitality and even ecstasy, seen in mania and acute drug intoxication.
- *Irritability* can be either expressed (as in a temper or impatience) or an internal feeling of exasperation or anger, seen in both mania and depressive illness, especially in men.
- *Blunting* of affect is a total absence of emotion, seen most commonly in chronic schizophrenia.

Fluctuating or labile mood

This occurs when different emotions rapidly follow one another, so that a patient is crying one moment and laughing the next. This can occur in mixed affective states (see p. 1198). Alternatively, the patient is easily and excessively emotional over banal events or news, but the emotion is transient. This is seen both in a pseudobulbar palsy, commonly following a stroke (see p. 1111), and with mild depressive illnesses.

Incongruous mood

This occurs when emotional expression fails to match thoughts and actions. For example, a patient may laugh when describing the death of a close relative. This can occur in schizophrenia.

Speech

Disorders of thinking are usually recognized from the patient's speech.

Disorders of the stream of thought

These are abnormalities in the amount and speed of the thoughts experienced.

Pressure of speech occurs in mania and can be recognized by loudness, rapidity, and difficulty in interrupting speech.

Poverty of speech is the opposite experience, when there appears to be an absence of any thought and patients report their minds to be empty. It occurs in depressive illness.

Thought block occurs in schizophrenia. There is an abrupt and complete interruption of the stream of thought, so the mind goes blank. Patients may interpret the experience in an unusual way (e.g. thought withdrawal; see below).

Disorders of the form of thought

Flight of ideas. The patient's thoughts rapidly jump from one topic to another, such that one train of thought is not completed before another appears. It is often produced by clang associations (the use of two or more words with a similar sound: 'sun, son, song'), punning, rhyming, and responding to distracting cues in the immediate surroundings. Flight of ideas is characteristic of mania and often accompanies pressure of speech.

Perseveration is the persistent and inappropriate repetition of the same thoughts or actions. It occurs in frontal lobe disorders.

Thought broadcast is when the patient experiences their thoughts as being understood by others without talking, as though their thoughts are literally being broadcast to all around them.

Thought insertion occurs when a patient's thought is perceived as being planted in their mind by someone else.

Thought withdrawal occurs when a patient experiences their thoughts being taken away from them, without their control.

Thought content

Thought content refers to the worries and preoccupations expressed by the patient. Abnormal beliefs are, of course, part of the thought content, but are regarded as sufficient to be discussed separately (see below).

- *An obsessional rumination* is a recurrent, persistent thought, impulse, image or musical theme that occurs despite the patient's effort to resist it. The patient recognizes that the obsessional thought is their own, but it is usually unpleasant and often 'out of character'. Common obsessions concern dirt, contamination and orderliness.
- *A compulsion* is a repetitive and seemingly purposeful action performed in a stereotypical way, referred to as a compulsive ritual. Compulsions are accompanied by a subjective sense that they must be carried out (or the patient will be overwhelmed by either anxiety or a superstitious belief that something bad will occur) and by an urge to resist. Compulsive rituals are used to counteract ruminations, so patients repetitively wash their hands to diminish the fear of contamination with dirt. Obsessions and compulsions are characteristic of obsessive compulsive disorder (see below).

Insight and illness beliefs

Insight is the degree to which a person recognizes that he or she is unwell, and is minimal in patients with a psychosis. *Illness beliefs* are the patient's own explanations of their ill-health, including diagnosis and causes. These beliefs should be elicited because they can help to determine prognosis and adherence with treatment, with any disease.

Abnormal beliefs

The main form of abnormal belief is the delusion (Box 22.3). Delusions can be primary or secondary.

- *Primary delusions* are rare and appear suddenly and with full conviction but without any preceding event. For example, a patient on being offered a glass of wine suddenly believes that this means that he is Jesus Christ.

Box 22.3 Delusion

Delusion is defined as an abnormal belief that is:

- held with absolute conviction
- not amenable to reason or modifiable by experience
- not shared by those of a common cultural or social background
- experienced as a self-evident truth of great personal significance
- usually false.

- *Secondary delusions* are derived from a preceding morbid experience, such as an auditory hallucination.

Delusions are also classified according to their content, and include delusions of persecution, self-reference, guilt, worthlessness, nihilism, religion, grandeur, jealousy or control.

Feelings, thoughts or actions may also be interpreted by the patient as being under the control of some external power. Such passivity experiences are first rank symptoms and are regarded as diagnostic of schizophrenia (see p. 1216). Patients may develop secondary delusions that explain this alien control, such as witchcraft, hypnosis, radio waves or television – so-called 'delusions of passivity'.

Delusions should be distinguished from *overvalued ideas* – deeply held personal convictions that are understandable when the individual's background is known.

Ideas of self-reference that fall short of delusions are held by people who are particularly self-conscious. Such individuals cannot help feeling that people take particular notice of them in public places, laugh at them or pass comment about them. Such a feeling is not delusional in that patients realize that the feeling originates within themselves and that they are no more noticeable than anyone else, but nevertheless they cannot dismiss the feeling.

Abnormal perceptions

- *Illusions* are misperceptions of external stimuli and are most likely to occur when the general level of sensory stimulation is reduced.
- *Hallucinations* are defined in Box 22.4. Healthy people occasionally experience hallucinations, such as in normal grief, or during the transition between sleeping and waking (hypnogogic and hypnopompic). Hallucinations may affect any of the perceptions: auditory, visual, tactile, gustatory, olfactory or of deep sensation.
- *Pseudohallucinations* are usually auditory, and are either true externally sited hallucinations, but with insight into their imaginary nature, or are sited within internal space (e.g. 'I heard a voice in my head speaking to me'). They can occur in mood disorders and do not indicate a psychosis.
- *Depersonalization* is a change in self-awareness such that the person feels unreal or detached from their body. The individual is aware, however, of the subjective nature of this alteration.
- *Derealization* is the unpleasant feeling that the external environment has become unreal and/or remote; patients may describe themselves as though they are in a dream-like state. Both this and depersonalization can occur in healthy people when they are tired, after sensory deprivation and when using hallucinogenic drugs. They also occur in anxiety disorders, schizophrenia and temporal lobe epilepsy.
- *Increased sensitivity of perceptions*, such as photosensitivity and phonosensitivity, occurs in anxiety disorders (e.g. increased sensitivity to the neon strip-lights and noise in a supermarket in agoraphobia) as well as migraine.

Box 22.4 Hallucination

An hallucination is defined as a perception in the absence of a stimulus. It is:

- a false perception and not a distortion
- perceived as inhabiting objective space
- perceived as having qualities of normal perception
- perceived alongside normal perceptions
- independent of the individual's will.

Cognitive state

Examination of the cognitive state is necessary to diagnose organic brain disorders, such as delirium and dementia. Poor concentration, confusion and memory problems are the most common subjective complaints. Clinical testing involves the screening of cognitive functions, which may suggest the need for more formal psychometry. A premorbid estimate of intelligence, necessary to judge changes in cognitive abilities, can be made from asking the patient the final year level of education and the highest qualifications or skills achieved.

Testing can be divided into tests of diffuse and focal brain functions.

Diffuse functions

Orientation in time, place and person. Consciousness can be defined as the awareness of the self and the environment. Clouding of consciousness is more accurately a fluctuating level of awareness and is commonly seen in delirium.

Attention is tested by saying the months or days backwards.

Verbal memory. Ask the patient to repeat a name and address with 10 or so items, noting how many times it takes to recall it 100% accurately (normal is 1 or 2) (immediate recall or registration).

Ask the patient to try to remember it and then ask it of them again after 5 minutes (0 or 1 error is normal) (short-term memory).

Long-term memory. Ask the patient to recall the news of that morning or recently. If they are not interested in the news, find out their interests and ask relevant questions (about their football team or favourite soap opera). *Amnesia* is literally an absence of memory and *dysmnesia* indicates a dysfunctioning memory.

Focal functions

Frontal, temporal and parietal function tests are covered on page 1099. Note any disinhibited behaviour not explained by another psychiatric illness. *Sequential tasks* are tested by asking the patient to alternate making a fist with one hand at the same time as a flat hand with the other. Ask the patient to tap a table once if you tap twice and vice versa. Note any motor perseveration whereby the patient cannot change the movement once established. Observe for verbal perseveration, in which the patient repeats the same answer as given previously for a different question. *Abstract thinking* is measured by asking the meaning of common proverbs, a literal meaning suggesting frontal lobe dysfunction, assuming reasonable premorbid intelligence.

Mini-mental state examination

Box 22.5 gives the 'mini-mental state' examination of cognitive function. This is a 5-minute bedside test that is useful as a screen and in assessing the degree of cognitive dysfunc-

Box 22.5 The mini-mental state examination

Orientation	
Score one point for each correct answer:	
What is the: time, date, day, month, year?	
Maximum:	5 points
What is the name of: this ward, hospital, district, town, country?	5 points
Registration	
Name three objects only once. Score up to a maximum of 3 points for each correct repetition.	3 points
Repeat the objects until the patient can repeat them accurately (in order to test recall later).	
Attention and calculation	
Ask the patient to subtract 7 from 100 and then 7 from the result four more times. Score 1 point for each correct subtraction.	5 points
Recall	
Ask the patient to repeat the names of the three objects learnt in the registration test.	3 points
Language	
Score 1 point for each of two simple objects named (e.g. pen and a watch).	2 points
Score 1 point for an accurate repetition of the phrase: 'No ifs, ands or buts'.	1 point
Give a 3-stage command, scoring 1 point for each part correctly carried out; e.g. 'With the index finger of your right hand touch your nose and then your left ear'.	3 points
Write 'Close your eyes' on a blank piece of paper and ask the patient to follow the written command. Score 1 point if the patient closes the eyes.	1 point
Ask the patient to write a sentence. Score 1 point if the sentence is sensible and contains a noun and a verb.	1 point
Draw a pair of intersecting pentagons with each side approximately 1 inch long. Score 1 point if it is correctly copied	1 point
TOTAL MAXIMUM SCORE	30 POINTS

From: Folstein MF, Folstein SE, McHough PR. 'Mini-mental state': a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* 1975; **12**: 189–198

tion in patients with diffuse brain disorders. It correlates well with more time-consuming intelligence quotient (IQ) tests, but it will not as easily pick up cognitive problems caused by focal brain lesions. A score of 23 or less out of a maximum score of 30 will pick up about 90% of patients with cognitive impairments, with about 10% false positives.

Defence mechanisms

Although not strictly part of the mental state examination, it is useful to be able to identify psychological defences in ourselves and our patients. Defence mechanisms are mental processes that are usually unconscious. The defence mechanisms described below are among the most commonly used and are useful in understanding many aspects of behaviour.

- *Denial* is similar to repression and occurs when patients behave as though unaware of something that they might be expected to know. One example would be a patient who, despite being told that a close relative has died, continues to behave as though the relative were still alive.
- *Displacement* involves the transferring of emotion from a situation or object with which it is properly associated to another that gives less distress.
- *Identification* refers to the unconscious process of taking on some of the characteristics or behaviours of another person, often to reduce the pain of separation or loss.
- *Projection* involves the attribution to another person of thoughts or feelings that are in fact one's own.
- *Regression* is the adoption of primitive patterns of behaviour appropriate to an earlier stage of development. It can be seen in ill people who become child-like and highly dependent.
- *Repression* is the exclusion from awareness of memories, emotions and/or impulses that would cause anxiety or distress if allowed to enter consciousness.
- *Sublimation* refers to the unconscious diversion of unacceptable behaviours into acceptable ones.

The relevant physical examination

This should be guided by the history and mental state examination. Particular attention should usually be paid to the neurological and endocrinological examinations when organic brain syndromes and affective illnesses are suspected.

Summary or formulation

When the full history and mental state have been assessed, the doctor should make a concise assessment of the case, which is termed a formulation. In addition to summarizing the essential features of the history and examination, the formulation includes a differential diagnosis, a discussion of possible causal factors, and an outline of further investigations or interviews needed. It concludes with a concise plan of treatment and a statement of the likely prognosis.

CLASSIFICATION OF PSYCHIATRIC DISORDERS

The classification of psychiatric disorders into categories is mainly based on symptoms and behaviours, since there are currently few diagnostic tests for psychiatric disorders. There currently exists an unhelpful dualistic division of psychiatric disorders from neurological diseases, since the pathologies of at least the majority of each group of conditions are located in the brain, such as Alzheimer's disease causing dementia and a pseudobulbar palsy causing emotional lability.

Psychiatric classifications have traditionally divided up disorders into neuroses and psychoses.

Neuroses are illnesses in which symptoms vary only in severity from normal experiences, such as depressive illness. *Psychoses* are illnesses in which symptoms are qualitatively different from normal experience, with little insight into their nature, such as schizophrenia. There are several problems with a neurotic–psychotic dichotomy. First, neuroses may be as severe in their effects on the patient and their family as psychoses. Second, neuroses may cause symptoms that fulfil the definition of psychotic symptoms. For instance,

Table 22.1 International classification of psychiatric disorders (ICD-10)

Organic disorders
Mental and behavioural disorders due to psychoactive substance use
Schizophrenia and delusional disorders
Mood (affective) disorders
Neurotic, stress-related and somatoform disorders
Behavioural syndromes
Disorders of adult personality and behaviour
Mental retardation

World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders.* Geneva: World Health Organization, 1992

someone with anorexia nervosa may be convinced that they are fat when they are thin, and this belief would meet all the criteria for a delusional belief. Yet we would traditionally classify the illness as a neurosis.

The International Classification of Mental and Behavioural Disorders (ICD-10) has been published by the World Health Organization. This system has largely abandoned the traditional division between neurosis and psychosis, although the terms are still used. The disorders are now arranged in groups according to major common themes (e.g. mood disorders and delusional disorders). A classification of psychiatric disorders derived from ICD-10 is shown in Table 22.1, and this is the classification mainly used in this chapter.

The *Diagnostic and Statistical Manual of the American Psychiatric Association (DSM-IV-TR)* is an alternative classification system.

FURTHER READING

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition Text Revision (DSM-IV-TR).* Washington, DC: APA, 2000.

World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines.* Geneva: World Health Organization, 1992.

CAUSES OF A PSYCHIATRIC DISORDER

A psychiatric disorder may result from several causes which may interact. It is most helpful to divide causes into the three 'P's: predisposing, precipitating and perpetuating factors.

- *Predisposing factors* often stem from early life and include genetic, pregnancy and delivery, previous emotional traumas and personality factors.
- *Precipitating (triggering) factors* may be physical, psychological or social in nature. Whether they produce a disorder depends on their nature, severity and the presence of predisposing factors. For instance a death of a close, rather than distant, family member is more likely to precipitate a pathological grief reaction in someone who has not come to terms with a previous bereavement.
- *Perpetuating (maintaining) factors* prolong the course of a disorder after it has occurred. Again they may be physical, psychological or social, and several are often active and interacting at the same time. For example, high levels of criticism at home combined with taking cannabis, as relief from the criticism, may help to maintain schizophrenia.

PSYCHIATRIC ASPECTS OF PHYSICAL DISEASE

Patients with non-psychiatric, 'physical' diseases are more likely to suffer from psychiatric disorders than those who are well. The most common psychiatric disorders in physically unwell patients are mood or adjustment disorders and acute organic brain disorders (delirium). The relationship between psychological and physical symptoms may be understood in one of four ways:

- Psychological distress and disorders can precipitate physical diseases (e.g. anorexia nervosa causing cardiac arrhythmias, due to hypokalaemia).
- Physical diseases and their treatments can cause psychological symptoms or ill-health (Table 22.2).
- Both the psychological and physical symptoms are caused by a common disease process (e.g. Huntington's chorea).
- Physical and psychological symptoms and disorders may be independently co-morbid, particularly in the elderly.

Factors that increase the risk of a psychiatric disorder in someone with a physical disease are shown in Table 22.3.

Differences in treatment

Although the basic principles are the same as in treating psychiatric illnesses in the physically healthy, there are some differences:

- *Uncertainty* regarding the physical diagnosis or prognosis, with its attendant tendency to imagine the worst, is often a triggering or maintaining factor, particularly in an adjustment or mood disorder. Good two-way communication between doctor and patient, with time taken to listen to the patient's concerns, is often the most effective 'antidepressant' available.
- The history may reveal the role of a *physical disease* or *treatment* exacerbating the psychiatric condition, which should then be addressed (see Table 22.2). For example, the dopamine agonist bromocriptine can precipitate a psychosis.
- When prescribing psychotropic drugs, the dose should be reduced in disorders affecting *pharmacokinetics*, e.g. fluoxetine in renal or hepatic failure.

Table 22.2 Psychiatric conditions sometimes caused by physical diseases

Psychiatric disorders/symptom	Physical disease
Depressive illness	Hypothyroidism Cushing's syndrome Steroid treatment Brain tumour
Anxiety disorder	Thyrotoxicosis Hypoglycaemia (transient) Phaeochromocytoma Complex partial seizures (transient) Alcohol withdrawal
Irritability	Post-concussion syndrome Frontal lobe syndrome Hypoglycaemia (transient)
Memory problem	Brain tumour Hypothyroidism
Altered behaviour	Acute drug intoxication Post-ictal state Acute delirium Dementia Brain tumour

Classification of psychiatric disorders
Causes of a psychiatric disorder
Psychiatric aspects of physical disease
Severe behavioural disturbance
The sick role and illness behaviour

Table 22.3 Factors increasing the risk of psychiatric disorders in the general hospital

Patient factors
Previous psychiatric history
Current social or interpersonal stresses
Homelessness
Recent alcohol misuse
Setting
A&E department
Neurology, oncology and endocrinology wards
Intensive care unit
Renal dialysis unit
Physical conditions
Chronic ill-health
Chronic pain
Life-threatening illness
Recent bad prognostic news
Disabling condition
Brain disease
Recent live birth, stillbirth or miscarriage
Functional illness
Treatment
Certain drugs (e.g. dopamine agonists)
Second postoperative day
Surgery affecting body image (e.g. emergency stomata)

- *Drug interactions* should be of particular concern, e.g. lithium and non-steroidal anti-inflammatory drugs.
- Sometimes a physical treatment may be planned that may exacerbate the psychiatric condition. An example would be high-dose steroids as part of chemotherapy in a patient with leukaemia and depressive illness.
- Always consider the risk of *suicide* in an inpatient with a mood disorder and take steps to reduce that risk; for example, moving the patient to a room on the ground floor and/or having a registered mental health nurse attend the patient while at risk.

SEVERE BEHAVIOURAL DISTURBANCE

Patients with aggressive or violent behaviour cause understandable apprehension in all staff, and are most commonly seen in the accident and emergency department. Information from anyone accompanying the patient, including police or carers, can help considerably. Box 22.6 gives the main causes of disturbed behaviour.

Management of the severely disturbed patient

The primary aims of management are control of dangerous behaviour and establishment of a provisional diagnosis. Three specific strategies may be necessary when dealing with the violent patient:

Box 22.6 Main causes of disturbed behaviour

- Drug intoxication (especially alcohol)
- Delirium (acute confusional state)
- Acute psychosis
- Personality disorder

- reassurance and explanation
- physical restraint
- medication.

The majority of disturbed patients are themselves frightened, as well as frightening, and may feel threatened by those around them, misinterpreting the actions of others. Staff should always explain the situation and their intentions. This simple strategy may calm a patient sufficiently to be interviewed and allow an appropriate examination.

If the behaviour remains severely disturbed, it may be necessary to restrain patients from harming themselves or others. If planned, this should be done with sufficient numbers of trained staff; at least one person per limb and another two in charge or delivering medication. Once brought under physical control, the patient should be held in the prone position, in order to protect the airway and allow access for intramuscular medication. Care must be taken to ensure that neither the airway nor breathing is impeded, by having someone always present at the head of the patient.

It is usually necessary to administer medication while the patient is restrained, and they should not be released until they are visibly calmed. Management depends on the provisional diagnosis. 'Rapid tranquillization' should be employed when the patient has a psychosis, so long as the Mental Health Act has been used (see p. 1222) or the situation is so dangerous that the doctor is acting under 'common law'. A deep intramuscular injection of zuclopenthixol acetate in a dose ranging from 50 to 150 mg together with 0.5–1 mg lorazepam is now the treatment of choice for rapid effect. Used in combination, they have a synergistic action. Moderate doses of a neuroleptic or benzodiazepine can also be given at regular, comparatively short intervals (30–60 minutes) intramuscularly. Another alternative regimen is intramuscular butyrophenone (haloperidol 5–10 mg) in patients under 60 years old. This dose should be reduced in the elderly and those with known cardiac or hepatic disease. The patient should be observed for up to 1 hour before a further dose is administered. Breathing, pulse rate and blood pressure should be monitored for respiratory difficulty, arrhythmias and hypotension. In the case of continuing disturbance, it may be preferable to administer an adjunctive intramuscular benzodiazepine (lorazepam 2 mg) rather than a further dose of a neuroleptic. Both neuroleptic drugs and benzodiazepines may be used as tranquillizers.

THE SICK ROLE AND ILLNESS BEHAVIOUR

The *sick role* describes behaviour usually adopted by ill people. Such people are not expected to fulfil their normal social obligations. They are treated with sympathy by others and are only obliged to see their doctor and take medical advice or treatments.

Illness behaviour is the way in which given symptoms may be differentially perceived, evaluated and acted (or not acted) upon by different kinds of persons. We all have illness behaviour when we choose what to do about a symptom. Going to see a doctor is generally more likely with more severe, distressing and numerous symptoms. It is also more likely in introspective individuals who focus on their health.

Abnormal illness behaviour occurs when there is a discrepancy between the objective somatic pathology present and the patient's response to it, in spite of adequate medical investigation and explanation.

FUNCTIONAL OR PSYCHOSOMATIC DISORDERS: MEDICALLY UNEXPLAINED SYMPTOMS

So-called '*functional*' (in contrast to 'organic') disorders are illnesses in which there is no obvious pathology or anatomical change in an organ and there is a presumed dysfunction in an organ or system. The psychiatric classification of these disorders would be *somatoform disorders*, but they do not fit easily within either medical or psychiatric classification systems, since they occupy the hinterland between them. This classification also implies a dualistic 'mind or body' dichotomy, which is not supported by neuroscience. Since current classifications still support this outmoded understanding, this chapter will address these conditions in this way.

The word *psychosomatic* has had several meanings, including psychogenic, 'all in the mind', imaginary and malingering. The modern meaning is that psychosomatic disorders are syndromes of unknown aetiology in which both physical and psychological factors are likely to be causative. Medically unexplained symptoms and syndromes are very common in both primary care and the general hospital (over half the outpatients in gastroenterology and neurology clinics have these syndromes). Because orthodox medicine has not been particularly effective in treating or understanding these disorders, many patients perceive their doctors as unsympathetic and seek out complementary or even alternative treatments of uncertain efficacy. Examples of functional disorders are shown in Table 22.4.

Because epidemiological studies suggest that having one of these syndromes significantly increases the risk of having another, some doctors believe that these syndromes represent different manifestations of 'one functional syndrome', which is indicative of a somatization process. Functional disorders also have a significant association with psychiatric disorders, especially depressive and anxiety disorders. Against this view is the evidence that the majority of primary care patients with most of these disorders do not have either a psychiatric disorder or another functional disorder. It also seems that it requires a major stress or the development of a co-morbid psychiatric disorder in order for such sufferers to see their doctor, which might explain why doctors are so impressed with the associations with both stress and psychiatric disorders. Doctors have historically tended to diagnose 'stress' or 'psychosomatic disorders' in patients with symptoms that they cannot explain. History is full of such disorders being reclassified as research clarifies the pathology. An example is writer's cramp (p. 1150) which most neurologists now agree is a dystonia rather than a neurosis.

Table 22.4 'Functional' syndromes (medically unexplained symptoms)

'Tension' headaches
Atypical facial pain
Atypical chest pain
Fibromyalgia (chronic widespread pain)
Other chronic regional pain syndromes
Chronic fatigue syndrome
Multiple chemical sensitivity
Premenstrual syndrome
Irritable or functional bowel syndrome
Irritable bladder

The likelihood is that these so-called 'functional somatoform disorders' will be reclassified as their causes and pathophysiology are revealed. Functional brain scans suggest enhancement of brain activity during interoception in more than one syndrome. Interoception is the perception of internal (visceral) phenomena, such as a rapid heartbeat.

Chronic fatigue syndrome (CFS)

There has probably been more controversy over the existence and cause of CFS than any other 'functional' syndrome in recent years. This is reflected in its uncertain classification as *neurasthenia* in the psychiatric classification and *myalgic encephalomyelitis (ME)* under neurological diseases. There is now good evidence for the independent existence of this syndrome, although the diagnosis is made clinically and by exclusion of other fatiguing disorders. Its prevalence is 0.5% in the UK, but it is found world-wide, while varying in prevalence. The symptom of abnormal fatigue occurs in 10–20%. It occurs most commonly in women between the ages of 20 and 50 years, particularly in ethnic minority groups in the UK. The cardinal symptom is chronic fatigue made worse by minimal exertion. The fatigue is usually both physical and mental, associated most commonly with poor concentration, impaired registration of memory, alteration in sleep pattern (either insomnia or hypersomnia), and muscular pain. Mood disorders are present in a large minority of patients, and can cause problems in diagnosis because of the large overlap in symptoms. These mood disorders may be secondary, independent (co-morbid), or primary (with a misdiagnosis of CFS).

Aetiology

Functional disorders often have some aetiological factors in common with each other (see Table 22.5), as well as more

Table 22.5 Aetiological factors commonly seen in 'functional' disorders

Predisposing
Perfectionist, obsessional and introspective personality traits
Childhood traumas (physical and sexual abuse)
Similar illnesses in first-degree relatives
Precipitating (triggering)
Infections (chronic fatigue and irritable bowel syndromes)
Traumatic events (especially accidents)
Physical injuries ('fibromyalgia' and other chronic pain syndromes)
Life events that precipitate changed behaviours (e.g. going off sick)
Incidents where the patient believes others are responsible
Perpetuating (maintaining)
Inactivity with consequent physiological adaptation (CFS and 'fibromyalgia')
Avoidant behaviours – multiple chemical sensitivities, CFS
Maladaptive illness beliefs (that maintain maladaptive behaviours) (CFS, MCS)
Excessive dietary restrictions ('food allergies')
Stimulant drugs (e.g. caffeine)
Sleep disturbance
Mood disorders
Somatization disorder
Unresolved anger or guilt
Unresolved compensation claim

specific aetiologies. For instance, CFS can be triggered by certain infections, such as infectious mononucleosis and viral hepatitis. About 10% of patients with infectious mononucleosis have CFS 6 months after the onset of infection, yet there is no evidence of persistent infection in these patients. Those fatigue states which clearly do follow on a viral infection can also be classified as *post-viral fatigue syndromes*. Other aetiological factors are uncertain. Immune and endocrine abnormalities noted in CFS may be secondary to the inactivity or sleep disturbance commonly seen. The role of stress is uncertain, with some indication that the influence of stress is mediated through consequent psychiatric disorders exacerbating fatigue, rather than any direct effect.

Management

The general principles of the management of functional disorders are given in Box 22.7. Specific management of CFS should include a mutually agreed and supervised programme of gradually increasing activity. However, only a quarter of patients recover after treatment. It is sometimes difficult to persuade a patient to accept what are inappropriately perceived as 'psychological therapies' for such a physically manifested condition. Antidepressants do not work in the absence of a mood disorder or insomnia.

Prognosis

This is poor without treatment, with less than 10% of hospital attenders recovered after 1 year. Outcomes are worse with greater severity, increasing age, co-morbid mood disorders, and the conviction that the illness is entirely physical. With rehabilitative treatments some 60% make a significant improvement.

Fibromyalgia (chronic widespread pain: CWP)

This controversial condition of unknown aetiology overlaps with chronic fatigue syndrome, with both conditions causing fatigue and sleep disturbance (see p. 1192). Diffuse muscle and joint pains are more constant and severe in CWP, although the 'tender points', previously considered to be pathognomonic, are now known to be ubiquitous, associated with psychological distress, and of no diagnostic importance (p. 516). CWP occurs most commonly in women aged 40–65 years, with a prevalence in the community of between 1 and 11%. There are associations with depressive and anxiety disorders, other functional disorders, physical deconditioning and a possibly characteristic sleep disturbance (see Table 22.5). Functional brain scans suggest that patients actually perceive greater pain, supporting the idea of abnormal sensory processing.

Management

Apart from the general principles in Box 22.7, management also consists of symptomatic analgesia, reversing the sleep disturbance, and a physically orientated rehabilitation programme. A meta-analysis suggests that tricyclic antidepressants that inhibit reuptake of both serotonin (5-hydroxytryptamine – 5-HT) and norepinephrine (noradrenaline) (e.g. amitriptyline, dosulepin) have the greatest effect on sleep, fatigue and pain. The doses used are too low for antidepressant effect and the drugs may work through their hypnotic and analgesic effects.

Box 22.7 Management of functional disorders

The first principle is the identification and treatment of maintaining factors (e.g. dysfunctional beliefs and behaviours, mood and sleep disorders).

- Communication
  - Explanation of ill-health, including diagnosis and causes
  - Education about management (including self-help leaflets)
- Stopping drugs (e.g. caffeine causing insomnia, analgesics causing dependence)
- Rehabilitative therapies
  - Cognitive behaviour therapy (to challenge unhelpful beliefs and change coping strategies)
  - Supervised and graded exercise therapy for approximately 3 months (to reduce inactivity and improve fitness)
- Pharmacotherapies
  - Specific antidepressants for mood disorders, analgesia and sleep disturbance (e.g. 10–50 mg of amitriptyline at night for sleep and pain)
  - Symptomatic medicines (e.g. appropriate analgesia, taken only when necessary)

Other chronic pain syndromes

A chronic pain syndrome is a condition of chronic disabling pain for which no medical cause can be found. The psychiatric classification would be a persistent *somatoform pain disorder*, but this is unsatisfactory since the criteria include the stipulation that emotional factors must be the main cause, and it is clinically difficult to be that certain. The main sites of chronic pain syndromes are the head, face, neck, lower back, abdomen, genitalia or all over (CWP: fibromyalgia). 'Functional' low back pain is the commonest 'physical' reason for being off sick long-term in the UK (p. 509). Quite often a minor abnormality will be found on investigation (such as mild cervical spondylosis on the neck X-ray), but this will not be severe enough to explain the severity of the pain and resultant disability. These pains are often unremitting and respond poorly to analgesics. Sleep disturbance is almost universal and co-morbid psychiatric disorders are commonly found.

Aetiology

The perception of pain involves sensory (nociceptive), emotional and cognitive processing in the brain. Functional brain scans suggest that the brain responds abnormally to pain in these conditions, with increased activation in response to chronic pain. This could be related to conditioned behavioural and physiological responses to the initial acute pain. The brain may then adapt to the prolonged stimulus of the pain by changing its central processing. The prefrontal cortex, thalamus and cingulate gyrus seem to be particularly affected and some of these areas are involved in the emotional appreciation of pain in general. Thus it is possible to start to understand how beliefs, emotions and behaviours might influence the perception of chronic pain (see Table 22.5).

Management

Management involves the same principles as used in other functional syndromes (Box 22.7). Since analgesics are rarely effective, and can cause long-term harm, patients should be encouraged to gradually reduce their use. It is often helpful to involve the patient's immediate family or partner, to ensure

that the partner is also supported and not unconsciously discouraging progress.

Specific drug treatments are few. Nerve blocks are not usually effective. Anticonvulsants such as carbamazepine and gabapentin may be given a therapeutic trial if the pain is thought to be neuropathic (see p. 1118). The antidepressant *dosulepin* is an effective treatment in half of patients with atypical facial pain, and this effect seems to be independent of dosulepin's effect on mood. Another tricyclic antidepressant, *amitriptyline*, is helpful in tension headaches, which might be related to its independent analgesic effect. Amitriptyline has the added bonus of increasing slow wave sleep, which may be why it is more effective than NSAIDs in chronic widespread pain. Tricyclic antidepressants that affect both serotonin and norepinephrine (noradrenaline) reuptake (e.g. p. 1202) seem to be more effective than more selective norepinephrine reuptake inhibitors, e.g. in neuropathic pain. There is some evidence that tricyclics are generally superior to SSRIs in chronic pain syndromes.

Irritable bowel syndrome

This is one of the commonest functional syndromes, affecting some 10–30% of the population world-wide. The clinical features and management of the syndrome and the related functional bowel disorders are described in more detail on page 311. Although the majority of sufferers with the irritable bowel syndrome (IBS) do not have a psychiatric disorder, depressive illness should be excluded in patients with constipation and a poor appetite. Anxiety disorders should be excluded in patients with nausea and diarrhoea. Persistent abdominal pain or a feeling of emptiness may occasionally be the presenting symptom of a severe depressive illness, particularly in the elderly, with a *nihilistic delusion* that the body is empty or dead inside (see p. 1199).

Management

This is dealt with in more detail on page 312 and in Box 22.7. Seeing a physician who provides specific education that particularly addresses individual illness beliefs and concerns can provide lasting benefit. Psychological therapies that help the more severely affected include biofeedback, hypnotherapy, cognitive behaviour therapy and brief interpersonal psychotherapy. If indicated, the choice of antidepressant should be determined by the effects of these drugs on bowel transit times, with tricyclic antidepressants normally slowing and selective serotonin reuptake inhibitors (SSRIs) (p. 1202) normally speeding up transit times.

Multiple chemical sensitivity, *Candida* hypersensitivity, and food allergies

Some complementary health practitioners, doctors and patients themselves make diagnoses of multiple chemical sensitivities (MCS) (e.g. to foods, smoking, perfumes, petrol), *Candida* hypersensitivity, and allergies (to food, tap-water and even electricity). Symptoms and syndromes attributed to these putative disorders are numerous and variable and include all the functional disorders, *mood disorders* and *arthritis*. Scientific support for the existence of these disorders is weak, particularly when double-blind methodologies have been used.

Type 1 hypersensitivities to foods such as nuts certainly exist, although they are fortunately uncommon (approximately 3 per 1000) (see p. 72). Direct specific food intolerances also occur (e.g. chocolate with migraine, caffeine with IBS).

Candidiasis can occur in the gastrointestinal tract in immunocompromised individuals, such as those with AIDS. Vaginal candidiasis can occur after antibiotic treatment in otherwise healthy women. A double-blind and controlled study of nystatin in women diagnosed as having candidiasis hypersensitivity syndrome showed that vaginitis was the only condition relieved more by nystatin than placebo. There is little evidence of *Candida* having a systemic role in other symptoms.

In spite of this evidence, the patient is often convinced of the legitimacy and usefulness of these diagnoses and their treatments.

Aetiology

Surveys of patients diagnosed with MCS or food allergies have shown high rates of current and previous psychiatric disorders (especially mood and anxiety disorders) (see Table 22.5). Eating disorders (p. 1219) should be excluded in patients with food intolerances. Some patients taking very low carbohydrate diets as putative treatments may develop reactive hypoglycaemia after a high carbohydrate meal, which they then interpret as a food allergy. It has been shown that classical conditioning can produce intolerance to foods and smells in healthy people and this may be a causative mechanism in some patients with intolerance. This supports the existence of these intolerance conditions, but suggests they may be conditioned responses with attendant physiological consequences. This might explain why double-blinding abolishes the reaction to the stimulus.

Management

The general principles in Box 22.7 apply. If one assumes a phobic or conditioned response is responsible, graded exposure (systematic desensitization) to the conditioned stimulus may be worthwhile. Preliminary studies do suggest that this approach may successfully treat such intolerances, in the context of cognitive behaviour therapy.

Premenstrual syndrome

The premenstrual syndrome (PMS) consists of both physical and psychological symptoms that regularly occur during the premenstrual phase and substantially diminish or disappear soon after the period starts. Physical symptoms include headache, fatigue, breast tenderness, abdominal distension and fluid retention. Psychological symptoms can include irritability, emotional lability or low mood, and tension. The *premenstrual (late luteal) dysphoric disorder (PMDD)* is a severe form of PMS with marked mood swings, irritability, depression and anxiety accompanying the physical symptoms. Women who generally suffer from mood disorders may be more prone to experience this disorder. The prevalence of PMS does not vary between cultures and is reported by the majority (75%) of women at some time in their lives. Severely disabling PMS (PMDD) occurs in about 3–8% of women. The cause of the premenstrual syndrome remains unclear, although exacerbating factors include some of those outlined in Table 22.5. Research suggests that abnormalities of reproductive hormone receptors may play a role.

FURTHER READING

Royal College of Physicians and Royal College of Psychiatrists. *The Psychological Care of Medical Patients: A practical guide,* 2nd edn. London: Royal College of Physicians, 2003.

Wessly S, White PD. There is only one functional somatic syndrome. *British Journal of Psychiatry* 2004; **185:** 95–96.

Yonkers KA, O'Brien PM, Eriksson E. Premenstrual syndrome. *Lancet* 2008; **371:** 1200–1210.

Management

The general principles in Box 22.7 apply. Treatments with vitamin B_6 (p. 223), diuretics, progesterone, oral contraceptives, oil of evening primrose and oestrogen implants or patches (balanced by cyclical norethisterone) remain empirical. Psychotherapies aimed at enhancing the patient's coping skills can reduce disability. Two trials suggest that graded exercise therapy improves symptoms. Several studies have demonstrated that SSRIs (p. 1202) are effective treatments for the premenstrual dysphoric disorder.

The menopause

The clinical features and management of the menopause are described on page 997. A prospective study has shown that there is no increased incidence of depressive disorders at this time. Such a significant bodily change, sometimes occurring at the same time as children leaving home, is naturally accompanied by an emotional adjustment that does not normally amount to a pathological state.

SOMATOFORM DISORDERS

As explained in the section on functional disorders (p. 1192), the classification of somatoform disorders is unsatisfactory because of the uncertain nature and aetiology of these disorders. However, there are certain disorders, beyond those described in 'functional disorders', that present frequently and coherently enough to be usefully recognized.

Somatization disorder

One in ten patients presenting with a functional disorder will fulfil the criteria of a chronic somatization disorder, sometimes known as *Briquet's syndrome*. The condition is composed of multiple, recurrent, medically unexplained physical symptoms, usually starting early in adult life. Exhaustion, dizzy spells, headaches, hypersensitivity to light and noise, paraesthesiae, abdominal, neck and back pain, nausea, sexual symptoms, and abnormal skin sensations are among the most common complaints, but symptoms may be referred to almost any part or bodily system. The patient, usually female, has often had multiple medical opinions and repeated negative investigations. Medical reassurance that the symptoms do not have a demonstrable physical cause fails to reassure the patient, who will continue to 'doctor-shop'. The patient is usually reluctant to accept a psychological and/or social explanation for the symptoms even when such a link seems obvious. Abnormal illness behaviour is evident and patients can be attention-seeking and dependent on doctors. Yet they can complain about the medical care and attention they have previously received.

The aetiology is unknown, but both mood and personality disorders are often also present. It is often associated with dependence upon or misuse of prescribed medication, usually sedatives and analgesics. There is often a history of significant childhood traumas, or chronic ill-health in the child or parent, which may play an aetiological role (see Table 22.5). The condition is probably the somatic presentation of psychological distress, although iatrogenic damage (from postoperative and drug-related problems) soon complicates the clinical picture. The course of the disorder is chronic and disabling, with long-standing family, marital and/or occupational problems.

Hypochondriasis

The conspicuous feature is a preoccupation with an assumed serious disease and its consequences. Patients commonly believe that they suffer from cancer or AIDS, or some other serious condition. Characteristically, such patients repeatedly request laboratory and other investigations to either prove they are ill or reassure themselves that they are well. Such reassurance rarely lasts long before another cycle of worry and requests begins. The symptom of hypochondriasis may be secondary to or associated with a variety of psychiatric disorders, particularly depressive and anxiety disorders. Occasionally the hypochondriasis is delusional, secondary to schizophrenia or a depressive psychosis. Hypochondriasis may coexist with physical disease but the diagnostic point is that the patient's concern is disproportionate and unjustified.

Management of somatoform disorders

The principles outlined in Box 22.7 apply to these disorders. Patients appreciate a discussion and explanation of their symptoms. Further management consists of ceasing reassurance that no serious disease has been uncovered, since this simply reinforces dependence on the doctor. The doctor should sensitively explore possible psychological and social difficulties, if possible by demonstrating links between symptoms and stresses. Useful questions to ask include:

'When were you last completely well and happy?'

Such a patient may have trouble remembering such a time, which helps to support the diagnosis, and leads to a discussion as to why they have never been well or happy.

'What can't you do now because you are unwell?'

'What changes has your ill-health caused in your close relationships?'

These questions usually give information that can be used to formulate an agreed plan of management. Repeated laboratory investigations should be discouraged. It is vital that all members of staff and close family members adopt the same approach to the patient's problems. Such patients often consciously or unconsciously split both medical staff and family members into 'good' and 'bad' (or caring and uncaring) people, as a way of projecting their distress. Since these disorders have a poor prognosis, the aim is to minimize disability. A contract of mutually agreed care involving the appropriate professionals (general practitioner, and a choice of psychotherapist, health psychologist, complementary health professional, physician or psychiatrist), with agreed frequency of visits and a review date, can be helpful in managing the condition.

Cognitive behaviour therapy has been shown to provide effective rehabilitation in significant numbers of patients suffering from a somatoform disorder.

DISSOCIATIVE/CONVERSION DISORDERS

Until recently these disorders were known as 'hysteria', but because the word hysteria is sometimes used pejoratively to describe extravagant behaviour, the term is inappropriate.

FURTHER READING

Boorsky A, Ahern D. Cognitive behaviour therapy for hypochondriasis. Journal of the American Medical Association 2004; **291:** 1464–1470.

Butler CC, Evans M, Greaves D et al. Medically unexplained symptoms: the biopsychosocial model found wanting. Journal of the Royal Society of Medicine 2004; **97:** 219–222.

Henningsen P, Zipfel S, Herzog W. Management of functional somatic disorders. *Lancet* 2007; **369**: 946–955.

A *dissociative disorder* is a condition in which there is a profound loss of awareness or cognitive ability without medical explanation. The term dissociative indicates the disintegration of different mental activities, and covers such phenomena as amnesia, fugues and pseudoseizures (non-epileptic fits).

Conversion was introduced by Freud to explain how an unresolved conflict could be converted into usually symbolic physical symptoms as a defence against it. Such symptoms commonly include paralysis, abnormal movements, sensory loss, aphonia, disorders of gait and pseudocyesis (false pregnancy). The lifetime prevalence has been estimated at 3–6 per 1000 in women, with a lower prevalence in men. Most cases begin before the age of 35 years. Dissociation is unusual in the elderly.

Clinical features

The various symptoms are usually divided into dissociative and conversion categories (Table 22.6). Dissociative disorders have the following four characteristics that are necessary in order to make the diagnosis:

- They occur in the absence of physical pathology that would fully explain the symptoms.
- They are produced unconsciously.
- The illness is triggered by one or more unresolved conflicts or life events.
- Symptoms are not caused by overactivity of the sympathetic nervous system.

Other characteristics include:

- Symptoms and signs often reflect a patient's ideas about illness.
- Patients may take up the symptoms of a relative/friend who has been ill.
- There is usually abnormal illness behaviour, with obvious exaggeration of disability.
- There may have been significant childhood traumas.
- *Primary gain* is the immediate relief from the emotional conflict.
- *Secondary gain* refers to the social advantage gained by the patient by being ill and disabled (sympathy of family and friends, being off work, disability pension).
- Physical disease is not uncommonly also present (e.g. pseudoseizures are more common in someone with epilepsy).

Dissociative amnesia commences suddenly. Patients are unable to recall long periods of their lives and may even deny any knowledge of their previous life or personal identity. In a *dissociative fugue*, patients not only lose their memory but wander away from their usual surroundings, and, when found, deny all memory of their whereabouts during this wandering. The differential diagnosis of a fugue state includes post-ictal automatism, depressive illness and alcohol abuse.

Multiple personality disorder is rare, but dramatic, and may be triggered by suggestion on the part of a psychotherapist. There are rapid alterations between two or more 'personalities' in the same person, each of which is repressed and dissociated from the other 'personalities'. A differential diagnosis is rapid cycling manic depressive disorder which would explain sudden apparent changes in personality.

Epidemic or '*mass hysteria*' usually occurs in institutions for girls or young women, in which the combined effects of suggestion and shared anxiety produce outbreaks of sickness or disturbed behaviour, often following sudden illnesses in leaders of the group at a time of threatened or actual social change.

Table 22.6 Common dissociative/conversion symptoms

Dissociative (mental)	Conversion (physical)
Amnesia	Paralysis
Fugue	Disorders of gait
Pseudodemential	Tremor
Dissociative identity disorder	Aphonia
Psychosis	Mutism
	Sensory symptoms
	Globus hystericus
	Hysterical fits
	Blindness

Differential diagnosis

Dissociation is usually a stable and reliable diagnosis over time, although high rates of co-morbid mood and personality disorders are found in chronic sufferers. Particular care should be taken to make the diagnosis on positive grounds, and not simply on the basis of an absence of a medical diagnosis. Care should also be taken to exclude or treat co-morbid psychiatric disorders.

Aetiology

Functional brain scans differ between healthy controls feigning a motor abnormality and patients with a similar conversion motor symptom, which suggests that dissociation involves different areas of the brain from simulation (Fig. 22.1). This supports the theory of unconscious mechanisms

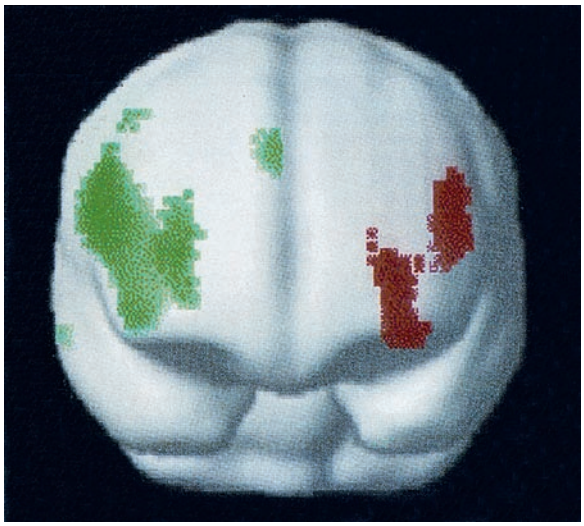


Fig. 22.1 **Statistical parametric maps superimposed on an MRI scan of the anterior surface of the brain, orientated as though looking at a person head on.** Red region shows hypofunction of patients with conversion motor symptoms. Green region shows hypofunction of healthy controls feigning the same motor abnormality. Reproduced from Spence SA, Crimlisk HL, Cope H et al. Discrete neurophysiological correlates in prefrontal cortex during hysterical and feigned disorders of movement. *Lancet* 2000; **355**: 1243–1244, © The Lancet Ltd. 2000, with permission.

first suggested by Charcot. Functional brain scanning of a patient with conversion paralysis has shown that recalling a past trauma not only activated the emotional areas, such as the amygdala, but also reduced motor cortex activity. This would suggest that conversion involves a disinhibition of voluntary will at an unconscious level, so that the patient can no longer *will* something to happen.

The psychoanalytical theory of dissociation is that it is the result of emotionally charged memories that are repressed into the unconscious at some point in the past. Symptoms are explained as the combined effects of repression and the symbolic conversion of this emotional energy into physical symptoms. This hypothesis is difficult to test, although there is some evidence that patients with dissociative disorders are more likely to have suffered childhood abuse, particularly when the abuse was both sexual and physical and started early in childhood. Caution should be taken with such a history obtained by therapies that 'recover' childhood memories that were previously completely unknown to the patient.

Patients with dissociative disorders by definition adopt both the sick role and abnormal illness behaviour, with consequent secondary gains that help to maintain the illness.

Management

The treatment of dissociation is similar to the treatment of somatoform disorders in general, outlined above and in Box 22.7. The first task is to engage the patient and their family with an explanation of the illness that makes sense to them, is acceptable, and leads to the appropriate management. An invented example of a suitable explanation is given below:

You told me about the tremendous shock you felt when your mother suddenly died. This was particularly the case since you hadn't spoken to her for so long beforehand, after that big disagreement with her over your wedding to John. You weren't able to say good-bye before she died. Your brain was overloaded with grief, guilt and anger all at once. I wonder whether that is why you aren't able to speak now. I wonder whether it's difficult to think of anything to say that would make things right, particularly since you can't speak with your mother now.

Such an explanation would be modified by mutual discussion until an agreed understanding was achieved, which would serve as a working model for the illness. Provision of a rehabilitation programme that addresses both the physical and psychological needs and problems of the patient would then be planned. A graded and mutually agreed plan for a return to normal function can usually be led by the appropriate therapist (e.g. speech therapist for dysphonia, physiotherapist for paralysis). At the same time, a psychotherapeutic assessment should be made in order to determine the appropriate form of psychotherapy. For instance, couple therapy will address a significant relationship difficulty; individual psychotherapy could ease an unresolved conflict from childhood.

Abreaction brought about by *hypnosis* or by intravenous injections of small amounts of midazolam may produce a dramatic, if short-lived, recovery. In the abreactive state, the patient is encouraged to relive the stressful events that provoked the disorder and to express the accompanying emotions, i.e. to abreact. Such an approach has been useful in the treatment of acute dissociative states in wartime, but appears to be of much less value in civilian life. It should only be contemplated in the presence of an anaesthetist with suitable resuscitation equipment to hand.

Hypnotherapy is psychotherapy while the patient is in a hypnotic trance, the idea being that therapy is more possible because the patient is relaxed and not using repression. This may allow the therapist access to the previously unconscious emotional conflicts or memories. There are no published trials of this technique in dissociation, which Freud gave up as unsuccessful in order to found psychoanalysis, but some hypnotherapists claim good results. Care should be taken to avoid a catastrophic emotional reaction when the patient is suddenly faced with the previously repressed memories.

Prognosis

Most cases of recent onset recover quickly with treatment, which is why a positive diagnosis should be made early. Those cases that last longer than a year are likely to persist, with entrenched abnormal illness behaviour patterns. A recent study found that 83% were still unwell at 12 years' follow-up.

FURTHER READING

Crimlisk HL, Bhatia K, Cope H et al. Slater revisited: 6 year follow up study of patients with medically unexplained symptoms. *British Medical Journal* 1998; **316**: 582–586.

SLEEP DIFFICULTIES (p. 1145)

Sleep is divided into *rapid eye movement* (REM) and non-REM sleep. As drowsiness begins, the alpha rhythm on an EEG disappears and is replaced by deepening slow wave activity (non-REM). After 60–90 minutes, this slow wave pattern is replaced by low amplitude waves on which are superimposed rapid eye movements lasting a few minutes. This cycle is repeated during the duration of sleep, with the REM periods becoming longer and the slow-wave periods shorter and less deep (Fig. 22.2). REM sleep is accompanied by dreaming and physiological arousal. Slow wave sleep is associated with release of anabolic hormones and cytokines, with an increased cellular mitotic rate. It helps to maintain host defences, metabolism and repair of cells. For this reason slow wave sleep is increased in those conditions where growth or conservation is required (e.g. adolescence, pregnancy, thyrotoxicosis).

Insomnia is difficulty in sleeping; a third of adults complain of insomnia and in a third of these it can be severe.

Primary sleep disorders include sleep apnoea (p. 1145), narcolepsy (see p. 1145), *the restless legs syndrome*

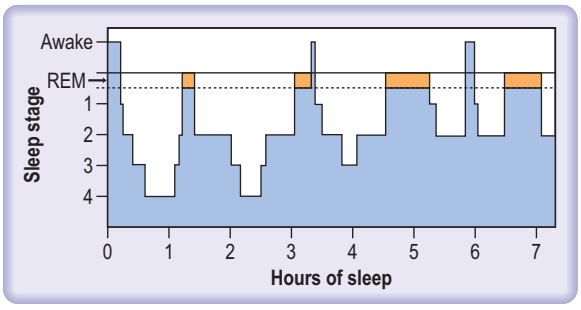


Fig. 22.2 **Sleep architecture, showing cycles of slow wave sleep interspersed with rapid eye movement (REM) sleep.** Sleep is 'staged' by the electroencephalogram (EEG). Deeper sleep (stages 3 and 4) is demonstrated by slow waves on the EEG. Sleep occurs in cycles, light sleep (accompanied by REM sleep and dreaming) to deep sleep and back again, and these cycles last about 90 minutes.

(Ekbom's) (see p. 626) and the related *periodic leg movement disorder*, in which the legs (and sometimes the arms) jerk while asleep.

Delayed sleep phase syndrome occurs when the circadian pattern of sleep is delayed so that the patient sleeps from the early hours until mid-day or later, and is most common in young people. *Night terrors, sleep-walking* and *sleep-talking* are non-REM phenomena, most commonly found in children, which can recur in adults when under stress or suffering from a mood disorder.

Psychophysiological insomnia commonly occurs with functional, mood and substance misuse disorders, and when under stress (Box 22.8). It can often be triggered by one of these factors, but then become a habit on its own, driven by anticipation of insomnia and day-time naps. Insomnia causes day-time sleepiness and fatigue, with consequences such as road traffic accidents. Assessment should pay particular attention to mood, life difficulties and drug intake (especially alcohol, nicotine and caffeine). Initial insomnia (trouble getting off to sleep) is common in mania, anxiety, depressive disorders and substance misuse. Middle insomnia (waking up in the middle of the night) occurs with medical conditions such as sleep apnoea and prostatism. Late insomnia (early morning waking) is caused by depressive illness and malnutrition (anorexia nervosa).

Habitual alcohol consumption should be carefully estimated since even a small excess can be a potent cause of insomnia, as well as recent withdrawal. *Caffeine* is perhaps the most commonly taken drug in the UK, and its effects are easily underestimated. Six cups of real coffee a day are likely to cause insomnia in the average healthy adult. Caffeine is not only found in tea and coffee, but is also found in chocolate, cola drinks and some analgesics. Prescription drugs that can either disturb sleep or cause vivid dreams include most appetite suppressants, glucocorticoids, dopamine agonists, lipid-soluble beta-blockers (e.g. propranolol) and certain psychotropic drugs (especially when first prescribed, e.g. fluoxetine, reboxetine, risperidone).

Hypersomnia is not uncommon in adolescents with depressive illness, occurs in narcolepsy, and may temporarily follow infections such as infectious mononucleosis.

Management of insomnia

This is determined by diagnosis. Where none is immediately apparent, it is worth educating the patient about sleep hygiene. Simple measures such as decreasing alcohol intake, having supper earlier, exercising daily, having a hot bath prior to going to bed and establishing a routine of going to bed at the same time should all be tried. Relaxation techniques and cognitive behaviour therapy have a role in those with intractable insomnia. Short half-life benzodiazepines can be useful for acute insomnia, but should not be used for more than 2 weeks continuously to avoid dependence. Non-benzodiazepine hypnotics (zaleplon, zopiclone, zolpidem) act at the benzodiazepine receptors and occasional dependence has been reported. Certain antihistamines (e.g. diphenhydramine and promethazine) and antidepressants (e.g. amitriptyline, trimipramine, trazodone, mirtazapine) are not addictive and can be used as hypnotics in low dose, with the added advantage of improving slow wave sleep. The commonest side-effects are morning sedation and weight gain.

Box 22.8 Common causes of insomnia

Psychiatric disorders

Mood disorders (mania, depressive and anxiety disorders)
Delirium and dementia

Drug use or misuse

Addictive drug withdrawal (alcohol, benzodiazepines)
Stimulant drugs (caffeine, amfetamines)
Prescribed drugs (steroids, dopamine agonists)

Physical conditions

Pain (classically with carpal tunnel syndrome)
Nocturia (e.g. from prostatism)
Malnutrition

Primary sleep disorders

Periodic leg movements
Restless legs syndrome

FURTHER READING

Sateia MJ, Nowell PD. Insomnia. *Lancet* 2004; **364:** 1959–1973.

Wilson S, Nutt D. Assessment and management of insomnia. *Clinical Medicine* 2005; **5:** 101–104.

MOOD (AFFECTIVE) DISORDERS

Classification

The central and common feature of these disorders is an abnormality of mood. Mood is best considered in terms of a continuum ranging from severe depression at one extreme to severe mania at the other, with the normal, stable mood in the middle. Mood disorders are divided into bipolar and unipolar affective disorders.

Bipolar affective disorder (otherwise known as manic-depressive disorder). Patients suffer bouts of both depression and mania. Although mania can rarely occur by itself without depressive mood swings (thus being 'unipolar') it is far more commonly found in association with depressive swings, even if sometimes it takes several years for the first depressive illness to appear. Hypomania is mild mania.

- *Bipolar I* disorder is defined as depression alternating with mania.
- *Bipolar II* is defined as depression alternating with hypomania.

About 10% of patients with depressive illness are eventually found to have a bipolar illness.

Unipolar affective disorders. Patients suffer from depressive mood swings alone, although they are commonly recurrent.

DEPRESSIVE DISORDERS

Depressive disorders or 'episodes' are classified by the ICD-10 as mild, moderate or severe, with or without somatic symptoms. Severe depressive episodes are divided according to the presence or absence of psychotic symptoms. Dysthymia is a chronic low-grade unipolar depressive illness.

Clinical features of depressive disorder

Whereas everyone will at some time or other feel cheesed off, fed up or down in the dumps, it is when such symptoms become qualitatively different, pervasive or interfere with normal functioning that a depressive illness has occurred. Depressive disorder, clinical or 'major' depression, is characterized by disturbances of mood, speech, energy and ideas (Table 22.7). Patients often describe their symptoms in physi-

Table 22.7 Characteristic features of depressive illness

Characteristic	Clinical features
Mood	Depressed, miserable or irritable
Talk	Impoverished, slow, monotonous
Energy	Reduced, lethargic
Ideas	Feelings of futility, guilt, self-reproach, unworthiness, hypochondriacal preoccupations, worrying, suicidal thoughts, delusions of guilt, nihilism and persecution
Cognition	Impaired learning, pseudodementia in elderly patients
Physical	Insomnia (especially early waking), poor appetite and weight loss, constipation, loss of libido, erectile dysfunction, bodily pains
Behaviour	Retardation or agitation, poverty of movement and expression
Hallucinations	Auditory – often hostile, critical

Box 22.9 Screening questions for depressive illness

- During the last month, have you often been bothered by feeling down, depressed or hopeless?
- During the last month, have you often been bothered by having little interest or pleasure in doing things?

If one or both of the answers is 'yes', assess further for depressive illness.

cal terms. Marked fatigue and headache are the two most common physical symptoms in depressive illness and may be the first symptoms to appear. Patients describe the world as looking grey, themselves as lacking a zest for living, being devoid of pleasure and interest in life (anhedonia). Anxiety and panic attacks are common; secondary obsessional and phobic symptoms may emerge. Symptoms should last for at least 2 weeks and should cause significant incapacity (e.g. trouble working or relating to others) to be considered an illness.

In the more severe forms, diurnal variation in mood can occur, feeling worse in the morning, after waking in the early hours with apprehension. Suicidal ideas are more frequent, intrusive and prolonged. Delusions of guilt, persecution and bodily disease are not uncommon, along with second person auditory hallucinations insulting the patient or suggesting suicide. In severe depressive illness, particularly in the elderly, concentration and memory can be so badly affected that the patient appears to have dementia (pseudodementia). Delusions of poverty and non-existence (nihilism) occur particularly in this age group. Suicide is a real risk, with the lifetime risk being approximately 5% in primary care patients, but 15% in those with depressive illness severe enough to warrant admission to hospital. Patients with bipolar disorder are also at greater risk of suicide. Screening questions for depressive illness are shown in Box 22.9.

Epidemiology

About a third of the population will feel unhappy at any one time, but this is not the same as depressive illness; the middle-aged feel least happy compared to the young and elderly. The point prevalence of depressive illness is 5% in the community, with a further 3% having dysthymia (see below). It is more common in women, but there is no increase with age, and no difference by ethnic group or socioeconomic class (apart from an inverse relationship only with dysthymia). Married and never married people have similar prevalence rates, with separated and divorced people having two to three times the prevalence. Some studies have suggested that depressive illness is becoming more common.

Depressive illnesses are more common in the presence of:

- physical diseases, particularly if chronic, stigmatizing or painful
- excessive and chronic alcohol use (probably the most depressing drug humans use)
- social stresses, particularly loss events, such as separation, redundancy and bereavement
- interpersonal difficulties with those close to the patient, especially when socially humiliated
- lack of social support, with no confiding relationship.

Depressed patients with another physical disorder view themselves as more sick, visit their doctors almost four times as often as the non-depressed physically ill, stay in hospital longer, comply less with medical advice and medication, and undergo more medical and surgical procedures. Depressive illness may be associated with increased mortality (excluding suicide) in patients with physical illness, such as myocardial infarct.

Dysthymia

Dysthymia is a mild or moderate depressive illness that lasts intermittently for 2 years or more and is characterized by tiredness and low mood, lack of pleasure, low self-esteem and a feeling of discouragement. The mood relapses and remits, with several weeks of feeling well, soon followed by longer periods of being unwell. It can be punctuated by depressive episodes of greater severity, so-called 'double depression'.

Seasonal affective disorder

Seasonal affective disorder is characterized by recurrent episodes of depressive illness occurring during the winter months in the northern hemisphere. Symptoms are similar to those found with atypical depressive illness, in that patients complain of hypersomnia, increased appetite (with carbohydrate craving) and weight gain, with profound fatigue. Such patients have a higher prevalence of bipolar affective disorder, and some doctors are uncertain whether the condition is different from normal depressive illness, with the accentuation of mood that naturally occurs by season. However, there is evidence that seasonal depressive illness can be successfully treated with bright light therapy given in the early morning, which causes a phase advance in the circadian rhythm of melatonin. In contrast, the same treatment given in the early evening, with consequent phase delay of melatonin secretion, is less antidepressant. Selective serotonin reuptake inhibitors (SSRIs) are alternative treatments.

Puerperal affective disorders

Affective illnesses and distress are common in women soon after they have given birth.

'*Maternity blues*' describe the brief episodes of emotional lability, irritability and tearfulness that occur in about 50% of women 2–3 days postpartum and resolve spontaneously in a few days.

Postpartum psychosis occurs once in every 500–1000 births. Over 80% of cases are affective in type and the onset is usually within the first 2 weeks following delivery. In addition to the classical features of an affective psychosis, disorientation and confusion are often noted. Severely depressed patients may have delusional ideas that the child is deformed, evil or otherwise affected in some way, and such false ideas may lead to either attempts to kill the child or suicide. The response to speedy treatment is generally good. The recurrence rate for a psychosis in a subsequent puerperium is 20–30%.

Non-psychotic postnatal depressive disorders occur during the first postpartum year in 10% of mothers, especially in the first 3 months. Risk factors are first pregnancy, poor relationship with the partner, ambivalence about the pregnancy, and emotional personality traits. The Edinburgh Postnatal Depression Scale (EPDS) is a 10-item questionnaire and can be used as an effective screening tool. Depressive illness after childbirth is clinically similar to other depressive illnesses, but lack of emotional bonding with the baby is common.

Differential diagnosis

The differential diagnoses of depressive illness are shown in Table 22.8. Other psychiatric disorders are the most common misdiagnoses. Ninety per cent of patients presenting with a depressive illness while misusing alcohol will no longer be depressed 2 weeks after their last drink.

Pathological (abnormal) and normal grief are described on page 1209. Pathological grief is closely associated with depressive illness.

Investigations

A corroborative history can be valuable in helping to exclude differential diagnoses such as alcohol misuse and elucidating maintaining factors such as a poor relationship with a partner. Physical investigations should be guided by the history and examination. They will often include measurement of free T_4 and TSH (particularly in women), calcium, sodium, potassium, mean corpuscular volume, γ-glutamyl transpeptidase, haemoglobin, white cell count, ESR or plasma viscosity. Less commonly a chest X-ray, anti-nuclear antibody, morning and evening cortisols, electroencephalogram or a brain scan are indicated.

Table 22.8 Common differential diagnoses of depressive illness

Other psychiatric disorders
Alcohol misuse
Amfetamine (and derivatives) misuse and withdrawal
Borderline personality disorder
Dementia
Delirium
Schizophrenia
Normal and pathological grief
Organic (secondary) affective illness
Physical causes which are both necessary and sufficient as a cause
Cushing's syndrome
Thyroid disease (although sometimes depression persists after treatment)
Hyperparathyroidism
Corticosteroid treatment
Brain tumour (rarely without other neurological signs)

The aetiology of unipolar depressive disorders

The aetiology of unipolar depressive disorders is multifactorial and a mixture of genetic and environmental factors.

Genetic

Unipolar depression is probably polygenic, but no linkage has been firmly identified. The risk of unipolar depression in a first-degree relative of a patient is approximately three times the risk of the non-affected. The concordance of unipolar depression in monozygotic twins is between 30 and 60%, the concordance increasing with more recurrent illnesses. Polymorphisms that increase the risk of depression involve monoamines and their receptors. The issue is complicated by the genetic influence on sleep habits, 'neurotic' personality, and even life events, which are all involved in the genesis of depressive illness.

Biochemical

The monoamine deficiency theory of depressive illness is supported by the efficacy of monoamine reuptake inhibitors and the depressive effect of dietary tryptophan depletion. Neuroendocrine tests also suggest that the serotonin neurotransmitter system is downregulated. 5-HT_{1a} and 5-HT_2 receptor subtypes are thought most likely to be involved. Receptor-labelled functional brain scans suggest that dopamine underactivity is related to psychomotor retardation.

Hormonal

Cushing's syndrome is the most potent cause of so-called 'organic' depressive illness, with 50–80% of patients with Cushing's suffering from a depressive illness. Corticosteroid treatment causes significant mood disturbance. Nearly half of patients with 'functional' depressive illness have raised cortisol levels, and this is associated with adrenal gland enlargement. Hypercortisolaemia can cause hippocampal damage, which has been found in chronic severe depressive illness. All these data suggest that cortisol may play a role in causing depressive illness.

In contrast, *atypical depressive illness*, with prominent hypersomnia and weight gain, is associated with a downregulated hypothalamic–pituitary–adrenal axis, supporting the heterogeneity of depressive disorders.

Neural changes and neuronal growth

The use of functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) has revealed a number of abnormalities in the brains of patients with major depression. Increased brain ventricle volume, orbitofrontal, dorsolateral frontal and anterior cingulate cortex altered activation have all been implicated. The hippocampus is smaller in several stress-related neuropsychiatric disorders such as recurrent depression. This may be related to hypercortisolaemia. Both emotional and cognitive functions are therefore involved in the neural basis of depression.

Recent work suggests that depressive illness is associated with reduced growth of new neurones in parts of the brain involved with emotional expression, such as the hippocampus.

Sleep

A reduced time between onset of sleep and REM sleep (shortened REM latency) and reduced slow wave sleep both

occur in depressive illness. These abnormalities persist in some patients when they are not depressed. Families with several sufferers of depressive illness can share these traits, suggesting that sleep patterns may be inherited and predispose to depression.

Childhood traumas and personality

Physical, sexual and emotional abuse or neglect in childhood all predispose adults to depressive illness, but the effect is non-specific. Both 'neurotic' (emotional) and perfectionist personality traits are risks for depressive illness, and these may be determined as much by genetic factors as childhood environment.

Social

Thirty per cent of women will develop a depressive illness after a severe life event or difficulty, such as a divorce, and this is compounded by low self-esteem and a lack of a confiding relationship. Unemployment is a significant risk factor in men.

An integrated model of aetiology

Stress is more likely to trigger depressive illness in a person predisposed by lack of social support and/or certain personality traits. Stress in turn triggers various brain changes in both stress hormones (such as the release of corticotrophin-releasing hormone) and neurotransmitters (e.g. serotonin) that are both known to be altered in depressive illness. We can thus start to glimpse the model of an integrated bio-psychosocial model of depressive illness. This model challenges dualistic ideas that depressive illnesses are either psychological or physical; depressive illnesses involve both the mind and the body, which are themselves indivisible.

Treatment of depressive illness

The patient needs to know the diagnosis to provide understanding and rationalization of the overwhelming distress inherent in depressive illness. Knowing that self-loathing, guilt and suicidal thoughts are caused by the illness can be 'antidepressant' on its own. The further treatment of depressive disorders involves physical, psychological and social interventions (Box 22.10). Patients who are actively suicidal, severely depressed or with psychotic symptoms should be admitted to hospital (necessary for perhaps 1 in 1000 patients with clinical depression in primary care). This provides the patient a break from self-care, and allows support, listening, observation, prevention of suicide and close monitoring of treatments. Avoid the pitfall of not treating a depressive illness just because it seems an 'understandable' reaction to serious illness or difficult circumstances. This is particularly likely to happen if the patient is elderly, severely or even terminally ill.

Exercise

There is good evidence that regular exercise, particularly involving other people, can help relieve depressive illness of mild or moderate severity. The benefit is independent of a physical training effect.

Drugs used in the treatment of clinical depression

- Recreational drugs such as alcohol should be stopped. Prescribed medicines suspected of exacerbating depression, such as corticosteroids, should be gradually stopped or reduced to a safe minimum.
- The first course of *antidepressant drugs* is effective in relieving clinical depression in 60–70% of patients, if given in adequate doses for a sufficient time to the correctly diagnosed patient. Such treatment is more successful when accompanied by sufficient patient education and regular follow-up, particularly a week after starting treatment and throughout the following 6 weeks. Dysthymia responds less well to antidepressants than does a depressive episode.
- The commonest two pharmacological types of antidepressants are *tricyclic antidepressants (TCAs)* and *selective serotonin reuptake inhibitors (SSRIs)*. All antidepressants have similar efficacy and speed of onset. Choice depends on their side-effects, which can be used to positive effect (sedating drugs given at night to enhance sleep), and their safety. Patients should be warned about side-effects and that it will take two or more weeks before a positive benefit is apparent. This is much slower than the time taken to alter neurotransmitter release and reuptake. The longer time is consistent with receptor adaptation, although recent animal work shows that antidepressants are growth factors that increase the development of new neurones from progenitor cells in the hippocampus. Drugs should normally be started at a low dose and increased, depending on side-effects and efficacy. A course of antidepressants should be given until 4 months after recovery to prevent relapse. The two greatest problems with these drugs are persuading the patient to take them and adherence, since 80% of the UK public wrongly believe that they are addictive.
- *Psychotic depression* needs either electroconvulsive therapy or a combination of an antidepressant and an antipsychotic drug.
- *Bipolar depressive illness* is best treated initially with mood stabilizers (see p. 1205).

Box 22.10 Management of depressive illness

Physical

Stop depressing drugs (alcohol, steroids)
Regular exercise (good for mild to moderate depression)
Antidepressants (choice determined by side-effects, co-morbid illnesses and interactions)
Adjunctive drugs (e.g. lithium; if no response to two different antidepressants)
Electroconvulsive therapy (ECT) (if life-threatening or non-responsive)

Psychological

Education and regular follow-up by same professional
Cognitive behaviour therapy (CBT)
Other indicated psychotherapies (couple, family, interpersonal)

Social

Financial: eligible benefits, debt counselling
Employment: acquire or change job or career
Housing: adequate, secure tenancy, safe, social neighbours
Young children: child-care support

Treatments combined

The most effective treatment is a mixture of CBT and an antidepressant

Ask the authors
www.studentconsult.com

Selective serotonin reuptake inhibitors (SSRIs)

SSRIs selectively inhibit the reuptake of the monoamine serotonin (5-HT) within the synapse, and are thus termed 'selective serotonin reuptake inhibitors' or SSRIs. Citalopram and its laevo isomer escitalopram, fluvoxamine, fluoxetine, paroxetine and sertraline have the advantage of causing less serious or disabling side-effects than tricyclics. For instance, SSRIs do not usually cause significant weight gain. Because of their long half-lives they can also be given just once a day, normally in the morning after breakfast. For these reasons patients adhere more to treatment and therefore SSRIs are now first-line treatments for depressive disorders. Normal doses are between 20 and 60 mg, with sertraline and fluvoxamine needing higher doses. The most common side-effects resemble a 'hangover' and include nausea, vomiting, headache, diarrhoea and dry mouth. Insomnia and paradoxical agitation can occur when first starting the drugs. Recent studies suggest that adolescents, in particular, may develop suicidal thoughts with SSRIs. Further studies suggest that this is a small risk, if present, and no study has shown a significant increased risk of suicide itself. One in five patients also has sexual side-effects, such as erectile dysfunction and loss of libido. Uncommon side-effects include restless legs syndrome (see p. 626) and hyponatraemia.

A toxic hyperserotonergic state ('s*erotonin syndrome*') can be caused by the ingestion of two or more drugs that increase serotonin levels, e.g. an SSRI combined with a monoamine oxidase inhibitor (MAOI), dopaminergic drugs (e.g. selegiline) or a tricyclic antidepressant. Symptoms include agitation, confusion, tremor, diarrhoea, tachycardia and hypertension; hyperthermia is characteristic. Treatment is supportive.

SSRIs have also been associated with a specific withdrawal syndrome (*discontinuity syndrome*). This is characterized by shivering, anxiety, dizziness, 'electric shocks', headache and nausea. Patients should be warned not to leave out a dose and to gradually reduce SSRIs when stopping them. All antidepressants have the potential to cause a withdrawal syndrome if suddenly stopped.

Tricyclic antidepressants (TCAs)

Dosulepin, imipramine and amitriptyline are the three most commonly used in the UK, but many related compounds have been introduced, some having fewer autonomic and cardiotoxic effects (e.g. lofepramine). These drugs potentiate the action of the monoamines, noradrenaline (norepinephrine) and serotonin, by inhibiting their reuptake into nerve terminals (Fig. 22.3). Other tricyclics in common use include nortriptyline, doxepin and clomipramine. Depending on the particular drug, normal doses are between 75 and 150 mg. Having been available for more than 40 years, there is more evidence of the effectiveness of TCAs in depressive illness than for any other group of antidepressants. They are the drugs most commonly used in severe depressive illness.

TCAs have a number of side-effects (Table 22.9). In long-term treatment or prophylaxis, weight gain is most troublesome. Because of their toxicity in overdose, it is wisest not to prescribe them to outpatients with suicidal thoughts without monitoring or giving the drugs to a reliable family member to look after.

Trazodone is different from other tricyclic antidepressants and acts by blockade of $5H_2$ receptors.

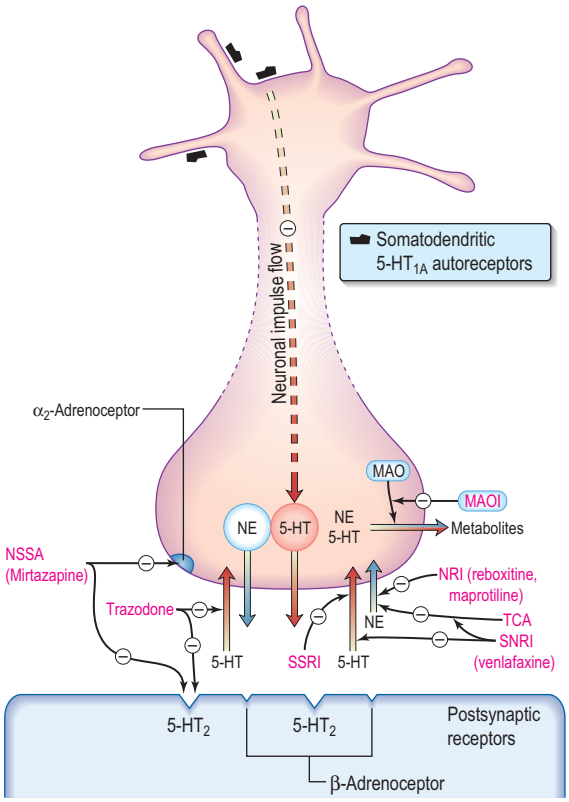


Fig. 22.3 **Sites of action of antidepressants with examples.** In a normal neurone, the stimulation of the somatodendritic 5-HT_{1A} autoreceptors inhibits the neuronal impulse flow down the axon, reducing 5-HT release. In depression, there is a reduction in amine neurotransmissions which results in upregulation of postsynaptic and somatodendritic receptors. Using 5-HT as an example, the depletion of 5-HT results in upregulation of postsynaptic 5-HT_2 receptors and presynaptic (somatodendritic) 5-HT_{1A} receptors. The many antidepressants now available have different actions on serotonin and noradrenaline (norepinephrine) neurotransmission. Classical antidepressants (TCA) have antimuscarinic and antihistaminergic activities. SSRI, selective serotonin reuptake inhibitors; SNRI, serotonin and noradrenaline (norepinephrine) reuptake inhibitors; NRI, noradrenaline (norepinephrine) reuptake inhibitor (selective); NE, noradrenaline (norepinephrine); MAO, monoamine oxidase; MAOI, monoamine oxidase inhibitor. From Waller DG, Renwick A, Hillier K (eds). *Medical Pharmacology and Therapeutics*, Edinburgh: Saunders, 2001, with permission from Elsevier.

SNRIs, NSSAs and NRIs: antidepressants

The latest generation of antidepressants block a number of different neurotransmitter receptors both at the synapse and elsewhere. Their different receptor profiles cause different side-effects.

Venlafaxine is a potent blocker of both *serotonin* and *noradrenaline (norepinephrine) reuptake (SNRI).* It has negligible affinity for other neurotransmitter receptor sites and so produces less sedation and fewer antimuscarinic effects. It can be given in slow-release form with the advantage of once-daily dosage. Nausea is the commonest side-effect and patients should be monitored for hypertension. It should not be prescribed in those with either uncontrolled hypertension or in those prone to cardiac arrhythmias.

Table 22.9 Side-effects of tricyclic antidepressants

Antimuscarinic effects
Dry mouth
Constipation
Tremor
Blurred vision
Urinary retention
Cardiovascular
QT prolongation
Arrhythmias
Postural hypotension
Convulsant activity
Lowered seizure threshold
Other effects
Weight gain
Sedation
Mania (rarely)

Mirtazapine is a 5-HT_2 and 5-HT_3 receptor antagonist and a potent α_2-adrenergic blocker. The consequent effect is to increase both *noradrenaline (norepinephrine) and selective serotonin transmission: an NSSA*. It can be given at night to aid sleep and rarely causes sexual side-effects. Mirtazapine can be sedating in low dose and can cause weight gain. An uncommon adverse effect is agranulocytosis.

Reboxetine is a *selective noradrenaline (norepinephrine) reuptake inhibitor (NaRI).* It is not sedating and may help reduced motivation and energy. Weight gain is not reported. Dry mouth, insomnia, constipation, urinary hesitancy and tachycardia are reported side-effects.

Monoamine oxidase inhibitors (MAOIs)

These act by irreversibly inhibiting the intracellular enzymes monoamine oxidase A and B, leading to an increase of norepinephrine (noradrenaline), dopamine and 5-HT in the brain (see Fig. 22.3). Because of their side-effects and restrictions while taking them, they are rarely used by non-psychiatrists. MAOIs also produce a dangerous hypertensive reaction with foods containing tyramine or dopamine and therefore a restricted diet is prescribed. Tyramine is present in cheese, pickled herrings, yeast extracts, certain red wines, and any food, such as game, that has undergone partial decomposition. Dopamine is present in broad beans. MAOIs interact with drugs such as pethidine and can also occasionally cause liver damage.

Reversible inhibitors of monoamine A (RIMAs)

An example is moclobemide; usual dose 300 mg daily. These drugs appear to have fewer side-effects (insomnia and headache, but some sexual problems) and constitute a low risk in overdose. Patients prescribed such antidepressants should be told that they can eat a normal diet, but should be careful to avoid excessive amounts of food rich in tyramine (see above).

Antidepressant augmentation

If two trials of antidepressants have failed, adding a second concomitant drug, e.g. lithium, can sometimes be helpful.

Antidepressant use in general medicine

- *Cardiac disease.* In patients with cardiac disease, SSRIs, lofepramine and trazodone are preferred over more quinidine-like compounds. MAOIs and mirtazapine do not affect epileptic thresholds. SSRIs are metabolized by the cytochrome P450 system, unlike venlafaxine, mirtazapine and reboxetine; the latter therefore have fewer drug interactions.
- *Herbal medicine.* Care should be taken not to prescribe antidepressants while a patient is taking the herbal antidepressant St John's wort, which interacts with serotonergic drugs in particular.
- *Elderly.* Doses of antidepressants should initially be halved in the elderly and in patients with renal or hepatic failure.
- *Pregnancy.* Antidepressants should be avoided if possible in pregnancy and breast-feeding. If other treatments are ineffective, the risks of drug therapy should be balanced against taking no treatment, since depression can affect fetal progress and future mother–child bonding. Tricyclic antidepressants are generally believed to be safe in pregnancy, with no significant increase in congenital malformations in fetuses exposed to them. However, occasionally their antimuscarinic side-effects produce jitteriness, sucking problems and hyperexcitability in the newborn. Postpartum plasma levels of babies breast-fed by treated mothers are negligible. SSRIs do not seem to be teratogenic but manufacturers advise against their use in pregnancy until more data are available. Pulmonary hypertension in the newborn is a rare complication. MAOIs should be avoided during pregnancy because of the possibility of a hypertensive reaction in the mother.

Electroconvulsive therapy (ECT)

Although the use of ECT is declining in the UK, it is still the treatment of choice in severe life-threatening depressive illness, particularly when psychotic symptoms are present. It is sometimes essential treatment when the patient is dangerously suicidal or refusing to eat and drink. The treatment involves the passage of an electric current across two electrodes applied to the anterior temporal areas of the scalp, in order to induce an epileptic fit. The fit is the essential part of the treatment. Before the treatment is given, the patient is given a general anaesthetic and receives a muscle relaxant to prevent injury during the fit. Treatments are normally given twice a week for 3–6 weeks.

ECT is a controversial treatment, yet it is remarkably safe and free of serious side-effects. Post-ictal confusion and headache are not uncommon, but transient. Short-term retrograde amnesia and a temporary defect in new learning can occur during the weeks of treatment, but these are short-lived effects.

Uncommonly used treatments

Transcranial magnetic stimulation (TMS) shows moderate efficacy, but is uncommonly used. *Psychosurgery* is very occasionally considered in patients with severe intractable depressive illness, when all other treatments have failed (see p. 1211). A third improve remarkably, while a further third improve somewhat. *Vagal nerve and deep brain stimulation* may represent major advances in the management of chronic and treatment-refractory depressive disorders, but definitive trials are awaited.

Psychological treatments

Cognitive behaviour therapy (CBT)

Beck developed CBT to reverse the negative cognitive triad with which patients regarded themselves, their situation and their futures. It involves the identification of the negative automatic thoughts that maintain the negative perceptions that feed depression. They commonly include catastrophizing (e.g. making a 'mountain out of a mole-hill'), overgeneralizing (e.g. 'I failed an exam; therefore I am a failure as a person.'), categorical ('black or white') thinking (e.g. 'My work is either perfect or abysmal.'). CBT then involves identifying the links between these thoughts, consequent behaviour, and feeling low, and then testing their logic. This is done by considering the evidence either in the therapy sessions (e.g. Q: 'Did you pass the other exams you took?' A: 'Yes; I guess I did.') or by behavioural 'experiments' (e.g. showing the 'abysmal' work to a colleague and asking their opinion).

There is good evidence that individual CBT is as effective as antidepressant drugs for mild and moderate depressive illness. CBT is also effective in preventing a relapse of clinical depression. Individual CBT is more effective than group delivered therapy, and there is preliminary evidence that computer delivered CBT programmes are also helpful.

Interpersonal psychotherapy

This psychotherapy is probably as effective as both antidepressants and CBT in mild and moderate clinical depression. The therapist focuses on a patient's interpersonal relationships involved in, or affected by, their illness (especially relationship changes or deficiencies), using problem-solving techniques to help the patient to find solutions.

Other psychotherapies

Couple therapy is particularly effective when a patient is in a relationship with problems; both the patient and partner attend therapy. *Family therapy* is effective not only in a family with problems, but also as a way of helping the family to help the patient get better. It may involve understanding one family member's 'depression' as a systemic 'solution' for a wider problem within the family.

Social treatments

Many patients with clinical depression have associated social problems (see Box 22.10). Assistance with social problems can make a significant contribution to clinical recovery. Other social interventions include the provision of group support, social clubs, occupational therapy and referral to a social worker. Educational programmes, self-help groups, and informed and supportive family members can help improve outcome.

Prognosis

People with major depressive illness are between 1.5 and 2 times more likely to die than not depressed people in the next 16 years, and the risks are not only suicide, but also cardiovascular, which makes effective treatment and prevention imperative.

The majority of patients have recovered by 6 months in primary care and 12 months in secondary care. About a quarter of patients attending hospital with depressive illnesses will have a recurrence within a year, and three-quarters will have a relapse within 10 years. Patients with recurrent depressive illnesses should be offered prevention. This may involve CBT that concentrates on relapse prevention, other forms of psychotherapy, antidepressant medication and advice on lifestyle activities such as regular exercise. Full-dose antidepressants are the most effective prophylaxis in recurrent depressive disorders.

FURTHER READING

Belmaker RH. Bipolar disorders. *New England Journal of Medicine* 2004; **351**: 476–486.

Belmaker RH, Agam G. Major depressive disorder. *New England Journal of Medicine* 2008; **358:** 55–68.

Benazzi F. Focus on bipolar disorder and mixed depression. *Lancet* 2007; **369**: 935–940.

Brockington I. Postpartum psychiatric disorders. *Lancet* 2004; **363**: 303–310.

Reynolds CF 3rd, Dew MA, Pollock BG et al. Maintenance treatment of major depression in old age. New England Journal of Medicine 2006; **354**(11): 1130–1138.

UK ECT Review Group. Efficacy and safety of electroconvulsive therapy in depressive disorders. Review. Lancet 2003; **361**: 799–808.

MANIA AND HYPOMANIA

Mania and hypomania almost always occurs as part of a *bipolar disorder*. The clinical features of mania include a marked elevation of mood, characterized by euphoria, overactivity and disinhibition (Table 22.10). Hypomania is the mild form of mania. Hypomania lasts a shorter time and is less severe, with no psychotic features and less disability. Hypomania can be distinguished from normal happiness by its persistence, non-reactivity (not provoked by good news and not affected by bad news) and social disability. The social disability of mania can be severe, with disinhibited behaviour leading to significant debts (from overspending), lost relationships (from promiscuity or irritability), social ostracism and lost employment (from reckless or disinhibited behaviour).

Some patients have a *rapid cycling* illness, with frequent swings from one mood state to another. A *mixed affective state* occurs when features of mania and depressive illness are seen in the same episode. Cyclothymia is a personality trait with spontaneous swings in mood not sufficiently severe or persistent to warrant another diagnosis.

Table 22.10 Clinical features of mania

Characteristic	Clinical feature
Mood	Elevated or irritable
Talk	Fast, pressurized, flight of ideas
Energy	Excessive
Ideas	Grandiose, self-confident, delusions of wealth, power, influence or of religious significance, sometimes persecutory
Cognition	Disturbance of registration of memories
Physical	Insomnia, mild to moderate weight loss, increased libido
Behaviour	Disinhibition, increased sexual activity, excessive drinking or spending
Hallucinations	Fleeting auditory

Differential diagnosis

Acute intoxication with recreational drugs such as amfetamines, amfetamine derivatives (MDMA: Ecstasy), and cocaine can mimic mania. Long-term use of cannabis can also induce an illness with manic features. In one study a quarter of patients with Cushing's syndrome had mania. Similarly corticosteroids can induce mania less commonly than depressive illness. Dopamine agonists (e.g. bromocriptine) are also known to sometimes induce mania. The excited phase of catatonic schizophrenia can sometimes be mistaken for mania.

Epidemiology

The lifetime prevalence of bipolar affective disorder is 1% across the world. Unlike unipolar depressive illness, it is equally common in men and women, supporting its different aetiology. There is no variation by socio-economic class or race. The mean age of onset is 21; earlier than unipolar depression. The higher prevalence found in divorced people is probably a consequence of the condition.

Aetiology

Genetic

There is strong evidence for a genetic aetiology in this disorder. There is a 60–80% concordance rate in monozygotic twins, compared to 15% in dizygotic twins, suggesting a high rate of heritability. Adoption studies show similar rates, so this high rate is probably genetic and not due to the family environment. Linkage studies have so far proved disappointing, with several suggestive chromosome linkages being found, suggesting there is no single gene with a large effect. Instead it is likely that the condition will prove to be polygenic.

Biochemical

Brain monoamines, e.g. serotonin, seem to be increased in mania. Dexamethasone tends not to suppress cortisol levels in patients with mania, suggesting a similar pattern of non-suppression to that seen in severe depressive illness. Antidepressants (especially tricyclics) may induce a manic episode in someone with bipolar disorder.

Psychological

The effect of life events is much weaker in bipolar compared to unipolar illnesses, with most effect apparent at first onset. Similarly, personality does not seem to be a major influence, in contrast to unipolar depression, although there is some evidence of a link with the creativity and divergent thinking that is an advantage in the right occupation.

Treatment

Acute mania

- First, stop any antidepressant being taken.
- Acute mania is treated with an atypical antipsychotic (neuroleptic), sodium valproate or lithium. The atypical antipsychotics olanzapine, quetiapine and risperidone are particularly recommended, especially with behavioural disturbance. Doses are similar to those used in schizophrenia. The behavioural excitement and overactivity are usually reduced within days, but elation, grandiosity and associated delusions often take longer to respond.
- Severe mania is best treated with a combination of valproic acid or lithium and a neuroleptic, allowing the neuroleptic to be withdrawn after the first 2 or 3 weeks. First attacks of mania usually require treatment for up to 3 months.
- Valproic acid is also helpful in hypomania or in rapidly cycling illnesses (see below). Recent work suggests it can be sometimes helpful to add to the regimen a benzodiazepine, such as lorazepam.

Prevention in bipolar disorders

Since bipolar illnesses tend to be relapsing and remitting, prevention of recurrence is the major therapeutic challenge in management. A patient who has experienced more than two episodes of affective disorder within a 5-year period is likely to benefit from preventative treatments. Recommendations include lithium, olanzapine, and valproic acid (so long as the patient is not a woman at risk of pregnancy).

Lithium (carbonate or citrate) is one of the two main agents used for prophylaxis in patients with repeated episodes of bipolar illness. It is rapidly absorbed into the gastrointestinal tract and more than 95% is excreted by the kidneys; small amounts are found in the saliva, sweat and breast milk. Renal clearance of lithium correlates with renal creatinine clearance. Lithium is a mood-stabilizing drug that prevents mania more than depression. It reduces the frequency and severity of relapses by half and significantly reduces the likelihood of suicide. Its mode of action is unknown, but lithium is known to act on the serotoninergic system. Poor responses to lithium are associated with a negative family history, an unstable premorbid personality, and a rapid cycling illness. Recent pharmacogenetic work suggests that certain polymorphisms may predict response.

Screening prior to starting lithium:

- *Thyroid disease* (free T_4, TSH and thyroid autoantibodies). Lithium interferes with thyroid function and can produce frank hypothyroidism. The presence of thyroid autoantibodies increases the risk.
- *Renal disease* (serum urea and creatinine, estimated creatinine clearance and 24-hour urinary volume). Long-term treatment with lithium causes two renal problems: nephrogenic diabetes insipidus (DI) and reduced glomerular function (see p. 575). The best screen for DI is to ask the patient about polyuria and polydipsia. Tests should include serum urea and creatinine, and 24-hour urinary volume if DI is suspected.

The therapeutic range for prophylaxis is 0.5–1.0 mmol/L. Lithium levels should be checked every 3 months, along with regular thyroid (free T_4 and TSH) and renal function tests.

Patients should carry a lithium card with them at all times, be advised to avoid dehydration, and be warned of drug interactions, such as NSAIDs and diuretics.

Other side-effects of lithium include:

- nausea and diarrhoea
- a fine tremor (15%)
- polyuria and polydipsia (see above)
- weight gain, mainly through increased appetite.

Lithium toxicity begins to occur when the serum concentration exceeds 1.5 mmol/L. Symptoms include drowsiness, nausea, vomiting, blurred vision, a coarse tremor, ataxia and dysarthria. Toxicity is more likely when the patient is dehydrated or with a drug interaction. Such symptoms progress to delirium and convulsions, and coma and death can occur. As a rule, lithium is not advised during pregnancy, particularly in the first trimester, because of an increased risk of fetal malformation (Ebstein's anomaly). Between 25% and 30% of women with a history of bipolar disorder relapse within 2

weeks of delivery. Restarting lithium within 24 hours of delivery (if the mother is prepared to forego breast-feeding) markedly reduces the risk of relapse.

Valproic acid (as the semisodium salt) is recommended both in prophylaxis and treatment of manic states. Second-line treatments include *carbamazepine* and *lamotrigine*. Some patients who do not respond to lithium may respond to these anticonvulsants or a combination of both. Patients with rapid cycling illnesses show a better response to anticonvulsants than to lithium. For antimanic treatment, dosage in the initial stage of treatment will be 200 mg twice daily of carbamazepine, increasing to a normal dose of 800–1000 mg. Other drugs which appear to exercise a prophylactic mood-stabilizing effect include olanzapine and risperidone. Both carbamazepine and valproate can be teratogenic (neural tube defects) and should be avoided in pregnancy. Other side-effects of these drugs are given on page 1143.

FURTHER READING

Goodwin FK, Jamieson KR. *Bipolar Disorders and Recurrent Depression,* 2nd edn. Oxford: Oxford University Press, 2007.

Prognosis

The average duration of a manic episode is 2 months, with 95% making a full recovery in time. Recurrence is the rule in bipolar disorders, with up to 90% relapsing within 10 years.

SUICIDE AND DELIBERATE SELF-HARM (see also p. 933)

Suicide accounts for 2% of male and 1% of female deaths in England and Wales each year, equivalent to a rate of 8 per 100 000. The rate increases with age, peaking for women in their sixties and for men in their seventies. Suicide is the second most common cause of mortality in 15- to 34-year-olds. Approximately 15% of people who have suffered a severe depressive disorder (requiring admission) will eventually commit suicide, with 6% doing so in the 10 years after their first admission. Suicide rates in schizophrenia sufferers are likewise high, being 20–50 times the rate in the general population; 20–40% of people with schizophrenia make suicide attempts, and 9–13% are successful. The highest rates of suicide have been reported in rural southern India (148 per 100 000 in young women and 58 per 100 000 in young men) and in eastern Europe (30–40 per 100 000), while the lowest are those of Spain (3.9 per 100 000) and Greece (2.8 per 100 000), but such variations may reflect differences in reporting, which may be related to religion, as much as genuine differences. Factors that increase the risk of suicide are indicated in Table 22.11.

Table 22.11 Factors that increase the risk of suicide

Male sex
Older age
Living alone
Immigrant status
Recent bereavement, separation or divorce
Recent loss of a job or retirement
Living in a socially disorganized area
Family history of affective disorder, suicide or alcohol abuse
Previous history of affective disorder, alcohol or drug abuse
Previous suicide attempt
Addiction to alcohol or drugs
Severe depression or early dementia
Incapacitating painful physical illness

FURTHER READING

Kim WJ, Singh T. Trends and dynamics of youth suicides in developing countries. *Lancet* 2004; **363**: 1090–1091.

A distinction must be drawn between those who attempt suicide – deliberate self-harm (DSH) – and those who succeed (suicides):

- The majority of cases of DSH occur in people under 35 years of age.
- The majority of suicides occur in people aged over 60.
- Suicides are more common in men, while DSH is more common in women.
- Suicides are more common in older men, although rates are falling. Rates in young men are rising fast throughout the UK and Western Europe.
- Suicides in women are slowly falling in the UK.
- Approximately 90% of cases of DSH involve self-poisoning.
- A formal psychiatric diagnosis usually can be made retrospectively in suicide, but is unusual in DSH.

There is, however, an overlap between DSH and suicide. Between 1% and 2% of people who attempt suicide will kill themselves in the year following DSH. The risk of suicide stays elevated in those with DSH, with 0.5% per annum committing suicide in the following 20 years. In the UK, there are 100 000 cases of DSH each year, and the overwhelming majority of these are seen and treated within accident and emergency departments.

The guidelines (Box 22.11) for the assessment of such patients will help ensure that the risk factors relating to suicide are covered. Indications for referral to a psychiatrist before discharge from hospital are also given.

In general, it is worth trying to interview a family member or close friend and check these points with them. Requests for immediate re-prescription on discharge should be denied, except in cases of essential medication. In such cases, however, only 3 days' supply of medication should be given, and the patient should be requested to report to their general practitioner or to their psychiatric outpatient clinic for further supplies. Occasionally involuntary admission under the Mental Health Act may be required (p. 1222).

ANXIETY DISORDERS

These are conditions in which anxiety dominates the clinical symptoms. They are classified according to whether the anxiety is persistent (general anxiety) or episodic, with the episodic conditions classified according to whether the episodes are regularly triggered by a cue (phobia) or not (panic disorder). The differential diagnoses of anxiety disorders are given in Table 22.12. A patient with one anxiety disorder may well in time develop others.

General anxiety disorder

This occurs in 4–6% of the population and is more common in women. Symptoms are persistent and often chronic. General anxiety disorder (GAD) and its related panic disorder are differential diagnoses for functional somatic syndromes, owing to the many physical symptoms that are caused by these conditions.

Clinical features

The physical and psychological symptoms are outlined in Table 22.13. The patient looks worried, has a tense posture,

Box 22.11 Guidelines for the assessment of patients who harm themselves

Questions to ask: be concerned if positive answer

- Was there a clear precipitant/cause for the attempt?
- Was the act premeditated or impulsive?
- Did the patient leave a suicide note?
- Had the patient taken pains not to be discovered?
- Did the patient make the attempt in strange surroundings (i.e. away from home)?
- Would the patient do it again?

Other relevant factors

- Has the precipitant or crisis resolved?
- Is there continuing suicidal intent?
- Does the patient have any psychiatric symptoms?
- What is the patient's social support system?
- Has the patient inflicted self-harm before?
- Has anyone in the family ever taken their life?
- Does the patient have a physical illness?

Indications for referral to a psychiatrist

Absolute indications include:

- Clinical depression
- Psychotic illness of any kind
- Clearly preplanned suicidal attempt which was not intended to be discovered
- Persistent suicidal intent (the more detailed the plans, the more serious the risk)
- A violent method used.

Other common indications include:

- Alcohol and drug abuse
- Patients over 45 years, especially if male, and young adolescents
- Those with a family history of suicide in first-degree relatives
- Those with serious (especially incurable) physical disease
- Those living alone or otherwise unsupported
- Those in whom there is a major unresolved crisis
- Persistent suicide attempts
- Any patients who give you cause for concern.

Table 22.12 Anxiety disorders – the differential diagnosis

Psychiatric disorder	Physical disorder
Depressive illness	Hyperthyroidism
Obsessive compulsive disorder	Hypoglycaemia
Presenile dementia	Phaeochromocytoma
Alcohol dependence	
Drug dependence	
Benzodiazepine withdrawal	

restless behaviour and a pale and sweaty skin. The patient takes time to go to sleep, and when asleep wakes intermittently with worry dreams. Associated conditions include the hyperventilation syndrome, which is even more common in panic disorder (Box 22.12). The patient will sigh deeply, particularly when talking about the stresses in their life.

Mixed anxiety and depressive disorder

This disorder is probably the commonest mood disorder in primary care, in which there are equal elements of both anxiety and depression, showing how closely associated these two abnormal mood states are.

Table 22.13 Physical and psychological symptoms of anxiety

Physical symptoms

Gastrointestinal

Dry mouth
Difficulty in swallowing
Epigastric discomfort
Aerophagy (swallowing air)
'Diarrhoea' (usually frequency)

Respiratory

Feeling of chest constriction
Difficulty in inhaling
Overbreathing
Choking

Cardiovascular

Palpitations
Awareness of missed beats
Chest pain

Genitourinary

Increased frequency
Failure of erection
Lack of libido

Nervous system

Fatigue
Blurred vision
Dizziness
Sensitivity to noise and/or light
Headache
Sleep disturbance
Trembling

Psychological symptoms

Apprehension and fear
Irritability
Difficulty in concentrating
Distractibility
Restlessness
Depersonalization
Derealization

Panic disorder

Panic disorder is diagnosed when the patient has repeated sudden attacks of overwhelming anxiety, accompanied by severe physical symptoms, usually related to both hyperventilation (Box 22.12) and sympathetic nervous system activity (palpitations, tremor, restlessness and sweating). The lifetime prevalence is 5%. Patients with panic disorder often have catastrophic illness beliefs during the panic attack, such as convictions that they are about to die from a stroke or heart attack. The fear of a stroke is related to dizziness and headache. Fear of a heart attack accompanies chest pain (atypical chest pain). The occasional patient with long-standing attacks may deny feeling anxious and simply report the physical symptoms.

Aetiology

General anxiety and panic disorders occur four or more times as commonly in first-degree relatives of affected patients,

Box 22.12 The hyperventilation syndrome

Features

Panic attacks – fear, terror and impending doom – accompanied by some or all of the following:

- dyspnoea (trouble getting a good breath in)
- palpitations
- chest pain or discomfort
- suffocating sensation
- dizziness
- paraesthesiae in hands and feet
- peri-oral paraesthesiae
- sweating
- carpopedal spasms

Cause

Overbreathing leading to a decrease in $PaCO_2$ and an increase in arterial pH, leading to relative hypocalcaemia.

Diagnosis

A provocation test – voluntary overbreathing for 2 minutes – provokes similar symptoms; rebreathing from a large paper bag relieves them.

Management

- Explanation and reassurance is given.
- The patient is trained in relaxation techniques and slow controlled breathing.
- The patient is asked to breathe into a closed paper bag.

Box 22.13 Phobias

A phobia is an abnormal fear and avoidance of an object or situation.
Phobias are common (8% prevalence), disabling and treatable with behaviour therapy.

suggesting a genetic influence. Sympathetic nervous system overactivity, increased muscle tension and hyperventilation are the common pathophysiological mechanisms. Anxiety is the emotional response to the threat of a loss, whereas depression is the response to the loss itself. There is some evidence that being bullied, with the explicit threats involved, leads to anxiety disorders in adolescents.

Phobic (anxiety) disorders

Phobias are common conditions in which intense fear is triggered by a stimulus, or group of stimuli, that are predictable and normally cause no particular concern to others (e.g. agoraphobia, claustrophobia, social phobia). This leads to avoidance of the stimulus (Box 22.13). The patient knows that the fear is irrational, but cannot control it. The prevalence of all phobias is 8%, with many patients having more than one. Many phobias of 'medical' stimuli exist (e.g. of doctors, dentists, hospitals, vomit, blood and injections) which affect the patient's ability to receive adequate healthcare.

Aetiology

Phobias may be caused by classical conditioning, in which a response (fear and avoidance) becomes conditioned to a previously benign stimulus (a lift), often after an initiating emotional shock (being stuck in a lift). In children, phobias can arise through imagined threats (e.g. stories of ghosts told in the playground). Women have twice the prevalence of most phobias than men. Phobias aggregate in families, with increasing evidence of the importance of genetic factors being published.

Agoraphobia

Translated as 'fear of the market place', this common phobia (4% prevalence) presents as a fear of being away from home, with avoidance of travelling, walking down a road and supermarkets being common cues. This can be a very disabling condition, since the patient can be too unwell to ever leave home, particularly by themselves. It is often associated with *claustrophobia*, a fear of enclosed spaces.

Social phobia

This is the fear and avoidance of social situations: crowds, strangers, parties and meetings. Public speaking would be the sufferer's worst nightmare. It is suffered by 2% of the population.

Simple phobias

The commonest is the phobia of spiders (arachnophobia), particularly in women. The prevalence of simple phobias is 7% in the general population. Other common phobias include insects, moths, bats, dogs, snakes, heights, thunderstorms and the dark. Children are particularly phobic about the dark, ghosts and burglars, but the large majority grow out of these fears.

Treatment of anxiety disorders

Psychological treatments

For many people with brief episodes of an anxiety disorder, a discussion with a doctor concerning the nature of anxiety is usually sufficient.

- *Relaxation techniques* can be effective in mild/moderate anxiety. Relaxation can be achieved in many ways, including complementary techniques such as meditation and yoga. Conventional relaxation training involves slowing down the rate of breathing, muscle relaxation and mental imagery.
- *Anxiety management* training involves two stages. In the first stage, verbal cues and mental imagery are used to arouse anxiety to demonstrate the link with symptoms. In the second stage, the patient is trained to reduce this anxiety by relaxation, distraction and reassuring self-statements.
- *Biofeedback* is useful for showing patients that they are not relaxed, even when they fail to recognize it, having become so used to anxiety. Biofeedback involves feeding back to the patient a physiological measure that is abnormal in anxiety. These measures may include electrical resistance of the skin of the palm, heart rate, muscle electromyography or breathing pattern.
- *Behaviour therapies* are treatments that are intended to change behaviour and thus symptoms. The most common and successful behaviour therapy (with 80% success in some phobias) is graded exposure, otherwise known as systematic desensitization. First, the patient rates the phobia into a hierarchy or 'ladder' of worsening fears (e.g. in agoraphobia: walking to the front door with a coat on; walking out into the garden; walking to the end of the road). Second, the patient practises exposure to the least fearful stimulus until no fear is felt. The patient then moves 'up the ladder' of fears until they are cured.

- *Cognitive behaviour therapy* (CBT) (see p. 1204) is the treatment of choice for panic disorder and general anxiety disorder because the therapist and patient need to identify the mental cues (thoughts and memories) that may subtly provoke exacerbations of anxiety or panic attacks. CBT also allows identification and alteration of the patient's 'schema', or way of looking at themselves and their situation, that feeds anxiety.

Drug treatments

Initial 'drug' treatment should involve advice to gradually cease taking anxiogenic recreational drugs such as caffeine and alcohol (which can cause a rebound anxiety and withdrawal). Prescribed drugs used in the treatment of anxiety can be divided into two groups: those that act primarily on the central nervous system, and those that block peripheral autonomic receptors.

- *Benzodiazepines* are centrally acting anxiolytic drugs. They bind to specific receptors that stimulate release of the inhibitory transmitter γ-aminobutyric acid (GABA). Diazepam (5 mg twice daily, up to 10 mg three times daily in severe cases), alprazolam (250–500 μg three times daily) and chlordiazepoxide have relatively long half-lives (20–40 hours) and are used as anti-anxiety drugs in the short term. Side-effects include sedation and memory problems, and patients should be advised not to drive while on treatment. They can cause dependence and tolerance within 4–6 weeks, particularly in dependent personalities. The withdrawal syndrome (Table 22.14) can occur after just 3 weeks of continuous use and is particularly severe when high doses have been given for a longer time. Thus, if a benzodiazepine drug is prescribed for anxiety, it should be given in as low a dose as possible, preferably on an 'as necessary' basis, and for not more than 2–4 weeks. A withdrawal programme from chronic use includes changing the drug to the long-acting diazepam, followed by a very gradual reduction in dosage.
- *Most SSRIs* (e.g. fluoxetine, paroxetine, sertraline, escitalopram, citalopram) are useful symptomatic treatments for general anxiety and panic disorders, as well as some phobias (social phobia). Imipramine and clomipramine are alternative symptomatic treatments for panic disorder and GAD. Treatment response is often delayed several weeks; a trial of treatment should last 3 months.
- Many of the symptoms of anxiety are due to an increased or sustained release of epinephrine (adrenaline) and norepinephrine (noradrenaline) from the adrenal medulla and sympathetic nerves. Thus, beta-blockers such as propranolol (20–40 mg two or three times daily) are effective in reducing peripheral symptoms such as palpitations, tremor and tachycardia, but do not help central symptoms such as anxiety.

Table 22.14 Withdrawal syndrome with benzodiazepines

Insomnia
Anxiety
Tremulousness
Muscle twitchings
Perceptual distortions
Hallucinations (which may be visual)
Hypersensitivities (light, sound, touch)
Convulsions

Acute stress reactions and adjustment disorders

Acute stress reaction

This occurs in individuals without any other psychiatric disorder, in response to exceptional physical and/or psychological stress. While severe, such a reaction usually subsides within days. The stress may be an overwhelming traumatic experience (e.g. accident, battle, physical assault, rape) or a sudden change in the social circumstances of the individual, such as a bereavement. Individual vulnerability and coping capacity play a role in the occurrence and severity of an acute stress reaction, as evidenced by the fact that not all people exposed to exceptional stress develop symptoms. Symptoms usually include an initial state of feeling 'dazed' or numb, with inability to comprehend the situation. This state may be followed either by further withdrawal from the situation or by anxiety and overactivity. No treatments beyond reassurance and support are normally necessary.

Adjustment disorder

This disorder can follow an acute stress reaction and is common in the general hospital. This is a more prolonged (up to 6 months) emotional reaction to a significant life event, with low mood joining the initial shock and consequent anxiety, but not of sufficient severity or persistence to fulfil a diagnosis of a mood or anxiety disorder. Supportive counselling is usually a successful treatment, allowing facilitation of unexpressed feelings, elucidation of unspoken fears, and education about the likely future.

Pathological (abnormal) grief

This is a particular kind of adjustment disorder. It can be characterized as excessive and/or prolonged grief, or even absent grieving with abnormal denial of the bereavement. Usually a relative will be stuck in grief, with insomnia and repeated dreams of the dead person, anger at doctors or even the patient for dying, consequent guilt in equal measure, and an inability to 'say good-bye' to the loved person by dealing with their effects. *Guided mourning* uses cognitive and behavioural techniques to allow the relative to stop grieving and move on in life.

Normal grief immediately follows bereavement, is expressed openly, and allows a person to go through the social ceremonies and personal processes of bereavement. The three stages are, first, shock and disbelief, second, the emotional phase (anger, guilt and sadness) and, third, acceptance and resolution. This normal process of adjustment may take up to a year, with movement between all three stages occurring in a sometimes haphazard fashion.

Post-traumatic stress disorder (PTSD)

This arises as a protracted response to a stressful event or situation of an exceptionally threatening nature, likely to cause pervasive distress in almost anyone. Causes include natural or human disasters, war, serious accidents, witnessing the violent death of others, being the victim of sexual abuse, rape, torture, terrorism or hostage-taking. Predisposing factors such as personality, previously unresolved

traumas, or a history of psychiatric illness may prolong the course of the syndrome. These factors are neither necessary nor sufficient to explain its occurrence, which is most related to the intensity of the trauma, the proximity of the patient to the traumatic event, and how prolonged or repeated it was. Functional brain scan research suggests a possible neurophysiological relationship with obsessive compulsive disorder (p. 1210).

Clinical features

The typical symptoms of PTSD include:

- '*flashbacks*' – repeated vivid reliving of the trauma in the form of intrusive memories, often triggered by a reminder of the trauma
- *insomnia*, usually accompanied by nightmares, the nocturnal equivalent of flashbacks
- *emotional blunting*, emptiness or 'numbness', alternating with
- *intense anxiety* at exposure to events that resemble an aspect of the traumatic event, including anniversaries of the trauma
- *avoidance* of activities and situations reminiscent of the trauma
- *emotional detachment* from other people
- *hypervigilance* with autonomic hyperarousal and an enhanced startle reaction.

This clinical picture represents the severe end of a spectrum of emotional reactions to trauma, which might alternatively take the form of an adjustment or mood disorder. The course is often fluctuating but recovery can be expected in two-thirds of cases at the end of the first year. Complications include depressive illness and alcohol misuse. In a small proportion of cases the condition may show a chronic course over many years and a transition to an enduring personality change.

Treatment and prevention

Compulsory *psychological debriefing* immediately after a trauma does not prevent PTSD and may be harmful. Prevention is better achieved by the support offered by others who were also involved. Trauma focused CBT is often effective. *Eye movement desensitization and reprocessing* (EMDR) is equally effective treatment and may require fewer sessions. SSRIs and venlafaxine have a place in the management of chronic PTSD, but drop-out from pharmacotherapy is common.

FURTHER READING

Roy-Byrne PP, Craske M, Stein M. Panic disorder. *Lancet* 2006; **368**: 1023–1032.

Stein D, Seedat S, Iversen A et al. Post-traumatic stress disorder. *Lancet* 2006; **369**: 139–144.

Tyrer P, Baldwin D. Generalised anxiety disorder. *Lancet* 2006; **368**: 2156–2166.

The adult consequences of childhood sexual and physical abuse

Estimates of the prevalence of childhood sexual abuse vary depending on definition but there is reasonable evidence that 20% of women and 10% of men suffered significant, coercive and inappropriate sexual activity in childhood. The abuser is usually a member of the family or known to the child, and preadolescent girls are at greatest risk. The likelihood of long-term consequences is determined by:

- an earlier age of onset
- the severity of the abuse
- the repeated nature and duration of the period of abuse
- the association with physical abuse.

Consequent adult psychiatric disorders include depressive illness, substance misuse, eating disorders, borderline personality disorder and deliberate self-harm. Other negative outcomes include a decline in socio-economic status, sexual problems, prostitution and difficulties in forming adult relationships.

Repeated childhood physical and emotional abuse or neglect may also affect emotional and personality development, predisposing the adult to similar psychiatric disorders. Those with repeated abuse are more likely to have long-term physical stress related consequences, such as hypothalamic–pituitary–adrenal axis downregulation and smaller brain hippocampal sizes.

Psychodynamic psychotherapy

Psychodynamic psychotherapy is derived from psychoanalysis and is based on a number of key analytical concepts. These include Freud's ideas about psychosexual development, defence mechanisms, free association as the method of recall, and the therapeutic techniques of interpretation, including that of transference, defences and dreams. Such therapy usually involves once-weekly sessions, the length of treatment varying between 3 months and 2 years. The long-term aim of such therapy is twofold: symptom relief and personality change. Psychodynamic psychotherapy is classically indicated in the treatment of unresolved conflicts in early life, as might be found in non-psychotic and personality disorders, but there is no convincing evidence concerning its superiority over alternative forms of treatment.

Cognitive analytical therapy is an integration of cognitive behaviour therapy and psychodynamic therapy. It is a short-term therapy that involves the patient and therapist recognizing the origins of a recurrent problem, reformulating how it continues to occur, and revising other ways of coping and internalizing it, using both the transference of the patient–therapist relationship and behavioural experiments.

OBSESSIVE-COMPULSIVE DISORDER

Obsessive-compulsive disorder (OCD) is characterized by obsessional ruminations and compulsive rituals. It is particularly associated with and/or secondary to both depressive illness and *Gilles de la Tourette syndrome* (p. 1150). The prevalence is between 1% and 2% in the general population, and patients often do not seek help. There is an equal distribution by gender, and the mean age of onset ranges from 20 to 40 years.

Clinical features

The obsessions and compulsions are time consuming and intrusive so that they affect functioning and cause considerable distress. Ruminations are often unpleasant repetitive thoughts, out of character, such as being dirty or violent. This can lead to a constant need to check that everything and everyone is alright and that things have been done correctly, and reassurance cannot remove the doubt that persists. Some rituals are derived from superstitions, such as actions repeated a fixed number of times, with the need to start again if interrupted. When severe and primary, OCD can last for many years and may be resistant to treatment. However, obsessional symptoms commonly occur in other disorders, most notably depressive illness and schizophrenia, and remit with the resolution of the primary disorder.

Minor degrees of obsessional symptoms and compulsive rituals or superstitions are common in people who are not ill or in need of treatment, particularly in times of stress.

The mildest grade is that of obsessional personality traits such as over-conscientiousness, tidiness, punctuality and other attitudes and behaviours indicating a strong tendency towards conformity and inflexibility. Such individuals are *perfectionists* who are intolerant of shortcomings in themselves and others, and take pride in their high standards. When such traits are so marked that they dominate other aspects of the personality, in the absence of clear-cut OCD, the diagnosis is obsessional (anankastic) personality (see p. 1222).

Aetiology

Genetic

OCD is found in 5–7% of the first-degree relatives. Twin studies showed 80–90% concordance in monozygotic twins and about 50% in dizygotic twins.

Basal ganglia dysfunction

OCD is associated with a number of neurological disorders involving dysfunction of the striatum, including Parkinson's disease, Sydenham's and Huntington's chorea. OCD can follow head trauma. Neuroimaging suggests that abnormalities exist in the frontal lobe and basal ganglia (Fig. 22.4). Hyperactivity of the orbitofrontal cortex has been a consistent finding in brain imaging research on OCD patients. Other work suggests the caudate nucleus is smaller than in healthy controls.

Serotonin

Serotonin function is probably abnormal in patients with OCD. Serotonin reuptake inhibitors are effective drugs. Postsynaptic serotonin receptor hypersensitivity may follow chronically low levels of synaptic serotonin.

Conditioning

This suggests that compulsive rituals are classically conditioned avoidance responses, which therefore lend themselves to treatment with graded exposure therapy.

Treatment

Psychological treatments

A behaviour therapy that is particularly effective for rituals is *response prevention*. Patients are instructed not to carry out their rituals. There is an initial rise in distress but with persistence both the rituals and the distress diminish. Patients are encouraged to practise response prevention, while returning to situations that normally make them worse.

Modelling involves the therapist demonstrating to the patient what is required and encouraging the patient to follow this example. In the case of hand-washing rituals, this might involve holding an allegedly contaminated object and carrying out other activities without washing, the patient being encouraged to follow suit.

Thought stopping can reduce obsessional ruminations. The patient is taught to arrest the obsessional thought by arranging a sudden intrusion (e.g. snapping an elastic band, clicking the fingers).

Cognitive behaviour therapy combines repeated exposure to the provocative stimulus and thought stopping techniques.

Physical treatment

Selective serotonin reuptake inhibitors are the mainstay of drug treatment. Their efficacy is independent of their antidepressant action but the doses required are usually some 50–100% higher than those effective in depression. Three months' treatment with high doses may be necessary for a positive response. Positive correlations between reduced severity of OCD and decreased orbitofrontal and caudate metabolism following behavioural and SSRI treatments have been demonstrated in a number of studies. Clomipramine (a tricyclic) is an alternative agent commonly used in the UK.

Psychosurgery

Psychosurgery is very occasionally recommended in cases of chronic and severe OCD that has not responded to other treatments. The development of stereotactic techniques has led to the replacement of the earlier, crude leucotomies with more precise surgical interventions such as subcaudate tractotomy and cingulotomy, with small yttrium radioactive implants, which induce lesions in the cingulate area or the ventromedial quadrant of the frontal lobe. Psychosurgery is performed only in a few specialist centres in the UK, and formal and detailed consent requirements are laid down in the appropriate mental health act.

Prognosis

Two-thirds of cases improve within a year. The remainder run a fluctuating or persistent course. The prognosis is worse when the personality is anankastic and the OCD is primary and severe.

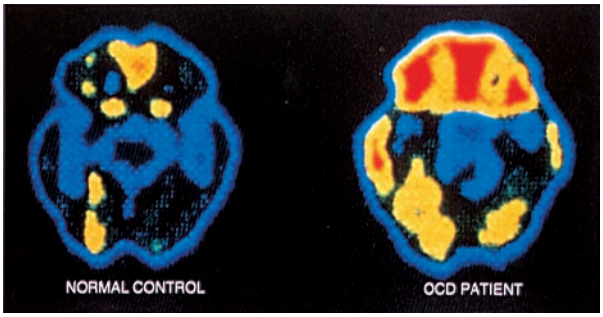


Fig. 22.4 **PET images of (left) a normal patient and (right) an obsessive-compulsive disorder (OCD) patient.** The right image shows the hyperactivity of the orbitofrontal cortex which is a consistent finding in this condition. From Baxter R, Phelps ME, Mazziotta JC et al. Local cerebral glucose metabolic rates in obsessive-compulsive disorder. A comparison with rates in unipolar depression and in normal controls. *Archives of General Psychiatry* 1987; **44**: 211, with permission.

FURTHER READING

Jenike MA. Obsessive-compulsive disorder. *New England Journal of Medicine* 2004; **350**: 259–265.

ALCOHOL MISUSE AND DEPENDENCE

A wide range of physical, social and psychiatric problems are associated with excessive drinking. Alcohol misuse occurs when a patient is drinking in a way that regularly causes problems to the patient or others.

- *The problem drinker* is one who causes or experiences physical, psychological and/or social harm as a consequence of drinking alcohol. Many problem drinkers, while heavy drinkers, are not physically addicted to alcohol.
- *Heavy drinkers* are those who drink significantly more in terms of quantity and/or frequency than is safe to do so long term.

- *Binge drinkers* are those who drink excessively in short bouts, usually 24–48 hours long, separated by often quite lengthy periods of abstinence. Their overall monthly or weekly alcohol intake may be relatively modest.
- *Alcohol dependence* is defined by a physical dependence on or addiction to alcohol. The term 'alcoholism' is a confusing one with off-putting connotations of vagrancy, 'meths' drinking and social disintegration. It has been replaced by the term 'alcohol dependence syndrome'.

Epidemiology of alcohol misuse

Twenty per cent of men and 10% of women drink more than double the recommended limits of 3 units a day of alcohol for men and 2 units for women in the UK. Some 4% of men and 2% of women report alcohol withdrawal symptoms, suggesting dependence. Approximately one in five male admissions to acute medical wards is directly or indirectly due to alcohol. Between 33% and 40% of accident and emergency attenders have blood alcohol concentrations above the present UK legal limit for driving. People with serious drinking problems have a two to three times increased risk of dying compared to members of the general population of the same age and sex.

Table 22.15 provides an approximate estimate of what can be expected in an average individual in the way of behavioural impairment resulting from a particular blood alcohol level. The amount of alcohol is measured for convenience in units which contain about 8 g of absolute alcohol and raises the blood alcohol concentration by about 15–20 mg/dL, the amount that is metabolized in 1 hour. One unit of alcohol is found in half a pint of ordinary beer (3.5% alcohol by volume – ABV) and 125 mL of 9% wine. However, some beer and most lager is now 5% ABV – 3 units per pint. Wine is often 13% ABV and sold in 175 mL glasses – 2–3 units per glass.

Detection

Alcohol misuse should be suspected in any patient presenting with one or more physical problems commonly associated with excessive drinking (see p. 238). Alcohol misuse may also be associated with a number of psychiatric symptoms/disorders and social problems (Table 22.16).

Guidelines

The patient's frequency of drinking and quantity drunk during a typical week should be established. Alcohol consumption can be assessed on the basis of units of alcohol.

- Drinking up to 21 units of alcohol a week for men and 14 units for women carries no long-term health risk.
- There is unlikely to be any long-term health damage with 21–35 units (men) and 14–24 units (women), provided the drinking is spread throughout the week.
- Beyond 36 units a week in men and 24 units a week in women, damage to health becomes increasingly likely.
- Drinking above 50 units a week in men (35 units in women) is a definite health hazard.

Table 22.15 Behavioural effects of alcohol

Rising blood alcohol (mg/dL)	Expected effect
20–99	Impaired coordination, euphoria
100–199	Ataxia, poor judgement, labile mood
200–299	Marked ataxia and slurred speech; poor judgement, labile mood, nausea and vomiting
300–399	Stage 1 anaesthesia, memory lapse, labile mood
400+	Respiratory failure, coma, death

Table 22.16 Common alcohol-related psychological and social problems

Psychological	Social
Depression	Domestic violence
Anxiety	Marital and sexual difficulties
Memory problems	Child abuse
Delirium tremens	Employment problems
Attempted suicide	Financial difficulties
Suicide	Accidents at home, on the roads, at work
Pathological jealousy	Delinquency and crime
	Homelessness

Diagnostic markers of alcohol misuse

Laboratory parameters indicating alcohol misuse in recent weeks include elevated γ-glutamyl transpeptidase (γ-GT) and mean corpuscular volume (MCV). Blood or breath alcohol tests are useful in anyone suspected of very recent drinking.

Alcohol dependence syndrome

Dependence is a pattern of repeated self-administration that usually results in tolerance, withdrawal and compulsive drug-taking behaviour, the essential element of which is the continued use of the substance despite significant substance-related problems. Figure 22.5 outlines the main characteristics of the syndrome but these do not necessarily present in any particular order. Symptoms of alcohol dependence in a typical order of occurrence are shown in Table 22.17. Diagnostic criteria for alcohol withdrawal syndrome are shown in Table 22.18.

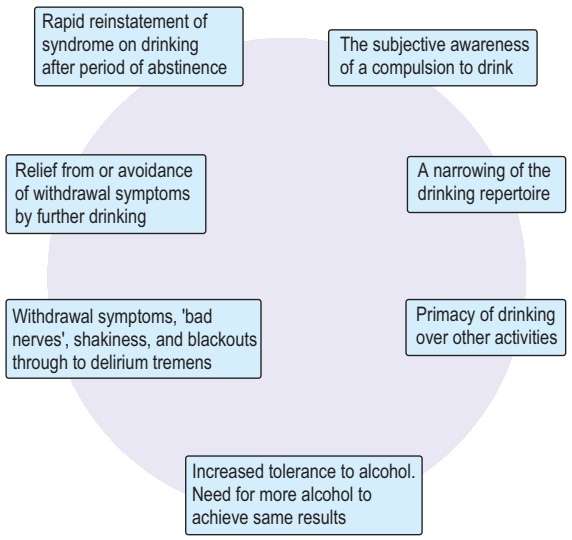


Fig. 22.5 **Elements of the alcohol dependence syndrome.**

Table 22.17 Symptoms of alcohol dependence

Unable to keep a drink limit
Difficulty in avoiding getting drunk
Spending a considerable time drinking
Missing meals
Memory lapses, blackouts
Restless without drink
Organizing day around drink
Trembling after drinking the day before
Morning retching and vomiting
Sweating excessively at night
Withdrawal fits
Morning drinking
Increased tolerance
Hallucinations, frank delirium tremens

Table 22.18 Diagnostic criteria for alcohol withdrawal syndrome

Any three of the following:
Tremor of outstretched hands, tongue or eyelids
Sweating
Nausea, retching or vomiting
Tachycardia or hypertension
Anxiety
Psychomotor agitation
Headache
Insomnia
Malaise or weakness
Transient visual, tactile or auditory hallucinations or illusions
Grand mal convulsions

The course of the alcohol dependence syndrome

About 25% of all cases of alcohol misuse will lead to chronic alcohol dependence. This most commonly ends in social incapacity, death or abstinence. Alcohol dependence syndrome usually develops after 10 years of heavy drinking (3–4 years in women). In some individuals who use alcohol to alter consciousness, obliterate conscience and defy social mores, dependence and loss of control may appear in only a few months or years.

Delirium tremens (DTs)

Delirium tremens is the most serious withdrawal state and occurs 1–3 days after alcohol cessation, so is commonly seen a day or two after admission to hospital. Patients are disorientated, agitated, and have a marked tremor and visual hallucinations (e.g. insects or small animals coming menacingly towards them). Signs include sweating, tachycardia, tachypnoea and pyrexia. Complications include dehydration, infection, hepatic disease or the Wernicke–Korsakoff syndrome (p. 221).

Causes of alcohol dependence

Genetic factors. Sons of alcohol-dependent people who are adopted by other families are four times more likely to develop drinking problems than are the adopted sons of non-alcohol misusers. Genetic markers include dopamine-2 receptor allele A1, alcohol dehydrogenase subtypes and monoamine oxidase B activity, but they are not specific.

Environmental factors. One in ten boys who grow up in a household where neither parent misused alcohol subsequently become alcohol dependent, compared with one in four of those reared by alcohol-misusing fathers and one in three of those reared by alcohol-misusing mothers.

Biochemical factors. Several factors have been suggested, including abnormalities in alcohol dehydrogenase, neurotransmitter substances and brain amino acids, such as GABA. There is no conclusive evidence that these or other biochemical factors play a causal role.

Psychiatric illness. This is an uncommon cause of addictive drinking but it is a treatable one. Some depressed patients drink excessively in the hope of raising their mood. Patients with anxiety states or phobias are also at risk.

Excess consumption in society. The prevalence of alcohol dependence and problems correlates with the general level (per capita consumption) of alcohol use in a society. This, in turn, is determined by factors that may control overall consumption – including price, licensing laws, availability, and the societal norms concerning the use and misuse of alcohol.

Treatment

Psychological treatment of problem drinking

Successful identification at an early stage can be a helpful intervention in its own right. It should lead to:

- the provision of information concerning safe drinking levels
- a recommendation to cut down where indicated
- simple support and advice concerning associated problems.

Such a brief intervention is effective in its own right. Successful alcohol misuse treatment involves *motivational enhancement (motivational therapy),* feedback, education about adverse effects of alcohol, and agreeing drinking goals. A motivational approach is based on five stages of change: precontemplation, contemplation, determination, action and maintenance. The therapist uses motivational interviewing and reflective listening to allow the patient to persuade himself along the five stages to change.

This technique, cognitive behaviour therapy and 12-step facilitation (as used by Alcoholics Anonymous (AA)) have all been shown to reduce harmful drinking. With addictive drinking, self-help group therapy, which involves the long-term support by fellow members of the group (e.g. AA), is helpful in maintaining abstinence. Family and marital therapy involving both the alcohol misuser and spouse may also be helpful. Families of drinkers find meeting others in a similar situation helpful (Al-Anon).

Drug treatments of problem drinking

Alcohol withdrawal and DTs

Addicted drinkers often experience considerable difficulty when they attempt to reduce or stop their drinking. Withdrawal symptoms are a particular problem and delirium tremens needs urgent treatment (Box 22.14). In the absence of DTs, alcohol withdrawal can be treated on an outpatient basis, using one of the fixed schedules in Box 22.14, so long as the patient attends daily for medication and monitoring,

Box 22.14 Management of delirium tremens (DTs)

General measures

Admit the patient to a medical bed.
Correct electrolyte abnormalities and dehydration.
Treat any co-morbid disorder (e.g. infection).
Give parenteral thiamin slowly (250 mg daily for 3–5 days in the absence of Wernicke–Korsakoff (W–K) syndrome).
Give parenteral thiamin slowly (500 mg daily for 3–5 days with W–K encephalopathy. NB beware anaphylaxis).
Give prophylactic phenytoin or carbamazepine, if previous history of withdrawal fits.

Specific drug treatment

One of the following orally:

- Diazepam 10–20 mg
- Chlordiazepoxide 30–60 mg

Repeat 1 hour after last dose depending on response.

Fixed-schedule regimens

- Diazepam 10 mg every 6 hours for 4 doses, then 5 mg 6-hourly for 8 doses OR
- Chlordiazepoxide 30 mg every 6 hours for 4 doses, then 15 mg 6-hourly for 8 doses

Provide additional benzodiazepine when symptoms and signs are not controlled.

Table 22.19 Commonly used drugs of misuse and dependence

Stimulants
Methylphenidate
Phenmetrazine
Phencyclidine ('angel dust')
Cocaine
Amfetamine derivates
Ecstasy (MDMA)
Hallucinogens
Cannabis preparations
Solvents
Lysergic acid diethylamide (LSD)
Mescaline
Narcotics
Morphine
Heroin
Codeine
Pethidine
Methadone
Tranquillizers
Barbiturates
Benzodiazepines

and has good social support. Outpatient schedules are sometimes given over 5 days. Long-term treatment with benzodiazepines should not be prescribed in those patients who continue to misuse alcohol. Some alcohol misusers add dependence on diazepam or clomethiazole to their problems.

Drugs for prevention of alcohol dependence

Naltrexone, the opioid antagonist (50 mg per day), reduces the risk of relapse into heavy drinking and the frequency of drinking. *Acamprosate* (1–2 g per day) acts on several receptors including those for GABA, norepinephrine (noradrenaline) and serotonin. There is reasonable evidence that it reduces drinking frequency. Neither drug seems particularly helpful in maintaining abstinence. Both drug effects are enhanced by combining them with counselling, but their moderate efficacy precludes regular use.

Drugs such as *disulfiram* react with alcohol to cause unpleasant acetaldehyde intoxication and histamine release. A daily maintenance dose means that the patient must wait until the disulfiram is eliminated from the body before drinking safely. There is mixed evidence of efficacy.

One trial has suggested that fluoxetine is helpful in the treatment of patients who have both a depressive illness and alcohol dependence.

Oral thiamin (300 mg a day) can prevent Wernicke–Korsakoff syndrome (p. 221) in heavy drinkers.

Outcome

Research suggests that 30–50% of alcohol-dependent drinkers are abstinent or drinking very much less up to 2 years following traditional intervention. It is too early to be certain of the long-term outcome of patients treated with the latest psychological and pharmacological therapies.

FURTHER READING

Anton RF. Naltrexone for the management of alcohol dependence. *New England Journal of Medicine* 2008; **359:** 715–721.

Moss M, Burnham EL. Alcohol abuse in the critically ill patient. *Lancet* 2006; **368**: 2231–2242.

DRUG MISUSE AND DEPENDENCE

In addition to alcohol and nicotine, there are a number of psychotropic substances that are taken for their effects on mood and other mental functions (Table 22.19).

Causes of drug misuse

There is no single cause of drug misuse and/or dependence. Three factors appear commonly, in a similar way to alcohol problems:

- the availability of drugs
- a vulnerable personality
- social pressures, particularly from peers.

Once regular drug-taking is established, pharmacological factors determine dependence.

Solvents

One per cent of adolescents in the UK sniff solvents for their intoxicating effects. Tolerance develops over weeks or months. Intoxication is characterized by euphoria, excitement, a floating sensation, dizziness, slurred speech and ataxia. Acute intoxication can cause amnesia and visual hallucinations. About 100 young people die in the UK each year from use of these drugs.

Amfetamines and related substances

These have temporary stimulant and euphoriant effects that are followed by fatigue and depression, with the latter sometimes prolonged for weeks. Psychological rather than true physical dependence occurs with 'Speed'. In addition to a

manic-like presentation, amfetamines can produce a paranoid psychosis indistinguishable from schizophrenia.

Ecstasy

'Ecstasy' (E, white burger, white dove) is the street name for 3,4-methylenedioxy-methamfetamine (MDMA), a psychoactive phenylisopropylamine, synthesized as an amfetamine derivative. It is a psychodelic drug which is often used as a 'dance drug'. It has a brief duration of action (4–6 hours). There is evidence that repeated use of MDMA can cause permanent neurotransmitter changes in the brain. Deaths have been reported from malignant hyperpyrexia and dehydration. Acute renal and liver failure can occur.

Cocaine

Cocaine is a central nervous system stimulant (with similar effects to amfetamines) derived from *Erythroxylon coca* trees grown in the Andes. In purified form it may be taken by mouth, snorted or injected. If cocaine hydrochloride is converted to its base ('crack') it can be smoked. This causes an intense stimulating effect, and 'free-basing' is common. Compulsive use and dependence occur more frequently among users who are free-basing. Dependent users take large doses and alternate between withdrawal phenomena of depression, tremor and muscle pains, and the hyperarousal produced by increasing doses. Prolonged use of high doses produces irritability, restlessness, paranoid ideation and occasionally convulsions. Persistent sniffing of the drug can cause perforation of the nasal septum. Overdoses cause death through myocardial infarction, cerebrovascular disease, hyperthermia and arrhythmias (p. 942).

Hallucinogenic drugs

Hallucinogenic drugs, such as lysergic acid diethylamide (LSD) and mescaline, produce distortions and intensifications of sensory perceptions, as well as frank hallucinations in acute intoxication. Psychosis is a long-term complication.

Cannabis

Cannabis (grass, pot, skunk, spliff, reefer) is a drug widely used in some subcultures. It is derived from the dried leaves and flowers of the plant *Cannabis sativa*. It can cause tolerance and dependence. Hashish is the dried resin from the flower tops whilst marijuana refers to any part of the plant. The drug, when smoked, seems to exaggerate the pre-existing mood, be it depression, euphoria or anxiety. It has specific analgesic properties. Cannabis use has increased in the UK recently, especially use of more potent 'skunk'. An amotivational syndrome with apathy and memory problems has been reported with chronic daily use. Cannabis may of itself sometimes cause psychosis in the right circumstances (see below).

Tranquillizers

Drugs causing dependence include barbiturates and benzodiazepines. Benzodiazepine dependence is common and may be iatrogenic, when the drugs are prescribed and not discontinued. Discontinuing treatment with benzodiazepines may cause withdrawal symptoms (see Table 22.14). For this reason, withdrawal should be supervised and gradual.

Opiates

Physical dependence occurs with morphine, heroin and codeine as well as with synthetic and semisynthetic opiates such as methadone, pethidine and fentanyl. These substances display cross-tolerance – the withdrawal effects of one are reduced by administration of another. The psychological effect of such substances is of a calm, slightly euphoric mood associated with freedom from physical discomfort and a flattening of emotional response. This is believed to be due to the attachment of morphine and its analogues to endorphin receptors in the CNS. Tolerance to this group of drugs is rapidly developed and marked, but is rapidly lost following abstinence. The opiate withdrawal syndrome consists of a constellation of signs and symptoms (Table 22.20) that reaches peak intensity on the second or third day after the last dose of the opiate. These rapidly subside over the next 7 days. Withdrawal is dangerous in patients with heart disease or other chronic debilitating conditions.

Opiate addicts have a relatively high mortality rate, owing to both the ease of accidental overdose and the blood-borne infections associated with shared needles. Heart disease (including infective endocarditis), tuberculosis and AIDS are common causes of death, while the complications of hepatitis B and C are also common.

Treatment of chronic misuse

Blood and urine screening for drugs is required in circumstances where drug misuse is suspected (Table 22.21). When a patient with an opiate addiction is admitted to hospital for another health problem, advice should be sought from a psychiatrist or the patient's drug clinic regarding management of their addiction while an inpatient.

The treatment of chronic dependence is directed towards helping the patient either to live without drugs or to prevent secondary ill-health. Patients need help and advice in order to avoid a withdrawal syndrome. Some patients with opiate

Table 22.20 Opiate withdrawal syndrome

Yawning	12–16 hours after last dose of opiate
Rhinorrhoea	
Lacrimation	
Pupillary dilatation	
Sweating	
Piloerection	
Restlessness	
Muscular twitches	24–72 hours after last dose of opiate
Aches and pains	
Abdominal cramps	
Vomiting	
Diarrhoea	
Hypertension	
Insomnia	
Anorexia	
Agitation	
Profuse sweating	

Ask the authors
www.studentconsult.com

Table 22.21 Length of time urine toxicology screens are likely to remain positive after abstinence

Substance	Usual time positive
Amfetamines	48 hours
Barbiturates	
Short-acting	24 hours
Long-acting	7+ days
Benzodiazepines	3+ days
Cannabinols	5+ days
Cocaine	3+ days
Codeine	48 hours
Morphine	48 hours

addiction who cannot manage abstinence may be maintained on oral methadone. In the UK, only specially licensed doctors may legally prescribe heroin and cocaine to an addict for maintenance treatment of addiction. An overdose should be treated immediately with the opioid antagonist naloxone.

FURTHER READING

Cami J, Farre M. Drug addiction. *New England Journal of Medicine* 2004; **349**: 975–986.

Moore TH, Zammit S, Lingford-Hughes A et al. Cannabis use and risk of psychotic or affective mental health outcomes: a systematic review. *Lancet* 2007; **370**: 319–328.

Drug psychoses

Drug-induced psychosis has been reported with amfetamine and its derivatives, cocaine and hallucinogens. It can occur acutely after drug use, but is more usually associated with chronic misuse. Psychoses are characterized by vivid hallucinations (usually auditory, but often in more than one sensory modality), misidentifications, delusions and/or ideas of reference (often of a persecutory nature), psychomotor disturbances (excitement or stupor) and an abnormal affect. ICD-10 requires that the condition occurs within 2 weeks and usually within 48 hours of drug use and that it should persist for more than 48 hours but not more than 6 months.

Cannabis use can result in anxiety, depression or hallucinations. Manic-like psychoses occurring after long-term cannabis use have been described, but seem more likely to be related to the toxic effects of heavy ingestion. However, a meta-analysis suggests that ever taking cannabis increases the risk of psychosis by 40%, daily cannabis use doubles the risk of psychosis, and that 14% of schizophrenia in the UK would be prevented if cannabis use ceased. The risk is higher in people taking cannabis early in their lives and with a heavy consumption.

SCHIZOPHRENIA

The group of illnesses conventionally referred to as 'schizophrenia' is diverse in nature and covers a broad range of perceptual, cognitive and behavioural disturbances. The point prevalence of the condition is 0.5% throughout the world, with equal gender distribution. A physician primarily needs to know how to recognize schizophrenia, what problems it might present with in the general hospital, and how it is treated.

Causes

No single cause has been identified. Schizophrenia is likely to be a disease of neurodevelopmental disconnection caused by an interaction of genetic and multiple environmental factors that affect brain development. It is likely that daily cannabis use is a risk factor. The genetic aetiology is likely to be polygenic and non-mendelian. Schizophrenia has a heritability of about 80%. The genetic risk in the general population and relatives of affected individuals is shown in Table 22.22. Brain scans and histology often show ventricular enlargement and disorganized cytoarchitecture in the hippocampus, supporting the neurodevelopmental theory of aetiology. Dopamine receptors are upregulated in the mesolimbic system, but the serotonin system may also be involved.

Table 22.22 Genetic risk for schizophrenia

General population	1%
Second-degree relative	2.5%
Parent	4%
Sibling	8%
Child of one affected parent	12%
Child of two affected parents	30–40%
Dizygotic twin	8–19%
Monozygotic twin	40–60%

Clinical features

The illness can begin at any age but is rare before puberty. The peak age of onset is in the early twenties. The symptoms that have been considered as diagnostic of the condition have been termed *first rank symptoms* and were described by the German psychiatrist Kurt Schneider. They consist of:

- auditory hallucinations in the third person, and/or voices commenting on their behaviour
- thought withdrawal, insertion and broadcast
- primary delusion (arising out of nothing)
- delusional perception
- somatic passivity and feelings – patients believe that thoughts, feelings or acts are controlled by others.

The more of these symptoms a patient has, the more likely the diagnosis is schizophrenia, but these symptoms may also occur in other psychoses. Other symptoms of *acute* schizophrenia include behavioural disturbances, other hallucinations, secondary (usually persecutory) delusions and blunting of mood. Schizophrenia is sometimes divided into 'positive' (type 1) and 'negative' (type 2) types:

Positive schizophrenia is characterized by acute onset, prominent delusions and hallucinations, normal brain structure, a biochemical disorder involving dopaminergic transmission, a good response to neuroleptics, and a better outcome.

Negative schizophrenia is characterized by a slow, insidious onset, a relative absence of acute symptoms, the presence of apathy, social withdrawal, lack of motivation, underlying brain structure abnormalities, and poor neuroleptic response.

Chronic schizophrenia

This is characterized by long duration and 'negative' symptoms of underactivity, lack of drive, social withdrawal and emotional emptiness.

Differential diagnosis

Schizophrenia should be distinguished from:

- organic mental disorders (e.g. partial complex epilepsy)
- mood (affective) disorders (e.g. mania)
- drug psychoses (e.g. amfetamine psychosis)
- personality disorders (schizotypal).

In older patients, any acute or chronic brain syndrome can present in a schizophrenia-like manner. A helpful diagnostic point is that altered consciousness and disturbances of memory do not occur in schizophrenia, and visual hallucinations are unusual.

A *schizoaffective psychosis* describes a clinical presentation in which clear-cut affective and schizophrenic symptoms coexist in the same episode.

Prognosis

The prognosis of schizophrenia is variable. A review of treatment studies suggests that 15–25% of people with schizophrenia recover completely, about 70% will have relapses and may develop mild to moderate negative symptoms, and about 10% will remain seriously disabled.

Treatment

The best results are obtained by combining drug and social treatments.

Antipsychotic (neuroleptic) drugs

These act by blocking the D_1 and D_2 groups of dopamine receptors. Such drugs are most effective against acute, positive symptoms and are least effective in the management of chronic, negative symptoms. Complete control of positive symptoms can take up to 3 months, and premature discontinuation of treatment can result in relapse.

As antipsychotic drugs block both D_1 and D_2 dopamine receptors, they usually produce extrapyramidal side-effects. This limits their use in maintenance therapy of many patients. They also block adrenergic and muscarinic receptors and thereby cause a number of unwanted effects (Table 22.23).

The neuroleptic malignant syndrome is an infrequent but potentially dangerous unwanted effect. This occurs in 0.2% of patients on *neuroleptic drugs*, particularly the potent dopaminergic antagonists, e.g. haloperidol. Symptoms occur a few days to a few weeks after starting therapy and consist of hyperthermia, muscle rigidity, autonomic instability (tachycardia, labile BP, pallor) and a fluctuating level of consciousness. Investigations show a raised creatine kinase, raised white cell count and abnormal liver biochemistry. Treatment consists of stopping the drug and general resuscitation, e.g. temperature reduction. Bromocriptine enhances dopaminergic activity and dantrolene will reduce muscle tone but no treatment has proven benefit.

Pregnancy. Data on the potential teratogenicity of antipsychotic (neuroleptic) medications are still limited. The disadvantages of not treating during pregnancy have to be balanced against possible developmental risks to the fetus. The butyrophenones (e.g. haloperidol) are probably safer than the phenothiazines. Subsequent management decisions on dosage will depend primarily on the ability to avoid side-effects, since the antiparkinsonian agents are still believed to be teratogenic and should be avoided.

Table 22.23 Unwanted effects of antipsychotic drugs

Common effects
Motor
Acute dystonia
Parkinsonism
Akathisia (motor restlessness)
Tardive dyskinesia
Autonomic
Hypotension
Failure of ejaculation
Antimuscarinic
Dry mouth
Urinary retention
Constipation
Blurred vision
Metabolic
Weight gain
Rare effects
Hypersensitivity
Cholestatic jaundice
Leucopenia
Skin reactions
Others
Precipitation of glaucoma
Galactorrhoea (due to hyperprolactinaemia)
Amenorrhoea
Cardiac arrhythmias
Seizures

Phenothiazines

Phenothiazines were the first group of neuroleptics used extensively. Chlorpromazine (100–1000 mg daily) is the drug of choice when a more sedating drug is required. Trifluoperazine is used when sedation is undesirable. Fluphenazine decanoate is used as a long-term prophylactic to prevent relapse, as a depot injection (25–100 mg i.m. every 1–4 weeks).

Butyrophenones

The butyrophenones (e.g. haloperidol 2–30 mg daily) are also powerful antipsychotics, used in the treatment of acute schizophrenia and mania. They are likely to cause a dystonia and/or extrapyramidal side-effects, but are much less sedating than the phenothiazines. A third of patients with acute schizophrenia will have a good response to haloperidol and a further third will make a partial response.

Atypical antipsychotics

These drugs are 'atypical' in that they block D_2 receptors less than D_1 and thus cause fewer extrapyramidal side-effects and less tardive dyskinesia. They are now recommended as first-line drugs for newly diagnosed schizophrenia and for patients taking typical antipsychotics but with significant adverse effects.

Clozapine. Clozapine is used in patients with intractable schizophrenia who have failed to respond to at least two conventional antipsychotic drugs or as first-line therapy. This drug is a dibenzodiazepine with a relative high affinity for D_1 rather than D_2 dopamine receptors, muscarinic and α-adrenergic receptors. It also blocks 5-HT_2 and 5-HT_1 receptors.

FURTHER READING

Mueser KT, McGurk SR. Schizophrenia. *Lancet* 2004; **363**: 2063–2072.

Functional brain scans have shown that clozapine selectively blocks limbic dopamine receptors more than striatal ones, which is probably why it causes considerably fewer extrapyramidal side-effects.

Clozapine has been shown to exercise a dramatic therapeutic effect on both intractable positive and negative symptoms. However, clozapine is expensive and produces severe agranulocytosis in 1–2% of patients. Therefore it can only be prescribed in the UK to registered patients by doctors and pharmacists registered with the Clozaril patient-monitoring service. The starting dose is 25 mg per day with a maintenance dose of 150–300 mg daily. White cell counts should be monitored weekly for 18 weeks and then 2-weekly for the length of treatment. In addition to its antipsychotic actions, clozapine may also help reduce aggressive and hostile behaviour and the risk of suicide. It can cause considerable weight gain and sialorrhoea. There is an increased risk of diabetes mellitus.

Risperidone is a benzisoxazole derivative with combined dopamine D_2 receptor and 5-HT_{2A}-receptor blocking properties. Dosage ranges from 6 to 10 mg per day. The drug is not markedly sedative and the overall incidence and severity of extrapyramidal side-effects is lower than with more conventional antipsychotics.

Olanzapine has affinity for 5-HT_2, D_1, D_2, D_4, and muscarinic receptor sites. Clinical studies indicate it to have a lower incidence of extrapyramidal side-effects. The apparent better compliance with the drug may be related to its lower side-effect profile and its once-daily dosage of 5–15 mg. Weight gain is a problem with long-term treatment and there is an increased risk of diabetes mellitus.

Neither risperidone nor olanzapine seems as specific a treatment for intractable chronic schizophrenia as clozapine.

Other atypical antipsychotics include amisulpride, sulpiride, zotepine, ziprasidone and quetiapine, the latter causing less hyperprolactinaemia. Those taking atypical antipsychotics should be regularly monitored for the development of diabetes mellitus. No more than one antipsychotic should be prescribed routinely.

Psychological treatment

This consists of reassurance, support and a good doctor–patient relationship. Psychotherapy of an intensive or exploratory kind is contraindicated. In contrast, individual cognitive behaviour therapy can help reduce the intensity of delusions.

Social treatment

Social treatment involves attention being paid to the patient's environment and social functioning. Family education can help relatives and partners to provide the optimal amount of emotional and social stimulation, so that not too much emotion is expressed (a risk for relapse). Sheltered employment is usually necessary for the majority of sufferers if they are to work. Assertive outreach mental health teams will follow up those poorly adhering to treatment.

Medical presentations related to treatment

The motor side-effects of neuroleptics are the commonest reason for a patient with schizophrenia to present to a physician, followed by deliberate self-harm. Acute dystonia normally arises in patients newly started on neuroleptics, causing a torticollis. Extrapyramidal side-effects are common and present in the same way as Parkinson's disease. They remit on stopping the drug and with antimuscarinic drugs, e.g. procyclidine. Akathisia is a motor restlessness, most commonly affecting the legs. It is similar to the restless legs syndrome, but apparent during the day.

Amenorrhoea and galactorrhoea can be caused by dopamine antagonists. Postural hypotension can affect the elderly, and neuroleptics can be the cause of delirium in the elderly, if their antimuscarinic effects are prominent.

ORGANIC MENTAL DISORDERS

Organic brain disorders result from structural pathology, as in dementia (see p. 1167), or from disturbed central nervous system (CNS) function, as in fever-induced delirium. They do not include mental and behavioural disorders due to alcohol and misuse of drugs, which are classified separately.

Delirium

Delirium, also termed *toxic confusional state* and *acute organic reaction*, is an acute or subacute brain failure in which impairment of attention is accompanied by abnormalities of perception and mood. It is the most common psychosis seen in the general hospital. Ten to 20 per cent of older surgical and medical inpatients have delirium during their admission. Confusion is usually worse at night, with consequent sleep reversal, so that the patient is asleep in the day and awake all night. During the acute phase, thought and speech are incoherent, memory is impaired and misperceptions occur. Episodic visual hallucinations (or illusions) and persecutory delusions occur. As a consequence, the patient may be frightened, suspicious, restless and uncooperative.

A developing, deteriorating or damaged brain predisposes a patient to develop delirium (Table 22.24).

A large number of diseases may cause delirium, particularly in elderly patients. Some causes of delirium are listed in Table 22.25. Delirium tremens should be considered in the differential diagnosis (p. 1167) as well as Lewy body dementia (p. 1167). Diagnostic criteria are given in Box 22.15.

Management

History should be taken from a witness. Examination may reveal the cause. Investigation and treatment of the underlying physical disease should be undertaken (Table 22.25). The

Table 22.24 Predisposing factors in delirium

Extremes of age (developing or deteriorating brain)
Damaged brain:
Any dementia (most common predisposition)
Previous head injury
Alcoholic brain damage
Previous stroke
Dislocation to an unfamiliar environment (e.g. hospital admission)
Sleep deprivation
Sensory extremes (overload or deprivation)
Immobilization
Visual or hearing impairment

Table 22.25 Some common causes of delirium

Systemic infection
Any infection, particularly with high fever (e.g. malaria, septicaemia)
Metabolic disturbance
Hepatic failure
Renal failure
Disorders of electrolyte balance, dehydration
Hypoxia
Vitamin deficiency
Thiamin (Wernicke–Korsakoff syndrome, beriberi)
Nicotinic acid (pellagra)
Vitamin B_{12}
Endocrine disease
Hypothyroidism
Cushing's syndrome
Intracranial causes
Trauma
Tumour
Abscess
Subarachnoid haemorrhage
Epilepsy
Drug intoxication
Anticonvulsants
Antimuscarinics
Anxiolytic/hypnotics
Tricyclic antidepressants
Dopamine agonists
Digoxin
Drug/alcohol withdrawal
Postoperative states
Terminal illness

Box 22.15 Delirium – diagnostic criteria (derived from DSM-IV-TR)

- Disturbance of consciousness:
  ↓ clarity of awareness of environment
  ↓ ability to focus, sustain or shift attention
- Change in cognition:
  memory deficit, disorientation, language disturbance, perceptual disturbance
- Disturbance develops over a short period (hours or days)
- Fluctuation over course of day

patient should be carefully nursed and rehydrated in a quiet single room with a window that does not allow exits. If a high fever is present, the patient's temperature should be reduced. All current drug therapy should be reviewed and, where possible, stopped. Psychoactive drugs should be avoided if possible (because of their own risk of exacerbating delirium). In severe delirium, haloperidol is an effective choice, the daily dose usually ranging between 1.5 mg (in the elderly) and 30 mg per day. If necessary, the first dose can be administered intramuscularly. Olanzapine is an effective alternative, especially if given at night for insomnia.

Prognosis of delirium

Delirium usually clears within a week or two, but brain recovery usually lags behind the recovery of the causative physical illness. The prognosis depends not only on the successful treatment of the causative disease, but also on the underlying state of the brain. Twenty-five per cent of the elderly with delirium will have an underlying dementia; 15% of patients do not survive their underlying illness; 40% are in institutional care at 6 months.

EATING DISORDERS

Obesity

This is the commonest eating disorder (see p. 228) and has become epidemic in some developed countries. It is usually caused by a combination of constitutional and social factors, but a binge eating disorder and psychological determinants of 'comfort eating' should be excluded.

Anorexia nervosa

The main clinical criteria for diagnosis are:

- a bodyweight more than 15% below the standard weight, or a body mass index (BMI) below 17.5 (ICD-10)
- weight loss is self-induced by avoidance of fattening foods, vomiting, purging, exercise, or appetite suppressants
- a distortion of body image so that the patient regards herself as fat when she is thin
- a morbid fear of fatness
- amenorrhoea in women.

Clinical features include:

- onset usually in adolescence
- a previous history of faddish eating
- the patient generally eats little, yet is obsessed by food
- excessive exercising.

The physical consequences of anorexia include sensitivity to cold, constipation, hypotension and bradycardia. In most cases, amenorrhoea is secondary to the weight loss. In those who also binge and vomit or abuse purgatives there may be hypokalaemia and alkalosis.

Prevalence

Case register data suggest an incidence rate of 19 per 100 000 females aged between 15 and 34 years. Surveys have suggested a prevalence rate of approximately 1% among schoolgirls and university students. However, many more young women have amenorrhoea accompanied by less weight loss than the 15% required for the diagnosis. The condition is much less common among men (ratio of 1:10).

Aetiology

Biological factors

Genetic. Six to 10 per cent of siblings of affected women suffer from anorexia nervosa. There is an increased concordance amongst monozygotic twins, suggesting a genetic predisposition.

Hormonal. The reductions in sex hormones and downregulation of the hypothalamic–pituitary–adrenal axis are secondary to malnutrition and usually reversed by refeeding.

FURTHER READING

Inouye SK. Delirium in older persons. *New England Journal of Medicine* 2006; **354**: 1157–1165.

Psychological factors

Individual. Anorexia nervosa has often been seen as an escape from the emotional problems of adolescence and a regression into childhood. Patients will often have had dietary problems in early life. Perfectionism and low self-esteem are common antecedents. Studies suggest that survivors of childhood sexual or other abuse are at greater risk of developing an eating disorder, usually anorexia nervosa, in adolescence.

Family. Families of such patients are allegedly characterized by overprotection, inflexibility and lack of conflict resolution. Anorexia is alleged to prevent dissension in families. However, a case control study suggested that there is no more evidence of these factors in families of patients with anorexia nervosa than in control families with a child with an established physical disease, suggesting that, if present, these features are secondary.

Social and cultural factors

There is a higher prevalence in higher social classes, westernized families and in certain occupational groups (e.g. ballet dancers and nurses) and in societies where cultural value is placed on thinness.

Prognosis

The condition runs a fluctuating course, with exacerbations and partial remissions. Long-term follow-up suggests that about two-thirds of patients maintain normal weight and that the remaining one-third are split between those who are moderately underweight and those who are seriously underweight. Indicators of a poor outcome include:

- a long initial illness
- severe weight loss
- older age at onset
- binging and purging
- personality difficulties
- difficulties in relationships.

Suicide has been reported in 2–5% of patients with chronic anorexia nervosa. The mortality rate per year is 0.5% from all causes. The illness can quickly cause osteopenia, and later osteoporosis. More than one-third have recurrent affective illness, and various family, genetic and endocrine studies have found associations between eating disorders and depression.

Treatment

Treatment can be conducted on an outpatient basis unless the weight loss is severe and accompanied by marked cardiovascular signs and/or electrolyte and vitamin disturbances. Hospital admission may then be unavoidable and may need to be on a medical ward initially. Uncommonly, the patient's weight loss may be so severe as to be life-threatening. If the patient cannot be persuaded to enter hospital, compulsory admission may have to be used. Inpatient treatment goals include:

- establishing a therapeutic relationship with both the patient and her family
- restoring the weight to a level between the ideal bodyweight and the patient's ideal weight
- the provision of a balanced diet, aimed at gaining 0.5–1 kg weight per week
- the elimination of purgative and/or laxative use and vomiting

Outpatient treatment should include cognitive behavioural or interpersonal psychotherapies. Setting up a therapeutic alliance is vital. Family therapy is more effective than individual psychotherapy in adolescents and those still at home, and less effective in those who have left home. Motivational enhancement techniques are being used with some success.

Drug treatment has met with limited success, and those drugs that increase the QTc interval should be avoided. Vitamins and minerals may need replacement.

Bulimia nervosa

This refers to episodes of uncontrolled excessive eating, which are also termed 'binges', accompanied by means to lose weight. There is a preoccupation with food and habitual behaviours to avoid the fattening effects of periodic binges. These behaviours include:

- self-induced vomiting
- misuse of drugs – laxatives, diuretics, thyroid extract or anorectics.

Additional clinical features include:

- physical complications of vomiting:
  (a) cardiac arrhythmias
  (b) renal impairment
  (c) muscular paralysis
  (d) tetany – from hypokalaemic alkalosis
  (e) swollen salivary glands – from vomiting
  (f) eroded dental enamel
- associated psychiatric disorders:
  (a) depressive illness
  (b) alcohol misuse
- fluctuations in bodyweight within normal limits
- menstrual function – periods irregular but amenorrhoea rare
- personality – perfectionism and/or low self-esteem present premorbidly.

The prevalence of bulimia in community studies is high; it affects between 5% and 30% of girls attending high schools, colleges or universities in the USA. Bulimia is sometimes associated with anorexia nervosa. A premorbid history of dieting is common. The prognosis for bulimia nervosa is better than for anorexia nervosa.

Treatment

Individual cognitive behaviour therapy is more effective than both interpersonal psychotherapy and drug treatments. SSRIs (e.g. fluoxetine) are also an effective treatment, even in the absence of a depressive illness, but may require high doses.

Atypical eating disorders

These include eating disorders that do not conform clinically to the diagnostic criteria for anorexia nervosa or bulimia nervosa. Binge eating disorders consist of bulimia without the vomiting and other weight-reducing strategies.

FURTHER READING

Mehler PS. Bulimia nervosa. *New England Journal of Medicine* 2003; **349**: 875–881.

Wilson GT, Shafran. Eating disorders guidelines from NICE. *Lancet* 2005; **365**: 79–81.

Yager J, Anderson AE. Anorexia nervosa. *Lancet* 2005; **353**: 1481–1488.

SEXUAL DISORDERS

Sexual disorders can be divided into sexual dysfunctions and deviations and gender role disorders (Table 22.26).

Sexual dysfunction

Sexual dysfunction in men refers to repeated inability to achieve normal sexual intercourse, whereas in women it refers to a repeatedly unsatisfactory quality of sexual satisfaction. Problems of sexual dysfunction can usefully be classified into those affecting sexual desire, arousal and orgasm. Among men presenting for treatment of sexual dysfunction, erectile dysfunction is the most frequent complaint. The prevalence of premature ejaculation is low, except in young men, while ejaculatory failure is rare.

Sexual drive is affected by constitutional factors, ignorance of sexual technique, anxiety about sexual performance, medical and psychiatric conditions and certain drugs (Tables 22.27 and 22.28).

The treatment of sexual dysfunction involves careful assessment of both physical and psychosocial factors, the participation (where appropriate) of the patient's partner, sex education, and specific therapeutic techniques, including relaxation, behavioural therapy and couple therapy. The introduction of phosphodiesterase type 5 inhibitors (e.g. sildenafil) has provided an effective therapy for the treatment of erectile dysfunction (see p. 1000).

Sexual deviation

Sexual deviations are regarded as unusual forms of behaviour rather than as an illness. Doctors are only likely to be involved when the behaviour involves breaking the law (e.g. paedophilia) and when there is a question of an associated mental or physical disorder. Men are more likely than women to have sexual deviations.

Table 22.26 Classification of sexual disorders

Sexual dysfunction	Sexual deviations	Disorders of the gender role
Affecting sexual desire Low libido	**Variations of the sexual 'object'** Fetishism Transvestism Paedophilia Bestiality Necrophilia	Transsexualism
Impaired sexual arousal Erectile dysfunction Failure of arousal in women		
Affecting orgasm Premature ejaculation Retarded ejaculation Orgasmic dysfunction in women	**Variations of the sexual act** Exhibitionism Voyeurism Sadism Masochism Frotteurism	

Table 22.27 Medical conditions affecting sexual performance

Endocrine
Diabetes mellitus
Hyperthyroidism
Hypothyroidism
Cardiovascular
Angina pectoris
Previous myocardial infarction
Disorders of peripheral circulation
Hepatic
Cirrhosis, particularly alcohol-related
Renal
Chronic kidney disease
Neurological
Neuropathy
Spinal cord lesions
Respiratory
Asthma
COPD
Psychiatric
Depressive illness
Substance misuse

Table 22.28 Drugs affecting sexual arousal

Male arousal	Female arousal
Alcohol	Alcohol
Benzodiazepines	CNS depressants
Neuroleptics	Antidepressants (SSRIs)
Cimetidine	Oral combined contraceptives
Opiate analgesics	Methyldopa
Methyldopa	Clonidine
Clonidine	
Spironolactone	
Antihistamines	
Metoclopramide	
Diuretics	
Beta-blockers	
Cannabis	

Alcohol increases the desire but diminishes the performance

Gender role disorders

Transsexualism involves a disturbance in gender identity in which the patient is convinced that their body is the wrong gender. A person's gender identity refers to the individual's sense of masculinity or femininity as distinct from sex. It is thought to arise from a biological component (prenatal endocrine influences), psychological imprinting and social conditioning. Disturbances in these three areas have variously been blamed for the cause of transsexualism.

For males, treatment includes oestrogen administration and, if surgery is to be recommended, a period of living as a woman as a trial beforehand. In the case of female transsexuals, treatment involves surgery and the use of methyltestosterone.

PERSONALITY DISORDERS

These disorders comprise enduring patterns of behaviour which manifest themselves as inflexible responses to a broad range of personal and social situations. Personality disorders are developmental conditions that appear in childhood or adolescence and continue into adult life. Their prevalence is about 15% in the population. They are not secondary to another psychiatric disorder or brain disease, although they may precede or coexist with other disorders. In contrast, personality change is acquired, usually in adult life, following severe or prolonged stress, extreme environmental deprivation, serious psychiatric disorder or brain injury or disease.

Personality disorders are usually subdivided according to clusters of traits that correspond to the most frequent or obvious behavioural manifestations, but many will show the characteristics of more than one category. The main categories of personality disorder are described below.

Borderline (emotionally unstable). Such people act impulsively and develop intense, but short-lived, emotional attachments to others. They describe chronic internal emptiness with frequent self-harm, self-abuse (eating disorders, substance misuse) and they may develop transient psychotic features of uncertain significance. There is often a strong family history of mood disorders.

Paranoid. Characterized by extreme sensitivity, suspiciousness, litigiousness, a tendency to excessive self-importance, and a preoccupation with unsubstantiated conspiracy theories.

Schizoid. Characterized by emotional coldness and detachment, a limited capacity to express emotions, indifference to praise or criticism, a preference for solitary activities, lack of close friendships, and a marked insensitivity to prevailing social norms and conventions.

Antisocial. An antisocial personality has a callous unconcern for the feelings of others, an incapacity to maintain enduring relationships, a very low tolerance of frustration, an incapacity to experience guilt and to profit from experience, and a tendency to blame others.

Histrionic. Characterized by self-dramatization, theatricality, suggestibility, shallow and labile emotions, a continual seeking for excitement and appreciation by others, and an inappropriate seductiveness in appearance or behaviour.

Narcissistic. Such individuals have an inflated sense of self-importance, a need to be admired, and are self referenced and arrogant.

Anankastic (obsessive-compulsive). Such a personality is characterized by feelings of excessive doubt and caution, preoccupation with details, rules and order, perfectionism, excessive conscientiousness, scrupulousness, pedantry and rigidity.

Dependent. People with such a personality encourage others to make their personal decisions, subordinate their needs to others on whom they are dependent, are unwilling to make demands on others, feel unable to care for themselves, and are preoccupied with fears of being abandoned.

MENTAL HEALTH ACT

The law in most developed countries provides for the compulsory admission and treatment of mentally disordered persons for their own protection and/or the protection of others and for mitigation in the case of mentally disordered individuals who commit a criminal offence (see also p. 5). In England and Wales a new Mental Health Act was implemented in 2008. It amends the previous 1983 Act. Changes include a widening of the definition of mental disorder to include 'any disorder or disability of the mind. It abolishes the treatability provision to simply appropriate medical treatment (including care) being available 'to alleviate, or prevent a worsening of, the disorder or one or more of its symptoms or manifestations', and a supervised community treatment order. The Responsible Medical Officer is now 'the responsible clinician', and an 'approved clinician' rather than just a doctor can also sign a medical recommendation. The Mental Health (Scotland) Act 2003 and the Mental Health (Northern Ireland) Order 2004 contain clauses broadly similar to those in England and Wales.

There are three conditions that need to be met before an appropriate compulsory section form is signed. The patient must be:

- suffering from a mental disorder
- at risk to his/her health or safety and/or other people's safety
- unwilling to accept hospitalization voluntarily.

Learning disability by itself and alcohol/drug dependence are specific exclusions. Any registered medical practitioner may sign a medical recommendation under the Act, but the added signature of an approved mental health professional is needed for compulsory orders lasting for more than 72 hours. Unless the patient is already in hospital, the nearest relative or an approved mental health professional is also required to sign the application form. Relevant sections of the Act are detailed in Table 22.29.

Physicians are likely to be involved in sections 5(2) and 2. It should be remembered that a section does not give a doctor the right to treat a physical disease, although it could be argued that a section would apply if the physical disease was causing the mental disorder (e.g. delirium). This has never been legally tested.

MENTAL CAPACITY ACT

Mental capacity is the ability to make decisions about all aspects of one's life, including but not exclusively healthcare. All doctors need to be able to assess capacity. In England and Wales a new act was passed in 2005, which, for the first time, formally protects patients who lack capacity. Some 3% of people in the UK are thought to lack capacity due to conditions affecting brain function, such as dementia. However, the assessment of capacity is specific to an individual decision, so it is possible to have capacity to make one decision but not another. The assessment of capacity is outlined in Box 22.16. Capacity is assumed unless there is evidence to the contrary. In the absence of capacity, the doctor should act in the best interests of the patient and provide the least restrictive management, after consulting an independent advocate.

Table 22.29 Commonly used sections of the Mental Health Act 2008

Section	Duration	Signatures required	Purpose
2	28 days	Two doctors or approved clinicians *plus* nearest relative or approved mental health worker	Assessment and treatment
3	6 months	Two doctors or approved clinicians *plus* nearest relative or approved mental health worker	Treatment
4	72 hours	One doctor or approved clinician *plus* relative or approved mental health worker	Emergency admission
5(2)	72 hours	Doctor in charge of patient's care or nominee, or approved clinician	Emergency detention of a patient already in hospital
5(4)	6 hours	Nurse (RMN)	Emergency detention of a patient already in hospital
136	72 hours	Police officer	Psychiatric assessment of those in public places thought by police to be mentally ill and in need of a place of safety

Box 22.16 Assessment of mental capacity to make a decision

The patient has a demonstrated impairment or disturbance of their mind/brain and a demonstrated inability to do any of the following:

- understand relevant information
- retain that information for sufficient time to make the decision
- use or weigh that information
- communicate their decision.

FURTHER READING

Nicholson TRJ, Cutter W, Hotopf M. Assessing mental capacity: the Mental Capacity Act. *British Medical Journal* 2008; **336:** 322–325.

SIGNIFICANT WEBSITES

http://www.rcpsych.ac.uk
UK Royal College of Psychiatrists

http://www.connects.org.uk
Website for mental health in general

http://www.cebmh.com
Centre for Evidence-Based Mental Health

http://www.mentalhealth.org.uk
Mental Health Foundation – charity

http://www.sleepfoundation.org
National Sleep Foundation

http://www.edauk.com
Eating Disorders Association

http://psych.org
American Psychiatric Association

The late Professor Anthony Clare was a co-author of the Sixth Edition of this chapter.

This page intentionally left blank

23 Skin disease

Structure and function of the skin 1225
Approach to the patient 1227
Infections 1228
Bacterial infections 1228
Mycobacterial infections 1230
Viral infections 1230
Fungal infections 1232
Infestations 1234
Papulo-squamous/inflammatory rashes 1236
Eczema 1236
Psoriasis 1240
Urticaria/angio-oedema 1243
Pityriasis rosea 1244
Lichen planus 1244
Granuloma annulare 1245
Lichen sclerosus 1245
Facial rashes 1245
Photodermatology 1248
Idiopathic photodermatoses 1248
Erythroderma 1249
Cutaneous signs of systemic disease 1250
Bullous disease 1255
Immunobullous disease 1255
Mechanobullous disease 1257
Skin tumours 1257
Benign cutaneous tumours 1257
Potentially pre-malignant cutaneous tumours 1258
Malignant cutaneous tumours 1259
Disorders of blood vessels/lymphatics 1261
Leg ulcers 1261
Pressure sores 1262
Vasculitis 1263
Lymphatics 1263
Disorders of collagen and elastic tissue 1263
Disorders of pigmentation 1264
Hypopigmentation 1264
Hyperpigmentation 1264
Drug-induced rashes 1265
Disorders of nails 1266
Disorders of hair 1266
Hair loss 1266
Increased hair growth 1267
Birth marks/neonatal rashes 1267
Human immunodeficiency virus and the skin 1268
Dermatoses of pregnancy 1269
Principles of topical therapy 1270

Introduction

Skin diseases have a high prevalence throughout the world. In developing countries infectious diseases such as tuberculosis, leprosy and onchocerciasis are common, whereas in developed countries inflammatory disorders such as eczema and acne are common. Skin disorders can be inherited, e.g. Ehlers–Danlos syndrome, a part of normal development, e.g. acne vulgaris, or may present as part of a systemic disorder, e.g. systemic lupus erythematosus (SLE). Approximately 25% of the UK population will develop a skin problem and although self-medication is common, skin disease still accounts for 10% of the workload of family doctors. The common reasons for this are itching or pain, which can interfere with people's ability to function normally or to sleep; rashes which cause anxiety, depression and lack of self-confidence and can lead to social isolation if obviously visible; and an inability to work, because certain dermatoses (such as allergic hand eczema in a builder or hairdresser) can interfere with or even prevent working.

Rarely skin disease can be fatal. Examples are malignant melanoma, toxic epidermal necrolysis and pemphigus.

STRUCTURE AND FUNCTION OF THE SKIN

The skin consists of four distinct layers: the epidermis, the basement membrane zone, the dermis and the subcutaneous layer (Fig. 23.1). The functions are summarized in Box 23.1.

The epidermis

The epidermis is a stratified epithelium of ectodermal origin that arises from dividing basal keratinocytes. The downward projections of the epidermis into the dermis are called the 'rete ridges'. The lower epidermal cells (basal layer) produce a variety of keratin filaments and desmosomal proteins (e.g. desmoglein and desmoplakin), which make up the 'cytoskeleton'. This confers strength to the epidermis and prevents it shedding off. Higher up in the granular layer, complex lipids are secreted by the keratinocytes and these form into intercellular lipid bilayers, which act as a semipermeable skin barrier. The upper cells (stratum corneum) lose their nuclei and become surrounded by a tough impermeable 'envelope'

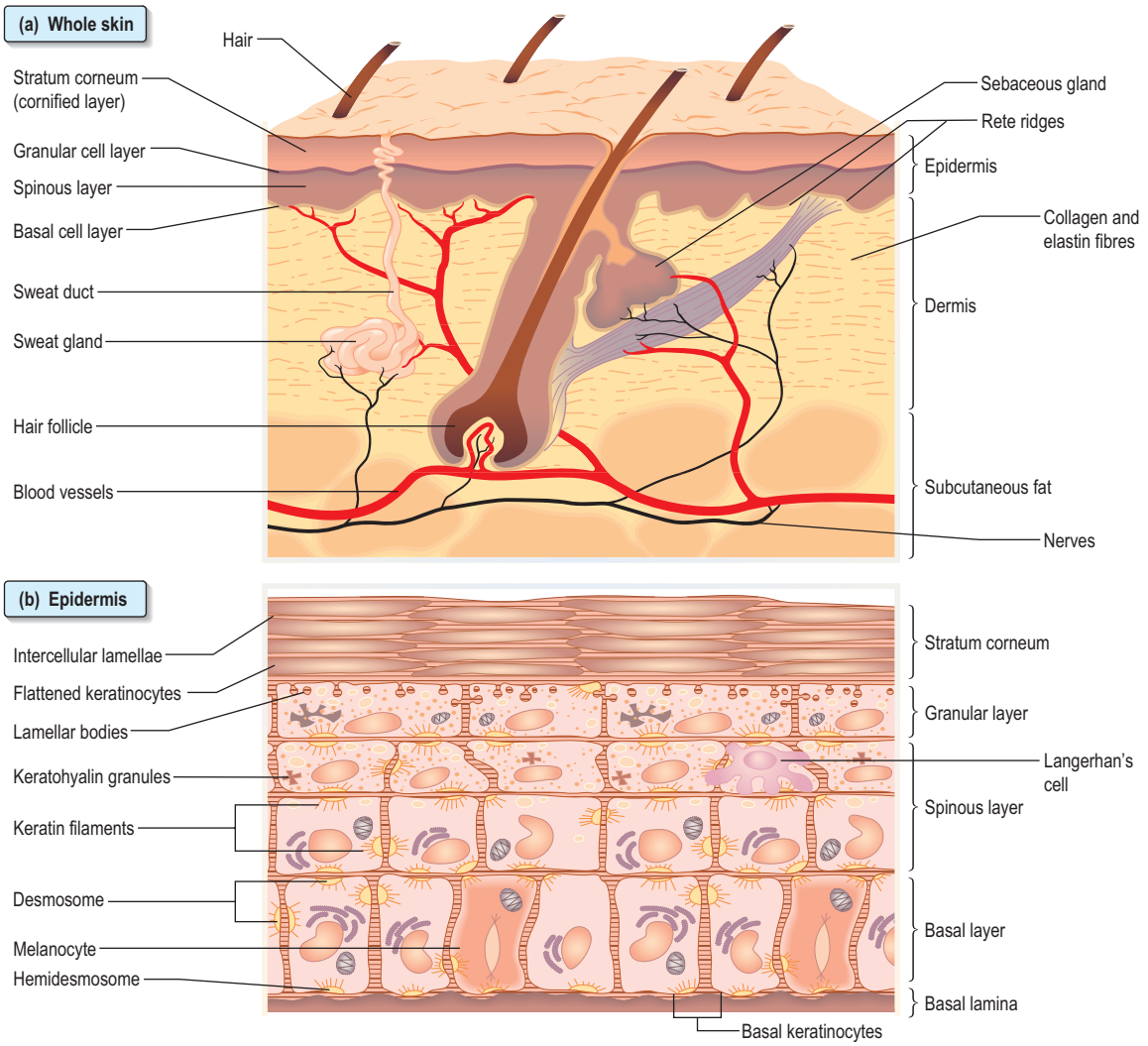


Fig. 23.1 **The structure of the skin** (see also Fig. 23.27).

Box 23.1 Functions of the skin

- Physical barrier against friction and shearing forces
- Protection against infection (immune and innate), chemicals, ultraviolet irradiation
- Prevention of excessive water loss or absorption
- Ultraviolet-induced synthesis of vitamin D
- Temperature regulation
- Sensation (pain, touch and temperature)
- Antigen presentation/immunological reactions/wound healing

of various proteins (loricrin, involucrin, filaggrin and keratin). Changes in lipid metabolism and protein expression in the outer layers allow normal shedding of keratinocytes.

Keratinocytes can secrete a variety of cytokines (e.g. interleukins, gamma-interferon, tumour necrosis factor alpha) in response to tissue injury or in certain skin diseases. These play a role in specific immune function, cutaneous inflammation and tissue repair. There is a further layer of protection against microbial invasion called the innate immune system of the skin. This comprises neutrophils and macrophages as well as keratinocyte-produced antimicrobial peptides (called β-defensins and cathelicidins). Expression of these peptides is both constitutive and induced by skin inflammation and they are active against bacterial, viral and fungal pathogens. There is evidence to suggest a deficiency of these peptides may account for the susceptibility of patients with atopic eczema to skin infection.

Other cells in the epidermis

Melanocytes are found in the basal layer and secrete the pigment melanin. These protect against UV irradiation. Racial differences are due to variation in melanin production, not melanocyte numbers.

Merkel cells are also found in the basal layer and originate either from neural crest or epidermal keratinocytes. They are numerous on finger tips and in the oral cavity and play a role in sensation.

Langerhans' cells are dendritic cells found in the suprabasal layer. They derive from the bone marrow and as they express the cytokine CCR6, they are guided to normal skin, which contains a CCR6 agonist called macrophage inflammatory protein 3α. Langerhans' cells endocytose extracellular antigens in the skin and then migrate to local lymph nodes for T-cell presentation and thus act as antigen-presenting cells.

Basement membrane zone (see Fig. 23.27)

The basement membrane zone is a complex proteinaceous structure consisting of type IV, VII and XVII (previously called BP180) collagen, hemidesmosomal proteins, integrins and laminin. Collectively they hold the skin together, keeping the epidermis firmly attached to the dermis. Inherited or autoimmune-induced deficiencies of these proteins can cause skin fragility and a variety of blistering diseases (see p. 1255).

The dermis

The dermis is of mesodermal origin and contains blood and lymphatic vessels, nerves, muscle, appendages (e.g. sweat glands, sebaceous glands and hair follicles) and a variety of immune cells such as mast cells and lymphocytes. It is a matrix of collagen and elastin in a ground substance.

The sweat glands

Eccrine sweat glands are found throughout the skin except the mucosal surfaces.

Apocrine sweat glands are found in the axillae, anogenital area and scalp and do not function until puberty.

The sweat glands and vasculature are involved in temperature control.

The sebaceous glands

These are inactive until puberty. They are responsible for secreting sebum or grease onto the skin surface (via the hair follicle) and are found in high numbers on the face and scalp.

Nerves

The skin is richly innervated. Nerve fibres enable sensation of touch, pain, itch, vibration and change in temperature.

Hair

Hairs arise from a downgrowth of epidermal keratinocytes into the dermis. The hair shaft has an inner and outer root sheath, a cortex and sometimes a medulla. The lower portion of the hair follicle consists of an expanded bulb (which also contains melanocytes) surrounding a richly innervated and vascularized dermal papilla. The hair regrows from the bulb after shedding. There are three types of hair:

- *terminal* – medullated coarse hair, e.g. scalp, beard, pubic
- *vellus* – non-medullated fine downy hairs seen on the face of women and in prepubertal children
- *lanugo* – non-medullated soft hair on newborns (most marked in premature babies) and occasionally in people with anorexia nervosa.

All hair follicles follow a growth cycle: anagen (growth phase), catagen (involution phase), telogen (shedding phase). At any one time most hairs (>90%) will be in the anagen phase which is typically 3–5 years for scalp hair.

Grey hair is due to decreased tyrosinase activity in the hair bulb melanocytes. White hair is due to total loss of these melanocytes.

Nails

Nails are tough plates of hardened keratin which arise from the nail matrix (just visible as the moon-shaped lunula) under the nail fold. It takes 6 months for a finger-nail to grow out fully and 1 year for a toe-nail.

The subcutaneous layer

The subcutaneous layer consists predominantly of adipose tissue as well as blood vessels and nerves. This layer provides insulation and acts as a lipid store.

APPROACH TO THE PATIENT

The *history* should aim to elicit the following points:

- time course of rash
- distribution of lesions
- symptoms (e.g. itch or pain)
- family history (especially of atopy and psoriasis)
- drug/allergy history
- past medical history
- provoking factors (e.g. sunlight or diet)
- previous skin treatments.

Examination entails looking *and* feeling a rash (for terminology, see Table 23.1). It should include an assessment of nails, hair and mucosal surfaces even if these are recorded as unaffected. The following terms are used to describe distribution: flexural, extensor, acral (hands and feet), symmetrical, localized, widespread, facial, unilateral, linear, centripetal

Table 23.1 Morphological description of skin lesions

Atrophy	Thinning of the skin
Bulla	A large fluid-filled blister
Crusted	Dried serum or exudate on the skin
Ecchymosis	Large confluent area of purpura ('bruise')
Erosion	Denuded area of skin (partial epidermal loss)
Excoriation	Scratch mark
Fissure	Deep linear crack or crevice (often in thickened skin)
Lichenified	Thickened epidermis with prominent normal skin markings
Macule	Flat, circumscribed non-palpable lesion
Nodule	Large papule (>0.5 cm)
Papule	Small palpable, circumscribed lesion (<0.5 cm)
Petechia	Pinpoint-sized macule of blood in the skin
Plaque	Large flat-topped, elevated, palpable lesion
Purpura	Larger macule or papule of blood in the skin which does not blanch on pressure
Pustule	Yellowish white pus-filled lesion
Scaly	Visible flaking and shedding of surface skin
Telangiectasia	Abnormal visible dilatation of blood vessels
Ulcer	Deeper denuded area of skin (full epidermal and dermal loss)
Vesicle	A small fluid-filled blister
Weal	Itchy raised 'nettle rash'-like swelling due to dermal oedema

FURTHER READING

Braff MH, Gallo RL. Antimicrobial peptides: an essential component of the skin defensive barrier. *Current Topics in Microbiology and Immunology* 2006; **306**: 91–110.

Krause K, Foitzik K. Biology of the hair follicle: the basics. *Seminars in Cutaneous Medicine and Surgery* 2006; **25**: 2–10.

Lin JY, Fisher DE. Melanocyte biology and skin pigmentation. *Nature* 2007; **445**: 843–850.

Table 23.2 Investigations used in skin disorders

Test	Use	Clinical example
Skin swabs	Bacterial culture	Impetigo
Blister fluid	Electron microscopy and viral culture	Herpes simplex
Skin scrapes	Fungal culture Microscopy	Tinea pedis Scabies
Nail sampling	Fungal culture	Onychomycosis
Wood's light	Fungal fluorescence	Scalp ringworm Erythrasma
Blood tests	Serology	Streptococcal cellulitis
	Autoantibodies	Systemic lupus erythematosus
	HLA typing	Dermatitis herpetiformis
	DNA analysis	Epidermolysis bullosa
Skin biopsy	Histology	General diagnosis
	Immunohistochemistry	Cutaneous lymphoma
	Immunofluorescence	Immunobullous disease
	Culture	Mycobacteria/fungi
Patch tests	Allergic contact eczema	Hand eczema
Urine	Dipstick (glucose) Cytology (red cells)	Diabetes mellitus Vasculitis
Dermatoscopy (direct microscopy of skin)	Assessment of pigmented lesions	Malignant melanoma

(trunk more than limbs), annular and reticulate (lacey network or mesh like).

Investigation. With regard to investigation, clinical acumen remains the most useful tool in dermatology but a variety of tests are useful in confirming a diagnosis (Table 23.2).

INFECTIONS

BACTERIAL INFECTIONS (see also p. 123)

The skin's normal bacterial flora prevents colonization by pathogenic organisms. A break in epidermal integrity by trauma, leg ulcers, fungal infections (e.g. athlete's foot) or abnormal scaling of the skin (e.g. in eczema) can allow infection. If reinfection occurs this may be due to asymptomatic nasal carriage of bacteria or the presence of other infected close contacts.

Impetigo

Impetigo is a highly infectious skin disease most common in children (Fig. 23.2). It presents as weeping, exudative areas with a typical honey-coloured crust on the surface. It is spread by direct contact. The term 'scrum pox' refers to impetigo spread between rugby players. Staphylococci or group A β-haemolytic streptococci are the causative agents: skin swabs should be taken.

Prevention

Prevention is by good personal hygiene, particularly hand washing with soap.

Treatment

Localized disease is treated with topical fusidic acid, and mupirocin is used for meticillin-resistant staphylococci (MRSA) three times daily. Extensive disease is treated with oral antibiotics for 7–10 days (flucloxacillin 500 mg four times daily for *Staphylococcus*; phenoxymethylpenicillin 500 mg four times daily for *Streptococcus*). Other close contacts should be examined and children should avoid school for 1 week after starting therapy. If impetigo appears resistant to treatment or recurrent, take nasal swabs and check other family members. Nasal mupirocin (three times daily for 1 week) is useful to eradicate nasal carriage, especially in hospital staff. Community acquired MRSA (p. 124) is increasingly recognized as a cause of superficial skin infections.

Fig. 23.2 **Impetigo** – crusted blistering lesions on the chin.

Bullous impetigo/staphylococcal scalded skin syndrome

Rarely *Staphylococcus* releases an exfoliating toxin which acts high up in the epidermis: toxin A causes blistering at the site of infection (bullous impetigo), toxin B spreads through the body causing more widespread blistering (staphylococcal scalded skin syndrome, SSSS). The latter is more common in childhood with very low mortality rates. In adults it is often associated with renal disease or immunosuppression, and mortality rates increase to 50%. Both these toxins cleave desmoglein 1 (a desmosomal protein) so mucosal involvement does not occur (this is analogous to pemphigus foliaceous which has the same target antigen, p. 1255).

SSSS can mimic toxic epidermal necrolysis (TEN, see p. 1266) on clinical grounds but can be differentiated in two ways: mucosal involvement does not occur in SSSS, and on skin biopsy a frozen section shows a superficial intraepidermal split in SSSS which is deeper and subepidermal in TEN.

Both bullous impetigo and SSSS are treated with anti-staphylococcal antibiotics (e.g. flucloxacillin) and supportive care.

Cellulitis

Cellulitis presents as a hot, sometimes tender area of confluent erythema of the skin due to infection of the deep subcutaneous layer. It often affects the lower leg, causing an upward-spreading, hot erythema, and occasionally will blister, especially if oedema is prominent. It may also be seen affecting one side of the face. Patients are often unwell with a high temperature. It is usually caused by a streptococcus, rarely a staphylococcus, and sometimes community acquired MRSA (p. 124). In the immunosuppressed or diabetic patient Gram-negative organisms or anaerobes should be suspected.

There may be an obvious portal of entry for infection such as a recent abrasion or a venous leg ulcer. The web spaces of the toes should be examined for evidence of fungal infection. Skin swabs are usually unhelpful. Confirmation of infection is best done serologically by streptococcal titres: antistreptolysin O titre (ASOT) and antiDNAse B titre (ADB).

Erysipelas is the term used for a more superficial infection (often on the face) of the dermis and upper subcutaneous layer that clinically presents with a well-defined edge. However, erysipelas and cellulitis overlap so it is often impossible to make a meaningful distinction.

Necrotizing fasciitis (see p. 125).

Treatment is with phenoxymethylpenicillin (or erythromycin) and flucloxacillin (all 500 mg four times daily). If disease is widespread, treatment with antibiotics should be given intravenously for 3–5 days followed by at least 2 weeks of oral therapy. It is also necessary to treat any identifiable underlying cause. If cellulitis is recurrent, low-dose antibiotic prophylaxis (e.g. phenoxymethylpenicillin 500 mg twice daily) is used as each episode will cause further lymphatic damage. Post-cellulitic oedema is common and predisposes to further episodes of cellulitis.

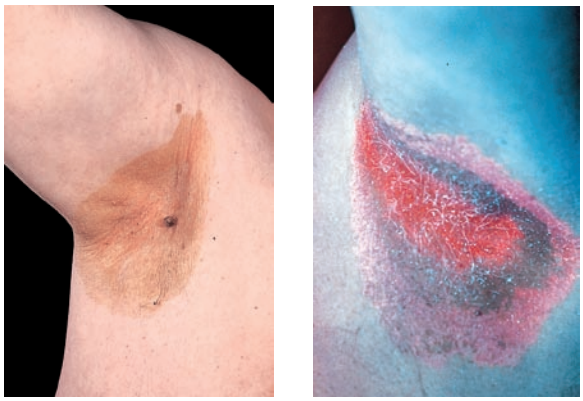
Ecthyma

Ecthyma is also an infection due to *Streptococcus* or *Staphylococcus aureus* or occasionally both. It presents as chronic well-demarcated, deeply ulcerative lesions sometimes with an exudative crust. It is commoner in developing countries, being associated with poor nutrition and hygiene. It is rare in the UK but is seen more commonly in intravenous drug abusers and people with HIV.

Treatment is with phenoxymethylpenicillin and flucloxacillin (both 500 mg four times daily) for 10–14 days.

Erythrasma

Erythrasma is caused by *Corynebacterium minutissimum*. It usually presents as an orange-brown flexural rash, and is often seen in the axillae or toe web spaces (Fig. 23.3). It is frequently misdiagnosed as a fungal infection. The rash shows a dramatic coral pink fluorescence under Wood's (ultraviolet) light.



(a) (b)

Fig. 23.3 **Erythrasma of the axilla**, showing pink fluorescence under a Wood's lamp.

Treatment is with topical sodium fusidate 3 times daily for 7 days or oral erythromycin 500 mg four times daily for 7–10 days.

Folliculitis

Folliculitis is an inflammation of the hair follicle. It presents as itchy or tender papules and pustules. *Staphylococcus aureus* is frequently implicated. It is commoner in humid climates and when occlusive clothes are worn. A variant occurs in the beard area (called 'sycosis barbae'), which is commoner in black Africans. This is probably caused by the ability of shaved hair to grow back into the skin, especially if the hair is naturally curly. Extensive, itchy folliculitis of the upper trunk and limbs should alert one to the possibility of underlying HIV infection. Folliculitis following use of hot tubs is due to *Pseudomonas ovale*.

Treatment is with topical antiseptics, topical antibiotics (e.g. sodium fusidate) or oral antibiotics (e.g. flucloxacillin 500 mg or erythromycin 500 mg both four times daily for 2–4 weeks).

Boils (furuncles)

Boils are a rather more deep-seated infection of the skin, often caused by *Staphylococcus*. These can cause painful red swellings. They are commoner in teenagers and often recurrent. Recurrent boils may rarely occur in diabetes mellitus or in immunosuppression. Large boils are sometimes called 'carbuncles'. Swabs should be taken to check antibiotic sensitivity as community acquired MRSA is an increasingly common cause.

Treatment is with oral antibiotics (e.g. erythromycin 500 mg four times daily for 10–14 days) and occasionally by incision and drainage.

Antiseptics such as povidone iodine or chlorhexidine (as soap) and using a bath oil can be useful in prophylaxis.

Hidradenitis suppurativa

This condition is characterized by a painful, discharging, chronic inflammation of the skin at sites rich in apocrine glands (axillae, groins, natal cleft). The cause is unknown but it is commoner in females and within some families it appears to be inherited in an autosomal dominant fashion. The initial lesion is occlusion of the hair follicle producing an inflammatory reaction that spreads to the apocrine glands. Clinically it presents after puberty with papules, nodules and abscesses which often progress to cysts and sinus formation. With time, scarring may arise. The condition follows a chronic relapsing/remitting course and is often worse in obese individuals.

Treatment is difficult but weight loss, 'prophylactic' antibiotics, oral retinoids, zinc and co-cyprindiol (2 mg cyproterone acetate + 35 µg ethinylestradiol in females only) have been tried. They should be used as for acne vulgaris (p. 1245). Severe recalcitrant cases have been treated with intravenous infliximab, a monoclonal antibody (p. 76). Surgery and skin grafting is sometimes required.

Pitted keratolysis

This is a superficial infection of the horny layer of the skin caused by a corynebacterium. It frequently involves the soles of the forefoot and appears as numerous small punched-out circular lesions of a rather macerated skin (e.g. as seen after prolonged immersion). There may be an associated hyperhidrosis of the feet and a prominent odour.

Treatment. Topical antibiotics (e.g. sodium fusidate or clindamycin, applied three times daily for 2–4 weeks) and topical anti-sweating lotions are effective therapies.

MYCOBACTERIAL INFECTIONS

Leprosy (Hansen's disease) (p. 140)

Leprosy usually involves the skin, and the clinical features depend on the body's immune response to the organism *Mycobacterium leprae*.

Indeterminate leprosy is the commonest clinical type, especially in children. This presents as small, hypopigmented or erythematous circular macules with occasional mild anaesthesia and scaling. This may resolve spontaneously or progress to one of the other types. Biopsy reveals a perineural granulomatous infiltrate and scant acid-fast bacilli.

Tuberculoid leprosy presents with a few larger, hypopigmented (see Fig. 4.30) or erythematous plaques with an active erythematous, often raised, rim. Lesions are usually markedly anaesthetic, dry and hairless, reflecting the nerve damage. Nerves may be enlarged and palpable. Biopsy shows a granulomatous infiltrate centred on nerves but no organisms.

Lepromatous leprosy presents with multiple inflammatory papules, plaques and nodules. Loss of the eyebrows ('madarosis') and nasal stuffiness are common. Skin thickening and severe disfigurement may follow. Anaesthesia is much less prominent. Biopsy shows numerous acid-fast bacilli.

Diagnosis and treatment are discussed on page 141.

Skin manifestations of tuberculosis

Tuberculosis can occasionally cause skin manifestations:

- *Lupus vulgaris* usually arises as a post-primary infection. It usually presents on the head or neck with red-brown nodules that look like apple jelly when pressed with a glass slide ('diascopy'). They heal with scarring and new lesions slowly spread out to form a chronic solitary erythematous plaque. Chronic lesions are at high risk of developing squamous cell carcinoma.
- *Tuberculosis verrucosa cutis* arises in people who are partially immune to tuberculosis but who suffer a further direct inoculation in the skin. It presents as warty lesions on a 'cold' erythematous base.
- *Scrofuloderma* arises when an infected lymph node spreads to the skin causing ulceration, scarring and discharge.
- *The tuberculides* are a group of rashes caused by an immune manifestation of tuberculosis rather than direct infection. Erythema nodosum is the commonest and is discussed on page 1250. Erythema induratum ('Bazin's disease') produces similar deep red nodules but these are usually found on the calves rather than the shins and they often ulcerate.

Atypical mycobacteria

Atypical (non-tuberculous) mycobacteria can occasionally infect the skin. *Mycobacterium marinum* is found in fish tanks and occasionally swimming pools. It can gain access via a break in the skin and then causes deep granulomatous nodules, often in a linear fashion.

VIRAL INFECTIONS

Viral exanthem

This is probably the commonest type of virus-induced rash and presents clinically as a widespread non-specific erythematous maculopapular rash. It probably arises due to circulating immune complexes of antibody and viral antigen localizing to dermal blood vessels. The rash can be caused by many different viruses (e.g. echovirus, erythrovirus, human herpes virus-6, Epstein–Barr virus; see p. 107) and so is rarely diagnostic. The rash will resolve spontaneously in 7–10 days.

Slapped cheek syndrome (erythema infectiosum, fifth disease)

This affects children and is caused by erythrovirus B19 (see p. 109). It is a mild viral illness which is followed by an intense erythema on the cheeks ('slapped cheeks') and a reticulate erythema on the proximal limbs.

Herpes simplex virus (see also p. 104)

Herpes simplex virus (HSV) occurs as two genomic subtypes. HSV type 1 is spread by direct contact and droplet infection. Most people are affected in early childhood but the infection is usually subclinical. Occasionally it can cause a self-limiting pyrexial primary illness with either clusters of painful blisters on the face or a painful gingivostomatitis.

Ask the authors
www.studentconsult.com

Once the person is infected, cell-mediated immunity develops. In some individuals this response is poor and they get recurrent attacks of HSV, often manifest as cold sores. Immunosuppression can also cause a recrudescence of HSV. HSV can also autoinoculate into sites of trauma and present as painful blisters/pustules which may be seen for example on the fingers of healthcare workers ('herpetic whitlow').

HSV type 2 infections occur mainly after puberty and usually affect the genital area. Infections are often symptomatic and transmitted sexually. However, HSV type 1 can also be found in the genital area due to orogenital contact.

Other rare complications of HSV infection include corneal ulceration, eczema herpeticum, chronic perianal ulceration in AIDS patients and erythema multiforme (p. 1250).

Treatment

Oral valaciclovir (500 mg twice daily for 5 days) is used for primary HSV and painful genital HSV. Cold sores are treated with aciclovir cream but this must be used early to be effective in shortening an attack; recurrent sores are treated with oral therapy. Attacks of genital herpes become less frequent with time. Intravenous aciclovir must be used in immunosuppressed patients.

Varicella zoster virus

Varicella zoster virus (VZV) causes the common childhood infection called chickenpox. It is discussed on page 107. It also causes herpes zoster.

Herpes zoster (shingles)

'Shingles' results from a reactivation of VZV. It may be preceded by a prodromal phase of tingling or pain, which is then followed by a painful and tender blistering eruption in a dermatomal distribution (Fig. 23.4 and Fig. 4.17, p. 106). The blisters occur in crops, may become pustular and then crust over. The rash lasts 2–4 weeks and is usually more severe in the elderly. Occasionally more than one dermatome is involved.

Complications of shingles include severe, persistent pain (post-herpetic neuralgia), ocular disease (if the ophthalmic nerve is involved) and rarely motor neuropathy.

Treatment

Herpes zoster requires adequate analgesia and antibiotics (if secondary bacterial infection is present). Valaciclovir 1 g or famciclovir 250 mg three times daily for 7 days is used, or oral aciclovir 800 mg, five times daily for 7 days helps shorten the attack if given early in the illness. High-dose intravenous aciclovir is used in immunosuppressed patients. It remains unclear how useful aciclovir therapy is in preventing prolonged post-herpetic neuralgia. Centers for Disease Control (CDC) have recommended herpes zoster vaccination in people over 60 years.

Fig. 23.4 **Herpes zoster in an African** (courtesy of Dr P Matondo, Lusaka, Zambia).

Human papilloma virus

Human papilloma virus (HPV) is responsible for the common cutaneous infection of 'viral warts'. There are more than 70 subtypes as detected by DNA hybridization. All can cause overgrowth of differentiated squamous epithelium.

Common warts are papular lesions with a coarse roughened surface, often seen on the hands and feet, but also on other sites. Small black dots (bleeding points) are often seen within the lesion (Fig. 23.5). Children and adolescents are usually affected. Spread is by direct contact and is also associated with trauma.

Plantar warts (verrucae) is the term used for lesions on the soles of the feet. They often appear flat ('inward growing') although they have the same papillomatous surface change and black dots are often revealed if the skin is pared down (unlike callosities). Warts may be painful or tender if they are over pressure points or around nail folds.

Filiform warts occur on the face, at the nasal vestibule or around the mouth. They are elongated with a horny cap.

Plane warts are much less common and are caused by certain HPV subtypes. They are clinically different and appear as very small, flesh-coloured or pigmented, flat-topped lesions (best seen with side-on lighting) with little in the way of surface change and no black dots within them. They are usually multiple and are frequently found on the face or the backs of the hands.

Anogenital warts are usually seen in adults and are normally transmitted by sexual contact. They are rare in childhood and, whilst child sex abuse should always be considered, it should be remembered they may well have been transmitted through non-sexual contact. HPV subtypes 16 and 18 are potentially oncogenic and are associated with cervical and anal carcinomas.

Treatment

Common warts on the skin are surprisingly difficult to treat effectively but they almost always resolve spontaneously after months to years (with no scarring), presumably due to cell-mediated immune recognition. When they do resolve, they tend to do so rapidly within a few days.

Fig. 23.5 **Viral wart.**

Regular use of a topical keratolytic agent (e.g. 2–10% salicylic acid) over many months with weekly paring of the lesion helps speed up resolution in some patients and remains the mainstay of treatment. A course of cryotherapy (freezing) can also help. Cautery, surgery, carbon dioxide laser, alpha-interferon injection and bleomycin injection have all been used with variable success but are not recommended, as treatments may be very painful and can cause permanent scarring.

Genital warts (see p. 181) are usually treated with cryotherapy, trichloroacetic acid, 5% imiquimod cream or topical podophyllin. Patients with genital warts (and their sexual partners) must be screened for other sexually transmitted diseases.

HPV vaccines against the high risk oncogenic subtypes (6, 11, 16 and 18) have recently become available; they should help protect against cervical neoplasia but will have no impact on common warts. It remains controversial whether this vaccine should be given to girls and boys routinely.

Molluscum contagiosum ('water blisters')

Molluscum contagiosum is a common cutaneous infection of childhood caused by the pox virus. Clinically, lesions are multiple small (1–3 mm) translucent papules which often look like fluid-filled vesicles but are in fact solid. Individual lesions may have a central depression called a punctum. They exhibit the Köbner phenomenon (p. 1241). They can occur at any body site including the genitalia. Transmission is by direct contact. Occasionally lesions may be up to 1 cm in diameter ('giant molluscum'). They are said to be more extensive in children with atopic eczema, which may just reflect that scratching aids their spread.

They usually continue to occur in crops over 6–12 months and rarely require treatment as they spontaneously resolve. Any form of localized trauma, including scratching, helps speed up resolution, and cryotherapy is used in older children. One per cent stabilized hydrogen peroxide cream or 5% imiquimod cream is used in younger children. Molluscum in an adult, especially if giant, should raise the underlying possibility of immunosuppression, especially HIV infection.

Orf

Orf is a disease of sheep (and occasionally goats) due to a pox virus infection. It causes a vesicular and pustular rash around the mouths of young lambs. People who come into contact with the affected fluid may develop lesions on the hands. Clinically these appear as 1–2 cm reddish papules with a surrounding erythema which usually become pustular. The lesion(s) resolves spontaneously after 4–6 weeks and immunity is lifelong. Occasionally orf is complicated by erythema multiforme (p. 1250).

FUNGAL INFECTIONS

Fungi are primitive, saprophytic organisms found throughout our environment. Fungal skin disease (mycosis) has a high prevalence in humans with 'thrush' and 'athlete's foot' being two of the commonest examples. In the immunosuppressed, mycoses can be widespread and life-threatening. There are three groups of pathogenic fungi that commonly affect the outer layer of skin or keratinizing epithelium: dermatophytes, *Candida albicans* and pityrosporum.

Dermatophyte infection

By definition, dermatophytes cause a 'ringworm' type of rash. The three main genera responsible are *Trichophyton, Microsporum* and *Epidermophyton*. These organisms are identified by microscopy and culture of skin, hair or nail samples. The clinical appearance of mycoses depends in part on the organism involved, the site affected and the host reaction. All are spread by direct contact from other humans or from infected animals. The use of communal showers and swimming baths and the sharing of towels or sportswear aids indirect fomite transmission.

Tinea corporis

Ringworm of the body usually presents with slightly itchy, asymmetrical, scaly patches which show central clearing and an advancing, scaly, raised edge. Occasionally vesicles or pustules may be seen in the edge. Central clearing is not a universal feature and it is recommended that all asymmetrical scaly lesions should be scraped for fungus. Ringworm of the face (tinea faciei) often arises after the use of topical steroids. It tends to be more erythematous and less scaly than trunk lesions and it may become itchy after sun exposure.

Tinea cruris

Ringworm of the groin is extremely common world-wide. Early on, the lesions appear as well-demarcated red plaques with an arc-like border extending down the upper thigh (Fig. 23.6). Central clearing may appear and a few pustules or vesicles may be seen if inflammation is intense. Satellite lesions, suggestive of *Candida*, are not present.

Tinea pedis

Athlete's foot may be confined to the toe clefts where the skin looks white, macerated and fissured. It may also be more diffuse, usually causing a diffuse scaly erythema of the soles spreading on to the sides of the foot. Annular lesions are rarely present. There may be an associated hyperhidrosis and fungal involvement of one or more toe nails. In severe infection a strong inflammatory reaction can occur causing pustules or blistering and this often leads to a misdiagnosis of pompholyx-type eczema.

Fig. 23.6 **Tinea cruris** – ringworm of the groin.

Tinea manuum

Ringworm of the hands also presents with a diffuse erythematous scaling of the palms with variable skin peeling and skin thickening. Annular lesions are rare at this site.

Tinea capitis

Scalp ringworm infections are continuing to increase in developed countries although the rate of increase is beginning to slow. Fungus may confine itself to within the hair shaft (endothrix) or spread out over the hair surface (ectothrix). The latter can cause fluorescence under a Wood's lamp (ultraviolet light). Scalp ringworm is spread by close contact (especially in schools and households) and may also be spread indirectly by hairdressers. The number of new cases has risen enormously in the large cities in developed countries. Increase in travel and immigration has allowed the spread of different pathogenic fungi (e.g. *Trichophyton tonsurans* from Central America, *Trichophyton violaceum* from India and Pakistan) into new countries where there is overcrowding and poor social conditions. The majority of UK cases are due to *T. tonsurans* (which does not show fluorescence).

Tinea capitis is much commoner in children, especially those of black African origin whose scalp and hair seems more susceptible to fungal invasion. The clinical appearance of scalp ringworm is highly variable from a mild diffuse scaling with no hair loss (similar to dandruff) to the more typical appearance of circular scaly patches in the scalp with associated alopecia and broken hairs. As the host's immune response increases a few pustules may appear and an exudate may be present. At worst, a full-blown 'kerion' develops; a boggy swollen mass with copious quantities of discharging pus and exudate accompanied by severe alopecia. This is still poorly recognized and inappropriately treated with antibiotics and attempted surgical drainage.

Extensive infection is occasionally accompanied by a widespread papulopustular rash on the trunk. This is a so-called 'id reaction' and probably relates to the host immune response to the fungus. It seems commoner in black African children. It resolves when the fungal infection is treated.

Tinea unguium

Ringworm of the nails is increasingly common with age and frequently ignored as it is often asymptomatic. Clinically this presents as asymmetrical whitening (or yellowish black discoloration) of one or more nails which usually starts at the distal or lateral edge before spreading throughout the nail (Fig. 23.7). The nail plate appears thickened. Crumbly white material appears under the nail plate and this is the best specimen to obtain for mycology sampling. The nail plate may become destroyed with advanced disease.

Fig. 23.7 **Dermatophyte infection of the nail showing white crumbling dystrophy.**

'Tinea incognito'

This is the term used to describe a fungal skin infection that has been modified by therapy with a topical steroid. The clinical appearance is variable but may show a non-specific erythema with little in the way of scaling or a few reddish nodules. The history of the rash improving with treatment (owing to the suppression of inflammation) but worsening and spreading every time it is stopped is typical. Skin scrapings for mycology or even a biopsy should confirm the diagnosis.

Treatment

Localized ringworm of the body or flexures is treated with topical antifungal creams (clotrimazole, miconazole, terbinafine or amorolfine applied three times daily for 1–2 weeks). More widespread infection, tinea pedis, tinea manuum and tinea capitis require oral antifungal therapy. Itraconazole (100 mg daily) and terbinafine (250 mg daily) are the most effective drugs used for periods of 1–2 months. High-dose griseofulvin is still used in children for scalp ringworm (15–20 mg/kg per day for 8 weeks).

Tinea unguium of the toe-nails is the most resistant to treatment. Itraconazole (100 mg daily or 200 mg twice daily for 1 week per month 'pulsed therapy') or terbinafine (250 mg daily) given for 3–6 months will cure up to about 80% of cases.

Candida albicans (see also p. 149)

Candida albicans is a yeast that is sometimes found as part of the body's flora, especially in the gastrointestinal tract. It acts as an opportunist, taking hold in the skin when there is a suitable warm moist environment such as in nappy rash (p. 1268) or intertrigo in obese individuals (Fig. 23.8).

The flexural areas affected are red with a rather ragged peeling edge that may contain a few small pustules. Small circular areas of erythema or small papules and pustules may be seen in front of the advancing edge (satellite lesions).

Fig. 23.8 **Intertrigo with satellite lesions typical of candidiasis.**

Candida may also affect the moist interdigital clefts of the toes and mimic tinea pedis. In people who have their hands immersed frequently in water (e.g. cleaners, nurses) *Candida* may cause infection in the macerated skin of the finger web spaces or the damaged skin around the nail folds ('chronic paronychia'). Nail infection may mimic tinea unguium. *Candida* can infect mucosal surfaces of the mouth or genital tract. This tends to occur in patients taking broad-spectrum antibiotics (due to suppression of protective bacterial flora) or in immunosuppressed patients. Clinically, superficial white or creamy pseudomembranous plaques appear which can be scraped off leaving raw areas underneath.

Treatment

Treatment is aimed at removing any underlying predisposing factor and applying topical antifungal creams, e.g. clotrimazole or miconazole (or the equivalent as mouth lozenges/pessaries). *Candida* nail infections require systemic antifungal therapy with an imidazole such as itraconazole (100 mg daily for 3–6 months). Recurrent candidiasis is relatively common, especially in women. Diabetes mellitus should always be excluded. Repeated topical treatment or an oral imidazole may be needed.

Pityrosporum

This yeast occurs as part of the normal flora of human skin. Colonization is prominent in the scalp, flexures and upper trunk. There are two morphological variants, *Pityrosporum ovale* and *Pityrosporum orbiculare*, and the mycelial form of this yeast is called *Malassezia furfur*. *Pityrosporum* can overgrow in some individuals and has been implicated in three dermatoses:

- pityriasis versicolor
- seborrhoeic eczema (p. 1238)
- pityrosporum folliculitis.

Pityriasis versicolor

This is a relatively common condition of young adults caused by infection with *Pityrosporum*. In Caucasians it presents most commonly on the trunk with reddish brown scaly macules, which are asymptomatic. In black-skinned individuals (or in whites who are sun-tanned) it more commonly presents as macular areas of hypopigmentation. Inappropriate use of topical steroids tends to spread the rash.

Diagnosis can be confirmed by skin scrapings or Wood's light examination (yellow fluorescence).

Treatment is with selenium sulphide or 2% ketoconazole shampoo (apply to body and remove after 30–60 minutes and repeat daily for 1 week) or a topical imidazole cream (twice daily for 10 days). Oral itraconazole (100 mg twice daily for 1 week) can be used for resistant cases. The pigmentation takes months to recover even after successful treatment. The condition may recur but can be re-treated.

Pityrosporum folliculitis

This is common in young adult males and characterized by small itchy papules and pustules on the upper back which are centred on hair follicles. It is commoner in people with Down's syndrome. It responds well to ketoconazole shampoo or a topical imidazole cream (twice daily for 2 weeks).

INFESTATIONS

Scabies

Scabies is an intensely itchy rash caused by the mite *Sarcoptes scabiei*. It can affect all races and people of any social class. It is most common in children and young adults but can affect any age group. There are 300 million cases of scabies in the world each year. It is commoner in poorer countries with social overcrowding.

Scabies is spread by prolonged close contact such as within households or institutions, and by sexual contact. It presents clinically with itchy red papules (or occasionally vesicles and pustules) which can occur anywhere in the skin but rarely on the face, except in neonates. The distribution of lesions is often suggestive of the diagnosis (Fig. 23.9). Sites of predilection are between the web spaces of the fingers and toes, on the palms and soles, around the wrists and axillae, on the male genitalia, and around the nipples and umbilicus.

The pathognomonic sign is of linear or curved skin burrows but these are not always present. The pruritus is normally worse at night. Excoriations and secondary bacterial infection may complicate the rash. Scabies can be confirmed by taking skin scrapings of a lesion and examining a potassium hydroxide preparation for the mite and/or its eggs by microscopy. Sometimes mites and eggs can be seen with a dermatoscope.

Treatment

This involves application of a topical scabicide (e.g. 5% permethrin) which is washed off after 10 hours. For the treatment to be successful the following factors should be noted:

- All the skin below the neck should be treated, including the genitalia, palms and soles, and under the nails. Treat the head and neck regions in infants (up to age 2 years).
- All close contacts should be treated at the same time even if asymptomatic.
- Reapply scabicide to the hands if they are washed during the treatment period.
- Washing or cleaning of recently worn clothes (preferably at 60°C is recommended to avoid re-infection.

Fig. 23.9 **Scabies** – itchy papules and pustules centred on the web spaces of the hand.

- Patients should be warned that the pruritus may persist for up to 4 weeks after successful treatment. Adjunct treatment with crotamiton cream, an emollient or a topical steroid is helpful.
- The application should be repeated after one week.
- A patient information leaflet about therapy helps improve compliance.

Malathion can be used if permethrin is unavailable; benzyl benzoate is used occasionally but is very irritant and should not be used in children. Lindane is a cheap therapy which is still used in many countries but there are concerns about resistance to this drug and possible neurotoxic side-effects.

Crusted scabies (Norwegian scabies)

Crusted scabies is a clinical variant that occurs in immunosuppressed individuals where huge numbers of mites are carried in the skin. Patients are not always itchy but they are extremely infectious after relatively minimal contact, which is unfortunate as the diagnosis is often delayed. Clinically the condition presents as hyperkeratotic crusted lesions, especially on the hands and feet. Lesions may progress such that the patient has a widespread erythema with irregular crusted plaques. It can therefore mimic infected eczema or psoriasis.

Treatment is with careful barrier nursing, repeated applications of a scabicide and, in resistant cases, oral ivermectin (200 μg/kg – two doses 1 week apart).

Lice infection

Lice are blood-sucking ectoparasites that can affect man in three ways.

Head lice (pediculosis capitis)

Head lice is a common infection world-wide affecting predominantly children and being commoner in females. Spread is by direct contact and encouraged by overcrowding. It usually presents with itch or scalp excoriations. Occasionally erythematous papules on the neck may be seen.

Diagnosis can be confirmed by the presence of eggs ('nits') seen tightly bound to the hair shaft. Adult lice may be seen rarely in heavy infection. School nurses and parents are usually adept at detecting the lice.

Treatment. Eradication is difficult because of non-compliance as well as resistance patterns. Malathion, carbaryl and phenothrin (two applications 7 days apart with a contact time of 12 hours) are the most commonly used agents. Nit combing of wet hair alone is not as effective as chemical treatments and is now not used. Policies of rotating insecticides are also outmoded. If one treatment fails, a different insecticide is used for the next course in an individual. Treatment is usually repeated after 7 days and metal nit combs may help remove the eggs.

Body lice (pediculosis corporis)

Infestation with body lice is a disease of poverty and neglect. It is rarely seen in developed countries except in homeless individuals and vagrants. It is spread by direct contact or sharing infested clothing. The lice and eggs are rarely seen on the patient but are commonly found on the clothing. It presents with itch, excoriations and sometimes post-inflammatory hyperpigmentation of the skin.

Treatment consists of malathion or permethrin for the patient and high-temperature washing and drying of clothing.

Pubic lice (crabs, phthiriasis pubis) (p. 183)

Pubic lice are transmitted by direct contact, usually sexual. Infestation presents with itching, especially at night. Lice can be seen near the base of the hair with eggs somewhat further up the shaft. Occasionally eyebrows, eyelashes and the beard area are affected.

Treatment is as for head lice but all sexual contacts should be treated and other sexually transmitted diseases should be screened for. Pubic lice of the eyelashes is treated with white soft paraffin three times daily for 1–2 weeks.

Arthropod-borne diseases ('insect bites' or papular urticaria)

These depend on contact with an animal (e.g. dog, cat, bird) that is infested with fleas (*Cheyletiella*) or on bites from flying insects (e.g. midges, mosquitoes). In the case of flea bites the animal itself may be itchy with scaly and thickened skin. These fleas can also live in soft furnishings (e.g. carpets and beds) even after the animal has been removed. Bites present as itchy urticated lesions, which are often grouped in clusters. The legs are most commonly affected. It is not unusual for an individual to react badly to bites when other family members seem unaffected. Anti-flea treatment of the animal and furnishings is required. Insect repellents and appropriate clothing help lessen bites from flying insects.

Infestations with bed bugs have increased enormously in the past decade in the developed world, possibly due to being imported from foreign travel/tourism. They can be seen as small brown/black lentil-sized insects at night-time as they are attracted to warmth and carbon dioxide of sleeping humans. Infestations require professional insecticide treatments of properties. Advice can be obtained from the National Pest Technicians Association (www.npta.org.uk).

FURTHER READING

Colgrove J. The ethics and politics of compulsory HPV vaccination. *New England Journal of Medicine* 2006; **355**: 2389–2391 (see also **356**: 1074–1075).

Cox NH. Oedema as a risk factor for multiple episodes of cellulitis/erysipelas of the lower leg: a series with community follow-up. *British Journal of Dermatology* 2006; **155**: 947–950.

Daum RS. Skin and soft tissue infections caused by methicillin-resistant staphylococcus aureus. *New England Journal of Medicine* 2007; **357**: 380–390.

Gibbs S, Harvey I. Topical treatments for cutaneous warts. *Cochrane Database of Systematic Reviews* 2006; **3**: CD0001781.

Heukelbach J, Feldmeier H. Scabies. *Lancet* 2006; **367**: 1767–1774.

Ko CJ, Elston DM. Pediculosis. *Journal of the American Academy of Dermatology* 2004; **50**: 1–12.

Steen CJ, Carbonaro. P, Schwartz R. Arthropods in dermatology. *Journal of the American Academy of Dermatology* 2004; **50**: 819–842.

PAPULO-SQUAMOUS/ INFLAMMATORY RASHES

ECZEMA

Introduction

The term 'eczema' derives from the Greek word for 'boiling', which reflects that the skin can become so acutely inflamed that fluid weeps out or vesicles appear. It is synonymous with the term dermatitis and the two words are interchangeable. In the developed world eczema accounts for a large proportion of skin disease, both in hospital and community based populations. It is estimated that 10% of people have some form of eczema at any one time, and up to 40% of the population will have an episode of eczema during their lifetime.

All eczemas (see Table 23.3) have some features in common and there is a spectrum of clinical presentation from acute through to chronic. Vesicles or bullae may appear in the acute stage if inflammation is intense. In subacute eczema the skin can be erythematous, dry and flaky, oedematous and crusted (especially if secondarily infected). Chronic persistent eczema is characterized by thickened or lichenified skin. Eczema is nearly always itchy. Histologically 'eczematous change' refers to a collection of fluid in the epidermis between the keratinocytes ('spongiosis') and an upper dermal perivascular infiltrate of lymphohistiocytic cells. In more chronic disease there is marked thickening of the epidermis ('acanthosis').

Atopic eczema

This type of eczema (often called 'endogenous eczema') occurs in individuals who are 'atopic' (p. 846). It is common, occurring in up to 5% of the UK population. It is commoner in early life, occurring at some stage during childhood in up to 10–20% of all children.

Aetiology

The exact pathophysiology is not fully understood but there is an initial selective activation of Th2-type CD4 lymphocytes in the skin which drives the inflammatory process. This precedes the chronic phase when Th0 and Th1 cells predominate. In at least 90% of cases there is a raised serum total IgE level. Atopic eczema is a genetically complex, familial disease with a strong maternal influence. A positive family history of atopic disease is often present: there is a 90% concordance in monozygotic twins but only 20% in dizygotic twins. If one parent has atopic disease, the risk for a child of developing eczema is about 20–30%. If both parents have atopic eczema the risk is greater than 50%.

Table 23.3 Classification of eczema

Endogenous	Exogenous
Atopic eczema	Contact eczema – irritant
Discoid eczema	Contact eczema – allergic
Hand eczema	Photosensitive eczema
Seborrhoeic eczema	Lichen simplex/nodular prurigo
Venous ('gravitational') eczema	
Asteatotic eczema	

Recent studies have shown abnormal skin-homing T cells in eczema compared to controls. However the latest genetic research points towards a primary problem of skin barrier function, suggesting the immunological changes are secondary. Loss-of-function mutations in the epidermal barrier protein filaggrin cause ichthyosis vulgaris but can predispose to atopic eczema in Caucasian individuals. Different mutations in the same gene filaggrin have been found in Japanese patients with eczema but this gene may not be significant in Bangladeshi or black African patients. Filaggrin is one of the genes in the epidermal differentiation complex on chromosome 1q. Very strong linkage to this region would suggest that other genes in this area are also involved in the development of atopic eczema (e.g. loricrin, involucrin, S100 calcium-binding proteins). Certainly filaggrin is only one of a number of genes involved in eczema – linkage is also found to 3q21, 3p22–24, 17q25, 20p. Other candidate genes under study include *SPINK 5*, mast cell chymotryptase and peptidoglycan recognition proteins.

Exacerbating factors

Infection either in the skin or systemically can lead to an exacerbation, possibly by a superantigen effect. Paradoxically, lack of infection (in infancy) may cause the immune system to follow a Th2 pathway and allow eczema to develop (the so-called 'hygiene hypothesis'). Strong detergents, chemicals and even woollen clothes can be irritant and exacerbate eczema. Teething is another factor in young children. Severe anxiety or stress appears to exacerbate eczema in some individuals. Cat and dog fur can certainly make eczema worse, possibly by both allergic and irritant mechanisms. The role of house dust mites and diet is less clear cut. There is some evidence that food allergens may play a role in triggering atopic eczema and that dairy products cause exacerbation of eczema in a few infants under 12 months of age.

Clinical features

Atopic eczema can present as a number of distinct morphological variants. The commonest presentation is of itchy erythematous scaly patches, especially in the flexures such as in front of the elbows and ankles, behind the knees (Fig. 23.10) and around the neck. In infants eczema often starts on the face before spreading to the body. Very acute lesions may weep or exude and can show small vesicles. Scratching can produce excoriations, and repeated rubbing produces skin thickening (lichenification) with exaggerated skin markings.

In patients with pigmented skin, eczema often shows a reverse pattern of extensor involvement. Also the eczema

Fig. 23.10 **Atopic eczema behind the knees.**

may be papular or follicular in nature, and lichenification is common. A final problem in pigmented skin is of post-inflammatory hyper- or hypopigmentation which is often very slow to fade after control of the eczema.

Associated features

Involvement of the nail bed may produce pitting and ridging of the nails. In some atopic individuals the skin of the upper arms and thighs may feel roughened due to follicular hyperkeratosis ('keratosis pilaris'). The palms may show very prominent skin creases ('hyperlinear palms'). There may be an associated dry 'fish-like' scaling of the skin which is non-inflammatory and often prominent on the lower legs ('ichthyosis vulgaris').

Complications

Broken skin commonly becomes secondarily infected by bacteria, usually *Staphylococcus aureus*, although streptococci can colonize eczema, especially in macerated flexural areas such as the neck and groin. Clinically this infection may appear as crusted, weeping impetigo-like lesions. Occasionally *Pseudomonas* can be grown from skin swabs but this rarely causes a clinical problem. Cutaneous viral infections (e.g. viral warts and molluscum) are often widespread in atopic eczema and are probably spread by scratching. HSV can cause a widespread eruption called eczema herpeticum (Kaposi's varicelliform eruption). This can occasionally be a very severe infection and rarely can be fatal. It appears as multiple small blisters or punched-out crusted lesions associated with malaise and pyrexia, and needs rapid treatment with oral (or intravenous if severe) aciclovir. Ocular complications of atopic eczema include conjunctival irritation and less commonly keratoconjunctivitis and cataract. Retarded growth may be seen in children with chronic severe eczema; it is due to the disease itself and not the use of topical steroids.

Investigations

The diagnosis of atopic eczema is normally clinical. Atopy is characterized by high serum IgE levels or high specific IgE levels to certain ingested or inhaled antigens. The latter can be tested by radio-immunoabsorbent assay (RAST tests) of blood, or indirectly by skin-prick testing (p. 853). A peripheral blood eosinophilia may also be seen.

Prognosis

The majority (80–90%) of children with early-onset atopic eczema will spontaneously improve and 'clear' before the teenage years, 50% being clear by the age of 6. A few will get a recurrence as adults, even if just as hand eczema. However, if the onset of eczema is late in childhood or in adulthood, the disorder follows a more chronic remitting/relapsing course.

Treatment *(Box 23.2)*

General measures

These include avoiding known irritants (especially soaps or furry animals), wearing cotton clothes, and not getting too hot. Manipulating the diet (e.g. dairy-free diet) is rarely beneficial except in a few children, especially those under 12 months of age. Any change in diet should be made under supervision, especially with growing children who may need supplements such as calcium.

Box 23.2 Management of atopic eczema

- Education and explanation
- Avoidance of irritants/allergens
- Emollients
- Bath oils/soap substitutes
- Topical therapies:
  - steroids
  - immunomodulators
- Adjunct therapies:
  - oral antibiotics
  - sedating antihistamines
  - bandaging
- Phototherapy/systemic therapy (for severe cases)

Table 23.4 Classification of topical steroids by potency

Very potent	0.05% clobetasol propionate 0.3% diflucortolone valerate
Potent	0.1% betamethasone valerate 0.025% fluocinolone acetonide
Diluted potent	0.025% betamethasone valerate 0.00625% fluocinolone acetonide
Moderately potent	0.05% clobetasone butyrate 0.05% alclometasone dipropionate
Mild	2.5% hydrocortisone 1% hydrocortisone

Topical therapies

Topical therapies (p. 1270) are sufficient to control atopic eczema in most people. The 'triple' combination of topical steroid, frequent emollients (see Table 23.17) and bath oil and soap substitute (e.g. aqueous cream) helps.

Written information or a practical demonstration of how to apply these treatments improves compliance.

Use of topical steroids. Unjustified fear about the dangers of topical steroids has often led to undertreatment of eczema. Providing appropriate-strength steroid preparations are used for the right body site, these compounds can be used quite safely on a long-term intermittent basis. Topical steroids can be divided into five groups depending on their potency (Table 23.4).

The following guidelines should be followed to allow their safe use in common chronic inflammatory skin conditions.

- The face should be treated only with mild steroids.
- In adults the body should be treated with either mild, moderately potent or diluted potent steroids.
- In young children the body should be treated with mild and moderately potent steroids.
- Potent steroids are used for short courses (7–10 days).
- Treatment of the palms and soles (but not the dorsal surfaces) may require potent or very potent steroids as the skin is much thicker.
- Regular use of emollients may lessen the need for steroid use.
- Only use steroids on inflamed skin. Do not use as an emollient.
- 'Apply sparingly' means use sufficient to leave a glistening surface to the skin after application.
- Use weaker steroid preparations in flexures (e.g. the groin, and under breasts) as apposition of the skin at

these sites tends to occlude the treatment and increase absorption.

Topical immunomodulators

Tacrolimus ointment (0.1% and 0.03%) and the less potent pimecrolimus cream are used for atopic eczema in patients over 2 years old. They have the advantage over potent steroids of not causing skin atrophy and are thus very useful for treating sensitive areas such as the face and eyelids. They can be very irritant when first used (although this settles with continued use) and 9% of patients develop flushing after alcohol. The long-term side-effects remain unknown. They do not work so well on lichenified eczema, probably due to poor absorption. Current advice is to avoid vaccinations and sun exposure when using these agents. The milder potency steroid creams should still be considered as first-line therapy but tacrolimus is a useful alternative to excessive use of potent steroids.

Antibiotics

These are needed for bacterial infection and are usually given orally for 7–10 days. Flucloxacillin (500 mg four times daily) is effective against *Staphylococcus* and phenoxymethylpenicillin (500 mg four times daily) acts against *Streptococcus*. Erythromycin (500 mg four times daily) is useful if there is allergy to penicillin. Topical antiseptics are used in cases of recurrent infection but they can be irritant. They are usually added to the bath water rather than applied directly to the skin. Combination topical steroid/antibiotic creams can also be used for short periods but there is little evidence that they are better than topical steroids alone.

Sedating antihistamines

These (e.g. oral hydroxyzine hydrochloride 25 mg) are useful at night-time. They help by their sedative properties not by their antihistamine activity.

Bandaging

Paste bandaging can be useful for resistant or lichenified eczema of the limbs. It helps absorption of treatment and acts as a barrier to prevent scratching. Wet tubular gauze bandages are used for in-patient therapy but are difficult and time-consuming to use at home.

Second-line agents

These are used in severe non-responsive cases, especially if the eczema is significantly interfering with an individual's life (e.g. growth, sleeping, schoolwork or job). Ultraviolet phototherapy (see p. 1248), prednisolone (initial doses up to 30 mg daily), ciclosporin (3–5 mg/kg daily) and azathioprine (1–3 mg/kg daily) (p. 290) can all be effective treatments. However they all have side-effects and the risk/benefit ratio must be openly discussed with the patient before they are used.

Use of ciclosporin. Ciclosporin is a selective immunosuppressant that inhibits interleukin-2 production by T lymphocytes. A large number of other drugs interact with ciclosporin (e.g. erythromycin, NSAIDs) and should be avoided. Renal damage and hypertension are the two most serious side-effects so blood pressure and serum creatinine and eGFR should be measured every 6–12 weeks. Renal damage becomes increasingly common with time and tends to be dose-dependent but is mostly reversible. Hypertrichosis, paraesthesia and nausea are less serious side-effects. Pregnancy should be avoided.

Discoid eczema (nummular eczema)

Discoid eczema is a morphological variant of eczema characterized by well-demarcated scaly patches, especially on the limbs, and this can be confused sometimes with psoriasis. It is commoner in adults and can occur in both atopic and non-atopic individuals. It tends to follow an acute/subacute course rather than a chronic pattern. There is often an infective component (*Staphylococcus aureus*).

Hand eczema

Eczema may be confined to the hands (and feet). It can present with:

- itchy vesicles or blisters on the palm and along the sides of the fingers (also called 'pompholyx') (Fig. 23.11)
- a diffuse erythematous scaling and hyperkeratosis of the palms
- a scaling and peeling, most marked at the finger tips.

Hand eczema is not unusual in atopics but more frequently occurs in non-atopic individuals, and a cause is not always found. A history of contact with irritants (e.g. detergents, chemicals) and an occupational history should be sought, especially in finger-tip eczema. Patch testing for specific allergic or contact eczema should always be performed as up to 10% of individuals with hand eczema will show a positive test. Finally look for evidence of fungal infection as this can occasionally induce a secondary pompholyx of the hands or feet (a so-called 'id reaction').

Seborrhoeic eczema

Aetiology

Overgrowth of *Pityrosporum ovale* (also called *Malassezia furfur* in its hyphal form) together with a strong cutaneous immune response to this yeast produces the characteristic inflammation and scaling of seborrhoeic eczema. The condition is more common in parkinsonism as well as in HIV disease.

Clinical features

Seborrhoeic eczema affects body sites rich in sebaceous glands, although these do not appear to be causal. Three age groups are affected:

Fig. 23.11 **Pompholyx eczema** (courtesy of Dr A Bewley, London).

- *In childhood* it is common and presents in the first few months of life as 'cradle cap' in most babies. This may in part be due to the effect of maternal androgens on infant sebaceous glands. Yellowish, greasy, thick crusts are seen on the scalp. A more widespread erythematous, scaly rash can be seen over the trunk, especially affecting the nappy area. Unlike with atopic eczema, the child is normally unbothered as there is little associated pruritus. The rash normally improves spontaneously after a few weeks.
- *In young adults* (especially males) it occurs in 1–3% of the population. The rash is more persistent and presents as an erythematous scaling along the sides of the nose (Fig. 23.12), in the eyebrows, around the eyes and extending into the scalp (which shows marked dandruff). It may affect the skin over the sternum, axillae and groins, and the glans penis. A blepharitis may also be present.
- *In elderly people* seborrhoeic eczema can be more severe and progress to involve large areas of the body and even cause erythroderma.

Treatment

This is suppressive rather than curative. A combination of a mild steroid ointment (e.g. 1% hydrocortisone applied twice daily) and a topical antifungal cream (e.g. miconazole cream applied twice daily) will help to control the eruption. Two per cent sulphur and 2% salicylic acid can be added to help control resistant cases. Recent studies have shown 0.1% tacrolimus ointment to be very effective. Ketoconazole shampoo and arachis oil are useful for the scalp. Emollients and a soap substitute are useful adjuncts.

Venous eczema (varicose eczema, gravitational eczema)

This type of eczema occurs on the lower legs due to chronic venous hypertension (usually of more than 2 years' duration) (see p. 1261). The exact cause remains unknown but it has been suggested that venous hypertension causes endothelial hyperplasia and extravasation of red and white blood cells which in turn causes inflammation, purpura and pigmentation.

Clinical features

Venous eczema tends to occur in older people, especially women. It usually appears on the lower legs around the ankles. There may be a past history of venous thrombosis or previous surgery for varicose veins. Brownish pigmentation (haemosiderin) may be seen in the skin and a venous leg ulcer or varicose veins may be present.

Superimposed contact eczema is common in venous eczema patients, especially when there have been chronic venous leg ulcers. This is usually due to an allergic reaction to topical therapies or skin dressings. Patch testing should always be done in treatment-resistant cases.

Treatment

This should include emollients and a moderately potent topical steroid. Support stockings or compression bandages, together with leg elevation, help reverse the underlying venous hypertension (p. 1261).

Asteatotic eczema (winter eczema, eczema craquelé, senile eczema)

This is a dry plate-like cracking of the skin with a red, eczematous component which occurs in elderly people. It occurs predominantly on the lower legs and the backs of the hands, especially in winter. The exact cause is unknown but the repeated use of soaps in the elderly with the loss of the stratum corneum lipids with age is probably involved. Rarely asteatotic eczema can be the presenting sign of hypothyroidism or can follow the commencement of diuretic therapy.

Treatment

Avoidance of soaps, and the regular use of emollients and bath oils should be encouraged. Humidifying centrally heated rooms may help. If the skin is very inflamed, a mild topical steroid can be used.

Contact and irritant eczema

If the eczema is in an unusual or localized distribution (Fig. 23.13), especially if there is no personal or family history of atopic disease, one of a variety of environmental agents (exogenous eczema) is the likely cause. A history of an exacerbation of eczema at the workplace is also suggestive. There are two possible mechanisms: direct irritation or an allergic reaction (type IV delayed hypersensitivity). A detailed history of occupation, hobbies, cosmetic products, clothing and contact with chemicals is necessary.

Allergic contact eczema occurs after repeated exposure to a chemical substance but only in those people who are

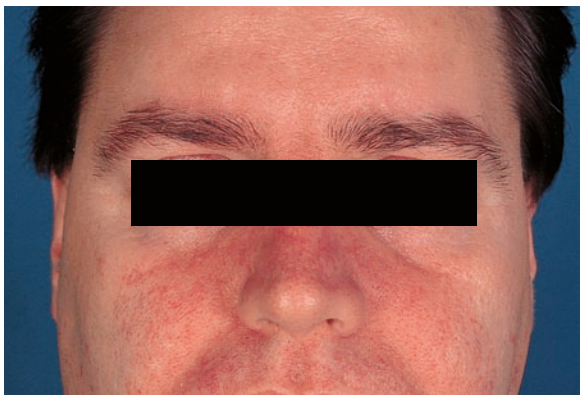


Fig. 23.12 **Seborrhoeic eczema affecting the sides of the nose.**

Fig. 23.13 **Contact eczema secondary to perfume allergy.**

susceptible to develop an allergic reaction. It is common, occurring in up to 4% of some populations. Many substances can cause this type of reaction but the commoner culprits are nickel (in costume jewellery and buckles), chromate (in cement), latex (in surgical gloves), perfume (in cosmetics and air fresheners), and plants (such as primula or compositae). A good history is necessary and if suspicious, patch testing should be arranged to prove any allergy.

Irritant contact eczema can occur in any individual. It often occurs on the hands after repeated exposures to irritants such as detergents, soaps or bleach. It is therefore common in housewives, cleaners, hairdressers, mechanics and nurses.

Treatment

Treatment is as for atopic eczema as well as strict avoidance of any causative agent. This may also involve wearing of protective clothing such as gloves or in extreme cases even changing occupation or hobbies.

Photosensitive eczema

This is discussed on page 1248.

Lichen simplex/nodular prurigo (neurodermatitis)

These terms are applied to a disorder characterized by chronic scratching or rubbing in the absence of an underlying dermatosis. They are more common in Asians and also in black African and Oriental patients.

Lichen simplex appears as thickened, scaly and hyperpigmented areas of lichenification (Fig. 23.14). It starts with intense itching that becomes tender with increased rubbing or scratching, and a self-perpetuating 'itch–scratch cycle' develops. It is rare before adolescence and is commoner in females. Common sites are the nape of the neck, the lateral calves, the upper thighs, the upper back and the scrotum or vulva but any accessible site can be affected.

Nodular prurigo is a different pattern of cutaneous response to scratching, rubbing or picking. Individual, itchy papules and domed nodules appear, especially on the upper trunk and the extensor surfaces of the limbs. They show significant surface damage from scratching. This is a chronic unremitting condition which is often resistant to treatment.

These two conditions overlap, with some patients showing mixed features. Atopic individuals seem predisposed to develop these conditions (in the absence of obviously active eczema). However they can occur in non-atopics. Emotional stress appears to be a contributory factor in many of these patients (hence the name neurodermatitis).

The diagnosis is made by exclusion of other pathologies and may require a skin biopsy. General medical causes of pruritus should be excluded (p. 1253). In the elderly, nodular prurigo may be an early sign of bullous pemphigoid, before the more typical blistering phase has appeared.

Fig. 23.14 **Lichen simplex from chronic rubbing.**

Treatment

Treatment is often difficult as symptoms can be intractable. Very potent topical steroids (e.g. 0.05% clobetasol propionate) with occlusive tar bandaging may sometimes help. Intralesional steroids can also be useful but there is a risk of atrophy. For resistant cases (especially of prurigo lesions) topical doxepin, phototherapy (p. 1248) and even ciclosporin (3–5 mg/kg/day) are used but the risk/benefit ratio must be discussed with the patient as these therapies are potentially toxic.

PSORIASIS

Psoriasis is a common papulo-squamous disorder affecting 2% of the population and is characterized by well-demarcated, red scaly plaques. The skin becomes inflamed and hyperproliferates at about ten times the normal rate. It affects males and females equally and can affect all races. The age of onset occurs in two peaks. Early onset (age 16–22) is commoner and is often associated with a positive family history. Late-onset disease peaks at age 55–60 years.

Aetiology

The condition appears to be polygenic but is also dependent on certain environmental triggers. Twin studies show 73% concordance in monozygotic twins compared with 20% in dizygotic pairs. Nine genetic psoriasis susceptibility loci have been identified (PSORS 1–9). Some loci seem shared with other diseases: atopic eczema (1q21, 3q21, 17q25, 20p), rheumatoid arthritis (3q21, 17q24–25) and Crohn's disease (16). The most studied locus, PSORS1, lies in the MHC region of chromosome 6 (HLA Cw6).

Infection (group A *Streptococcus*), drugs (e.g. lithium), ultraviolet light, alcohol abuse and possibly stress may be triggers or exacerbating factors in certain individuals. The exact aetiology is unknown but evidence suggests that psoriasis is a T-lymphocyte-driven disorder to an unidentified antigen(s). The initial T cell activation requires a complex interaction between T lymphocyte and antigen-presenting cell (dendritic/Langerhans) and this has provided a number of potential 'targets' for the newly developed biological therapies (Fig 23.15). This activation results in upregulation of Th1-type T cell cytokines, e.g. interferon-γ, interleukins (IL-1, -2, -8), growth factors (TGF-α and TNF-α) and adhesion molecules (ICAM-1). The pro-inflammatory cytokine TNF-α is also produced by keratinocytes and this may be involved in both initiation and maintenance of psoriatic lesions. TNF-α blockade seems to be the most promising of the new biological agents (Fig. 23.15).

Pathology

Skin biopsy shows epidermal acanthosis and parakeratosis, reflecting the increase in skin turnover, and the granular layer is often absent. Polymorphonuclear abscesses may be seen

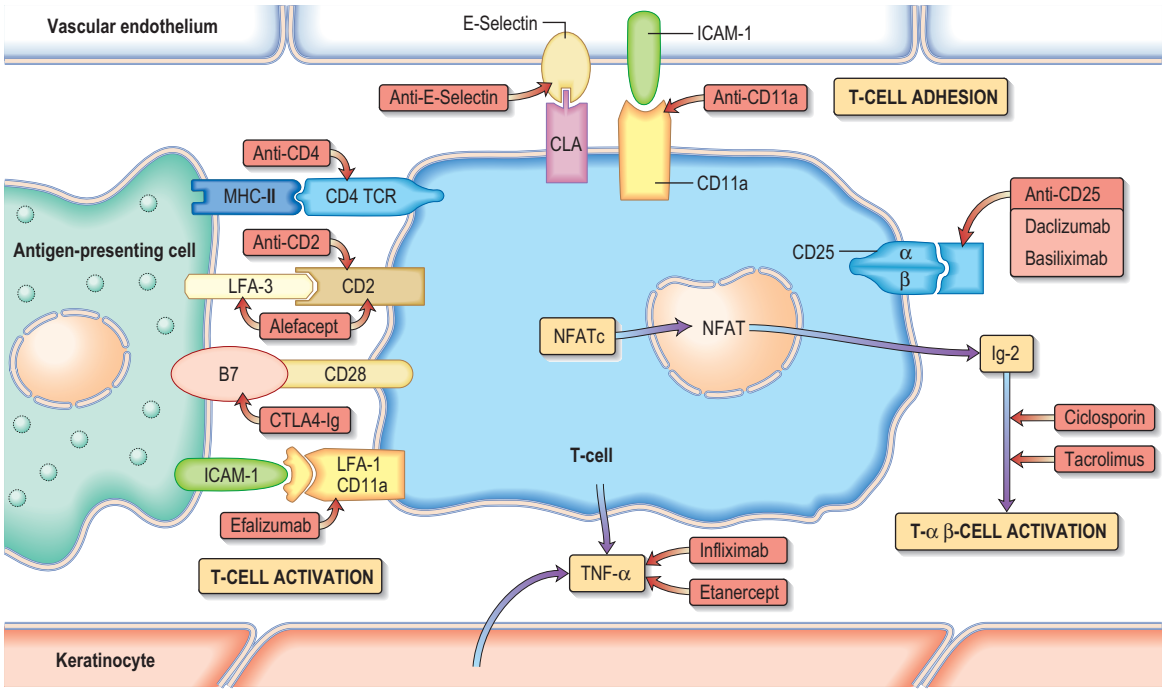


Fig. 23.15 **Psoriasis – pathogenesis.** This schematic figure shows two processes – how T cells/APCs react with keratinocytes in the epidermis and how T cells/APCs react with blood vessels in the dermis. Novel T-cell-targeted therapies are shown (red). NFAT, nuclear factor of activated T cells; ICAM, intracellular adhesion molecule; LFA, lymphocyte function-associated antigen; MHC, major histocompatibility complex; TNF, tumour necrosis factor; IL, interleukin; CLA, cutaneous lymphocyte antigen; B7, CD80.

in the upper epidermis. The epidermal rete ridges appear elongated and clubbed as they fold down into the dermis. Dermal changes include capillary dilation surrounded by a mixed neutrophilic and lymphohistiocytic perivascular infiltrate.

Clinical features

Psoriasis can present in different clinical patterns but there is overlap between the different forms. Certain drugs can make psoriasis worse, notably lithium, antimalarials and rarely beta-blockers.

Chronic plaque psoriasis

This is the 'common' type of psoriasis. It is characterized by pinkish red scaly plaques, with a silver scale seen, especially on extensor surfaces such as knees (Fig. 23.16a) and elbows. The lower back, ears and scalp are also commonly involved. New plaques of psoriasis occur at sites of skin trauma – the so-called 'Köbner phenomenon'. The lesions can become itchy or sore.

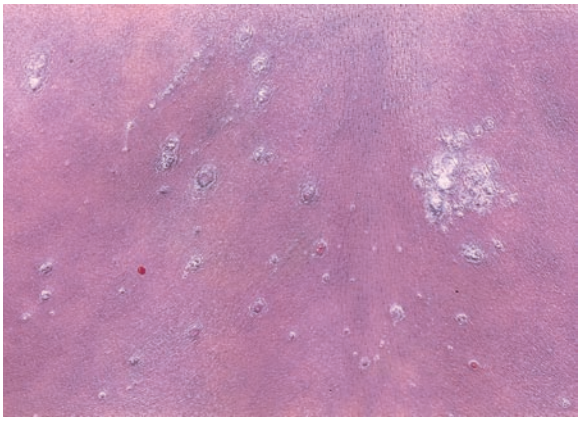
Flexural psoriasis

This tends to occur in later life. It is characterized by well-demarcated, red glazed plaques confined to flexures such as the groin, natal cleft and sub-mammary area. As these sites are apposed there is rarely any scaling. In the absence of psoriasis elsewhere the rash is often misdiagnosed as candida intertrigo but the latter will normally show satellite lesions.

Guttate psoriasis

'Raindrop-like' psoriasis is a variant most commonly seen in children and young adults (Fig. 23.16b). An explosive eruption of very small circular or oval plaques appears over the trunk about 2 weeks after a streptococcal sore throat.

(a)



(b)

Fig. 23.16 **(a) Psoriasis of the knees. (b) Guttate psoriasis in an African** (courtesy of Dr P Matondo, Lusaka, Zambia).

Erythrodermic and pustular psoriasis

These are the most severe types of psoriasis reflecting a widespread intense inflammation of the skin. They can occur together ('Von Zumbusch' psoriasis) and may be associated with malaise, pyrexia and circulatory disturbance. This form can be life-threatening. The pustules are not infected but are sterile collections of inflammatory cells. There is also a more localized variant of pustular psoriasis that confines itself to the hands and feet (palmo-plantar psoriasis) but is not associated with severe systemic symptoms. This latter type is more common in heavy cigarette smokers.

Associated features

Nails. Up to 50% of individuals with psoriasis develop nail changes (Fig. 23.17) and rarely these can precede the onset of skin disease. There are five types of nail change: (a) pitting of the nail plate; (b) distal separation of the nail plate (onycholysis); (c) yellow-brown discoloration; (d) subungual hyperkeratosis; (e) rarely a damaged nail matrix and lost nail plate. Treatment of nail dystrophy is very difficult.

Arthritis. Five to ten per cent of individuals develop psoriatic arthritis and most of these will have nail changes (p. 1266).

Treatment

This is concerned with control rather than cure. It should be tailored to the patient's wishes and not just to the doctor's assessment of disease severity. Severe cases may require hospitalization.

Chronic plaque psoriasis: emollients should always be used to hydrate the skin. Mild to moderate topical steroids, synthetic vitamin D_3 analogues (e.g. calcipotriol, calcitriol, tacalcitol), 0.05% tazarotene (a retinoid) and purified coal tar are the most popular therapies. Salicylic acid can be a useful adjunct. All should be applied once to twice daily to palpable lesions. Once lesions have flattened, therapy can be discontinued. Tazarotene and calcipotriol can be very irritant (calcitriol somewhat less so) so they are often used in combination with steroid creams. Vitamin D analogues should be used with caution in extensive psoriasis because there is a risk of hypercalcaemia if greater than 100 g is used per week.

Dithranol causes staining of the skin and clothing and it may prove difficult to use at home on a regular basis. It is normally applied for 20–60 minutes and then washed off. It must be applied carefully to the lesions as it causes irritation to normal skin. Dithranol is more likely to induce remission than other therapies but is being used less because of poor compliance.

Topical therapies are sometimes used in combination with UVB or PUVA. The 'Goeckerman regimen' consists of tar and UVB; the 'Ingram regimen' consists of dithranol and UVB. The latter has similar results to oral PUVA in terms of clearance rates and lengths of remission – approximately 75% in 6 weeks.

Flexural psoriasis is usually treated with mild steroid and/or tar topical creams. Calcitriol and 0.1% tacrolimus ointment are also useful for treating flexural (facial and genital) psoriasis where irritation can be a problem.

Guttate psoriasis is usually treated with topical therapies and/or UVB phototherapy.

Palmo-plantar psoriasis is treated with very potent topical steroids, coal tar paste or local hand and foot PUVA.

Systemic therapy. Agents such as methotrexate, acitretin, mycophenolate, ciclosporin or hydroxycarbamide (hydroxyurea) are used for resistant cases.

Erythrodermic psoriasis also requires systemic therapy (but not phototherapy) as well as general supportive measures (p. 1249).

All systemic treatments must be monitored for toxicity.

Use of methotrexate. Methotrexate is normally given once weekly. Some patients experience severe nausea on the day they take it which can be lessened by folic acid therapy. Both men and women should avoid conception during and for 3 months after therapy. Some patients are allergic to methotrexate and develop a pyrexia and mouth ulceration. Regular blood tests need to be done to monitor for bone marrow suppression and liver damage. Alcohol must be avoided as this increases the risk of hepatotoxicity. NSAIDs should also be avoided as these inhibit excretion. Lower doses should be used in the elderly. Long-term use causes hepatic fibrosis, and regular monitoring of patients' serum procollagen III peptide level or elastography (p. 326) is being used to assess fibrosis development. Patients with concomitant psoriatic arthritis are more likely to develop pulmonary fibrosis.

Biological agents (see Fig. 23.15)

- Etanercept (a TNF-α blocker) and efalizumab (an anti-CD11a monoclonal antibody) have recently been introduced in Europe. They are at present only used for patients with severe disease and those who have failed (or cannot tolerate due to toxicity) conventional systemic treatments. Two other TNF-α blockers, infliximab and adalimumab, have been approved.
- Alefacept (a dimeric fusion protein) has been licensed in some countries.
- Human interleukin-12/23 monoclonal antibody is being used in trials but may be limited by side-effects.

All these agents are given by injection and are very expensive.

The *TNF-α blockers* seem to be the most effective with almost 60–80% of patients having at least a 75% improvement within 12 weeks; efalizumab showed only a 50% similar improvement at 12 weeks but it had less serious side-effects.

Fig. 23.17 **Psoriasis of the nail** – yellowish brown discoloration and distal nail plate separation (onycholysis).

Long-term side-effects of these new biological agents are unknown. TNF-α blockers' side-effects are discussed on page 464.

Prognosis

Most individuals who develop chronic plaque psoriasis will have the condition lifelong but 80% will get a remission at times. It fluctuates in severity and there are no available tests to predict outcome. Guttate psoriasis resolves spontaneously and in up to a third of individuals does not recur. However, two-thirds will go on to get recurrent guttate attacks or will progress to chronic plaque psoriasis.

URTICARIA/ANGIO-OEDEMA

Urticaria (hives, 'nettle rash') is a common skin condition characterized by the acute development of itchy weals or swellings in the skin due to leaky dermal vessels (Fig. 23.18). Urticaria is described as 'acute' if it lasts less than 6 weeks and 'chronic' if it persists beyond this. Angio-oedema is a similar condition but involves sub-dermal vessels.

Aetiology

The final event in pathogenesis involves degranulation of cutaneous mast cells which releases a number of inflammatory mediators (including histamine) which in turn make the dermal and sub-dermal capillaries leaky. In most cases the underlying cause is unknown. Occasionally these conditions are secondary to viral or parasitic infection, drug reactions (e.g. NSAIDs, penicillin, ACE inhibitors, opiates), food allergy (e.g. to strawberries, food colourings or seafood), or rarely systemic lupus erythematosus. There is evidence for an autoimmune aetiology in some of the 'idiopathic' cases as certain individuals develop autoantibodies against the high-affinity IgE receptor α-subunit of the mast cell. Urticaria/angio-oedema is commoner in atopic individuals and usually presents in children and young adults.

Clinical features

Urticaria produces cutaneous swellings or weals developing acutely over a few minutes. They can occur anywhere on the skin and last between minutes and hours before resolving spontaneously. Lesions are intensely itchy and show no surface change or scaling. They are normally erythematous but if very acutely swollen, they may appear flesh-coloured or whitish and people often mistake them for blisters.

Angio-oedema with subcutaneous involvement presents as soft tissue swelling (oedema) especially around the eyes, the lips and the hands but this is rarely itchy. This can be very alarming to the patient. It can also be dangerous if mucosal areas such as the mouth and larynx are involved but fortunately this is very rare.

Fig. 23.18 **Urticaria.**

Physical urticarias

Occasionally urticaria can be caused by physical stimuli such as cold (cold urticaria), deep pressure (delayed pressure urticaria), stress or heat (cholinergic urticaria), sunlight (solar urticaria – p. 1249), water (aquagenic urticaria) or chemicals such as latex (contact urticaria).

Cholinergic urticaria is one of the commonest physical urticarias and has rather different clinical lesions from the other forms. Small itchy papules rather than weals appear on the upper trunk and arms after exercise or anxiety.

Pressure can cause two types of urticaria. More superficial pressure can cause *dermographism* which is relatively common. This presents as urticated weals occurring a few minutes after application of light pressure. Even scratching or rubbing will bring up linear weals in dermographic individuals. *Delayed pressure urticaria* is rare and occurs as deep swellings some hours after pressure is removed (e.g. on the soles of the feet or under a tight belt).

Investigations

The history is the most useful factor in diagnosing urticaria. Routine investigation is probably not justified unless the history suggests one of the underlying causes listed above. The physical urticarias should be reproducible by applying the relevant stimulus.

Treatment

Any identifiable underlying cause should be treated appropriately. Patients should avoid salicylates and opiates as they can degranulate mast cells. Oral antihistamines (H_1 blockers) are the most useful in treating the idiopathic cases. Therapy should be started with regular use of a non-sedating antihistamine (e.g. cetirizine 10 mg daily or loratadine 10 mg daily). If control proves difficult, addition of a sedating antihistamine, an H_2 blocker or dapsone may be helpful. Dietary manipulation (e.g. additive and colouring free diets) may help a small proportion of patients with chronic urticaria but it is generally unrewarding. Angio-oedema of the mouth and throat may require urgent treatment with intramuscular epinephrine (adrenaline) and intravenous steroids (see Emergency box 3.1).

Prognosis

Most cases of 'idiopathic' urticaria last a few weeks to months before disappearing spontaneously. The majority of these will be controlled with an antihistamine. A small percentage of people go on to develop chronic urticaria which can last for several months or years. The physical urticarias (especially cholinergic urticaria) are more persistent, often lasting for years, and they are often resistant to therapy.

Urticarial vasculitis

This is a variant of urticaria and should be suspected if individual urticarial lesions last longer than 24 hours and leave bruising behind after resolution. There is often an associated arthralgia or myalgia, and a small proportion may go on to develop a connective tissue disease. The diagnosis is confirmed by skin biopsy. A full vasculitis screen should be carried out for an underlying cause (p. 548).

Treatment is with antihistamines, oral dapsone (50–100 mg daily) or immunosuppressants.

Hereditary angio-oedema

This is an extremely rare autosomal dominant condition due to an inherited deficiency of C1 esterase inhibitor (CI-INH), a component of the complement system. The defect may be due to either reduced function or reduced levels (<50% of normal levels in 85% of cases). Serum C2 and C4 levels are normally low but C1 and C3 are normal. Rarely this condition is acquired and associated with lymphoma or SLE. These types show low C1 esterase inhibitor levels but also low C1 levels.

Clinical features

Hereditary angio-oedema presents with attacks of non-itchy cutaneous angio-oedema (but no urticaria), which may last up to 72 hours. It may also present with recurrent abdominal pain (due to intestinal oedema) and there may be a family history of sudden death (due to laryngeal involvement). A non-specific erythematous rash may precede an attack of angio-oedema but urticaria is not a feature.

Treatment

In the acute setting treatment is with C1 esterase inhibitor concentrates and fresh frozen plasma. Adrenaline and steroids are often ineffective. Maintenance treatment with the anabolic steroid stanozolol (or danazol) stimulates an increase in hepatic synthesis of C1 esterase inhibitor but this should not be used in children. Family members should be screened.

PITYRIASIS ROSEA

Pityriasis rosea is a self-limiting rash seen in adolescents and young adults. The cause is unknown but it is thought to be a viral or post-viral rash. There is an increased incidence in spring and autumn and outbreaks may occur in institutions.

Clinical features

The rash consists of circular or oval pink macules with a collarette of scale and is more prominent on the trunk than the limbs. The long axis of the oval lesions tends to run along dermatomal lines giving a 'Christmas tree' pattern on the back. The rash may be preceded by a large solitary patch with peripheral scaling ('herald patch') and this is most commonly found on the trunk. The rash is usually asymptomatic and spontaneously resolves over 4–8 weeks.

Treatment

Treatment is not normally required but 1% menthol in aqueous cream may help relieve any itch. In persistent cases UVB (p. 1248) may be helpful.

LICHEN PLANUS

Lichen planus is a pruritic inflammatory dermatosis that is commonly associated with mucosal involvement and rarely with nail dystrophy and scarring alopecia.

The cause is unknown but is possibly a T-cell-driven immune mechanism as an almost identical rash can be caused by certain drugs (e.g. beta-blockers, gold, levamisole, ACE inhibitors or antimalarials) or by graft-versus-host disease and in chronic HBV, HCV liver disease.

Pathology

Hyperkeratosis with thickening of the granular cell layer is seen in the epidermis. A dense T cell infiltrate is seen at the dermo-epidermal junction, which becomes ragged and saw-toothed. The basal layer shows liquefactive degeneration with colloid (apoptotic) bodies in the upper dermis.

Clinical features

The rash is characterized by small, purple flat-topped, polygonal papules that are intensely pruritic (Fig. 23.19). It is common on the flexors of the wrists and the lower legs but can occur anywhere. There may be a fine lacy white pattern on the surface of lesions (Wickham's striae). Lesions can fuse into plaques, especially on the lower legs and in black Africans. Hyperpigmentation is common after resolution of lesions, especially in patients with pigmented skin. Atrophic, hypertrophic and annular variants can occur. Lichen planus lesions often localize to scratch marks. If lesions occur in the scalp they may cause a scarring alopecia.

Mucosal involvement is common. The mouth is the most commonly affected but the anogenital region can be involved. It can present as lacy white streaks, white plaques or as ulceration. The prominent mucosal symptom is of severe pain rather than itch. Nails may be dystrophic and can be lost altogether (with scarring and 'wing' formation) in severe disease.

Prognosis

The condition often clears by 18 months but can recur at intervals. The hypertrophic and atrophic variants and mucosal disease are more persistent, lasting for years. Ulcerative mucosal disease is pre-malignant.

Fig. 23.19 Lichen planus.

Treatment

This requires the use of potent topical steroids (0.05% clobetasol propionate) and occasionally oral prednisolone (30 mg daily for 2–4 weeks). Occlusion of topical treatments can be helpful. Mouth lesions also require high potency steroids as an ointment, gel or sprays; candidiasis is a complication. Resistant cases may respond to PUVA, oral retinoids (0.5 mg/kg/day) or azathioprine (1–2 mg/kg daily). Topical 0.1% tacrolimus ointment or pimecrolimus has proved a very useful therapy for painful mucosal disease.

GRANULOMA ANNULARE

Granuloma annulare is a dermatosis predominantly of children and young adults. It is characterized by clusters of small dermal papules (with no surface change) that often form into rings or part of a ring. They are common on the dorsal surface of the hands and feet. They are flesh coloured or slightly erythematous and are usually asymptomatic. As they heal, the centre becomes dusky and altered in texture. A deep form, which is tender, exists in children. Diffuse granuloma annulare may be associated with diabetes mellitus. The pathology shows a granulomatous dermal infiltrate with foci of degeneration of collagen (necrobiosis). Spontaneous resolution often occurs after months to years but cryotherapy or triamcinolone injection may help localized disease.

LICHEN SCLEROSUS

Lichen sclerosus is a common inflammatory dermatosis that occurs in all age groups and particularly affects the anogenital region. It is more common in post-menopausal females. It presents with atrophic ivory-white patches with a well-defined edge on the vulva, glans penis, foreskin or perianal skin. Telangiectasia may be seen over the surface. Occasionally lesions involve the shaft of the penis and the urethral meatus. Lesions are often itchy but may be sore at times. Long-standing vulval lesions may be associated with fissuring and a marked loss of architecture, especially of the clitoral hood and the labia minora, which may become fused. Early lesions in young girls may present as haemorrhagic blisters and these are occasionally mistaken as signs of sexual abuse. Involvement of the foreskin can cause phimosis, and urethral disease may interfere with micturition. Perianal lesions may fissure and cause constipation.

Rarely lichen sclerosus can affect non-genital skin. This is most common in females and clinically it may show rather more hyperkeratosis and follicular plugging than is seen in the anogenital region.

Diagnosis requires biopsy if the clinical picture is unclear. This shows thinning of the epidermis with acanthosis and elongation of the rete ridges. Vulval scarring can also occur with cicatricial pemphigoid (p. 1256) so occasionally immunofluorescence studies are needed.

The underlying cause is unknown but HLA associations and recent studies showing antibodies to extracellular matrix protein-1 suggest an autoimmune aetiology.

Treatment with very potent topical steroids helps control the symptoms. Hydroxychloroquine (200 mg twice daily) is used in resistant cases. The condition may burn itself out after many years, especially in children. There is a risk of developing squamous cell carcinoma in long-standing lesions. Male patients may require circumcision if phimosis does not respond to medical therapy.

FURTHER READING

Bieber T. Mechanisms of disease: atopic dermatitis. *New England Journal of Medicine* 2008; **358**: 1483–1494.

Dahl MV. Granuloma annulare: long-term follow-up. *Archives of Dermatology* 2007; **143**: 946–947.

Krueger GG, Langley RG, Leonardi C et al. A human interleukin-12/23 monoclonal antibody for the treatment of psoriasis. *New England Journal of Medicine* 2007; **356**: 580–592.

Seneviratne SL, Black AP, Jones L et al. The role of skin-homing T cells in extrinsic atopic dermatitis. *Quarterly Journal of Medicine* 2007; **100**: 19–27.

Smith FJ, Irvine AD, Terron-Kwiatkowski A et al. Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nature Genetics* 2006; **38**: 337–342.

Zuraw BL. Hereditary angioedema. *New England Journal of Medicine* 2008; **359**; 1027–1036.

NICE UK treatment guidelines:

http://guidance.nice.org.uk/TA81 Atopic dermatitis (topical steroids)

http://guidance.nice.org.uk/TA82 Tacrolimus and pimecrolimus

http://guidance.nice.org.uk/TA103 Psoriasis (efalizumab and etanercept)

FACIAL RASHES

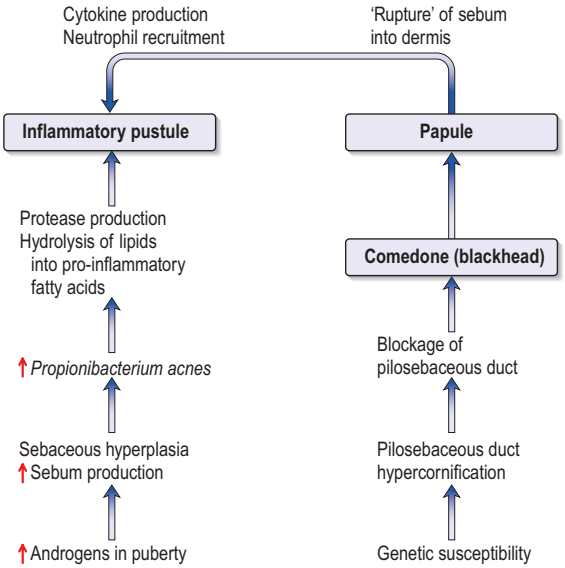
Facial rashes often cause diagnostic confusion but a close examination of the clinical signs should help differentiate the underlying cause (Table 23.5). All facial rashes, by virtue of their visibility, can cause significant distress to the patient and this should never be underestimated.

Acne vulgaris

Acne is a very common facial rash occurring in over 85% of adolescents and frequently continuing into early and mid-adult life. Occasionally it can cause profound psychological

Table 23.5 Causes of facial rashes

Acne vulgaris
Rosacea
Seborrhoeic eczema
Atopic eczema
Contact eczema
Dermatomyositis
Perioral dermatitis
Photosensitivity
Sarcoidosis
Chronic discoid lupus erythematosus
Systemic lupus erythematosus
Subacute lupus erythematosus



(a)

(b)

Fig. 23.20 **(a) Pathophysiology of acne vulgaris. (b) Skin lesion.**

disturbance and depression, even suicide. The cause is multifactorial but follicular epidermal hyperproliferation, blockage of pilosebaceous units with surrounding inflammation, increased sebum production are critical factors in the pathological process (Fig. 23.20a) as is infection with *Propionibacterium acnes*.

Propionibacterium acnes induced inflammation has recently been shown to occur via activation of Toll-like receptor 2, which leads to production of pro-inflammatory cytokines, e.g. IL-8, IL-12 and TNF-α.

Clinical features

Acne presents in areas rich in sebaceous glands such as the face, back and sternal area. The three cardinal features are:

- open comedones (blackheads) or closed comedones (whiteheads)
- inflammatory papules
- pustules (Fig. 23.20b).

The skin may be very greasy (seborrhoea). Rupture of the inflamed lesions may lead to deep-seated dermal inflammation and nodulocystic lesions, which are more likely to cause facial scarring. A premenstrual exacerbation of acne is sometimes noticed. There is a tendency for spontaneous improvement over a number of years but acne can persist unabated into adult life.

A number of clinical variants exist:

- *Infantile acne*. Facial acne is occasionally seen in infants and is sometimes cystic. It is thought to be due to the influence of maternal androgens and resolves spontaneously.
- *Steroid acne*. Acne may occur secondary to corticosteroid therapy or Cushing's syndrome. Comedones and cysts are rare in this variant but involvement of the back and shoulders (rather than the face) is common. Clinically the rash often appears as a pustular folliculitis.
- *Oil acne.* This is an industrial disease seen in workers who have prolonged contact with oils or other hydrocarbons and is common on the legs and other exposed sites.
- *Acne fulminans.* This is a rare variant seen most commonly in young male adolescents. Severe necrotic and crusted acne lesions appear associated with malaise, pyrexia, arthralgia and bone pain (due to sterile bone cysts). It requires urgent treatment with oral prednisolone (30–40 mg daily) and analgesics followed by a course of oral isotretinoin (see below).
- *'Follicular occlusion triad'.* This is a rare disorder most commonly seen in black Africans. It is characterized by the presence of severe nodulocystic acne, dissecting cellulitis of the scalp (p. 1267) and hidradenitis suppurativa (p. 1230). This may be caused by a problem of follicular occlusion rather than having an infective aetiology.

Treatment

This is aimed at decreasing sebum production, decreasing bacteria, normalizing duct keratinization and decreasing inflammation.

Regular washing with acne soaps to remove excess grease is helpful (normal soaps can be comedogenic). 'Picking' should be discouraged.

First-line therapy

Mild acne can respond to a variety of topical agents, e.g. keratolytics (benzoyl peroxide, azelaic acid) or topical retinoids (tretinoin or isotretinoin) or retinoid-like agents (adapalene). Topical antibiotics, e.g. erythromycin or clindamycin, are used for inflammatory acne. All topical agents can cause problems with irritation.

Second-line therapy

- *Low-dose oral antibiotic therapy* often helps but must be given for at least 3–4 months. Oxytetracycline 500 mg twice daily is often used first. Minocycline 100 mg daily, erythromycin 500 mg twice daily or trimethoprim 100 mg twice daily are also used.
- An extra treatment, cyproterone acetate 2 mg/ethinylestradiol 35 μg (co-cyprindol), is of value in females if there is no contraindication to oral contraception. This acts as a normal combined contraceptive but has antiandrogen activity. It may take 6–8 months to have its maximum effect.
- *UVB phototherapy* can be helpful but is rarely used now due to the development of retinoid drugs.

Third-line therapy

Third-line treatment with a retinoid drug (isotretinoin) should be given if:

www.studentconsult.com
Ask the authors

- the above measures fail
- there is nodulocystic acne with scarring
- there is severe psychological disturbance.

Use of retinoids (isotretinoin or acitretin). Retinoids are synthetic vitamin A analogues that affect cell growth and differentiation. They are very teratogenic. Isotretinoin is a 'hospital-only drug' in most countries due to its teratogenicity and is restricted to the use of dermatologists and a few trained family doctors. A pregnancy test, contraceptive advice and signed consent are mandatory prior to its use in fertile women, and pregnancy testing must be repeated monthly during therapy. It is given as a 4-month course at a dose of 0.5–1 mg/kg/day. Over 90% of individuals will respond to this therapy and 65% of people will obtain a long-term 'cure'.

Patients must avoid pregnancy during therapy and for 1 month after stopping isotretinoin (but for 2 years after stopping acitretin as it is very lipophilic). Both drugs cause drying of the skin, especially of the lips and nasal mucosa. Hair thinning and exercise-induced myalgia are not uncommon. Blood count, liver biochemistry and fasting lipids need to be monitored during therapy. In a few individuals retinoids may cause depression but it should also be remembered that acne itself has been a cause of suicide.

A number of physical techniques are currently under assessment (e.g. lasers, blue light, microdermabrasion) but they are not as effective as isotretinoin and can be very expensive as repeat treatments are often needed.

Rosacea

Rosacea (Fig. 23.21) is a common inflammatory rash predominantly affecting the face. The onset is usually in middle age and it is commoner in women. It often causes significant psychological distress.

The cause is unknown. Theories have suggested an underlying problem in vasomotor stability of blood vessels or a role of the skin mite *Demodex*.

Clinical features

The cardinal features are of facial flushing and inflammatory papules and pustules affecting the nose, forehead and cheeks. The flushing may precede the other signs by some years. There are no comedones. Additional features include dilated blood vessels (telangiectasia), inflammation of the eyelid margins (blepharitis), keratitis and sebaceous gland hypertrophy, especially of the nose. The latter is commoner in men and can cause a disfiguring enlargement of the nose called rhinophyma. The flushing may be exacerbated by alcohol, hot drinks, sunlight and changes in ambient temperature. Prolonged use of topical steroids can exacerbate or trigger the condition. As the disease progresses the flushing may be replaced by a permanent erythema.

Fig. 23.21 **Rosacea.** Papules and pustules on a background erythema. There are no comedones.

Treatment

This is suppressive rather than curative. Long-term use of topical 0.075% metronidazole or topical 15% azelaic acid may help. Avoid topical steroids. A 3-month course of oral tetracycline (500 mg twice daily) is also helpful. Oral metronidazole (400 mg twice daily) or oral isotretinoin (0.5 mg/kg/day) is occasionally given in resistant cases (p. 1247). The papules and pustules tend to respond best to therapy but repeat courses may be necessary. The flushing and erythema are often resistant to treatment but cosmetic camouflage can be helpful for these features. Intense pulse light or pulsed dye laser therapy can help the erythema and telangiectasia but often needs to be repeated as rosacea tends to recur. Rhinophyma can be treated with plastic surgery or by carbon dioxide laser.

Perioral dermatitis

Perioral dermatitis is a common rash found around the mouth, especially in young females. The exact cause is unknown but it often has an iatrogenic component as topical steroids often exacerbate the condition in the long term.

Clinical features

It presents with erythema, scaling, papules and occasionally pustules around the mouth. It usually spares a halo of skin immediately adjacent to the lips. Rarely there is involvement around the eyes.

Treatment

Treatment involves stopping topical steroids although they may have to be withdrawn slowly to prevent too severe a rebound after withdrawal. The mainstay of treatment is with a 3–4-month course of low-dose oxytetracycline or erythromycin (both 500 mg twice daily) and topical metronidazole.

Blushing

Blushing is characterized by facial flushing in response to emotional stimuli and would be regarded as a normal physiological response. In some individuals it is very frequent and can become a debilitating disorder interfering with work and social interaction. The cause may be both psychological and physiological. Non-emotional causes of flushing should be excluded, e.g. post-menopause, drugs, alcohol, diet, carcinoid syndrome, rosacea and mastocytosis.

Treatment includes cognitive behavioural therapy, cosmetic camouflage, beta-blockers and clonidine. Selective serotonin reuptake inhibitors may help associated depression and anxiety. Botulinum toxin and surgical sympathectomy have been advocated by some but neither is universally accepted and the latter can have major side-effects such as compensatory hyperhidrosis in over 50% of cases.

FURTHER READING

Crawford GH, Pelle MT, James WD. Rosacea I: Etiology, pathogenesis, and subtype classification. *Journal of the American Academy of Dermatology.* 2004; **51**: 327–341.

Heymann WR. Toll-like receptors in acne vulgaris. *Journal of the American Academy of Dermatology.* 2006; **55**: 691–692.

Nicolau M, Paes T, Wakelin S. Blushing: an embarrassing condition, but treatable. *Lancet.* 2006; **367**: 1297–1299 (see also page 2059).

Pelle MT, Crawford GH, James WD. Rosacea II: Therapy. *Journal of the American Academy of Dermatology.* 2004; **51**: 499–512.

Purdy S, de Berker D. Acne. *British Medical Journal.* 2006; **333**: 949–953.

PHOTODERMATOLOGY

Sunlight – light in the ultraviolet (UV) part of the spectrum – combines short, medium and long wavelengths (UVC, UVB and UVA respectively). Both UVB and UVA can penetrate the atmosphere and reach the skin. This light energy is potentially mutagenic and carcinogenic but it can also suppress cutaneous inflammation. Thus UV irradiation can both cause skin disease and be used to treat it.

Photosensitive rashes usually appear on sites exposed to the sun's rays, such as the face, the anterior 'V' of the chest, the ears and the backs of the hands. Certain 'protected' areas are characteristically spared such as under the chin or the upper eyelid and between the finger webs. Porphyria, drug sensitivity and lupus erythematosus should be excluded in all photosensitive patients.

Photosensitive rashes may be divided into photoexacerbated/provoked rashes and the idiopathic photodermatoses (Table 23.6). The former are discussed on pages 542 (SLE), 222 (pellagra) and 1075 (porphyria).

Table 23.6 Differential diagnosis of photosensitive rashes

Photoexacerbated/ provoked rashes	
Systemic disease	SLE, CDLE, SCLE (p. 541)
Metabolic disease	Porphyrias (pp. 1073 and 1254), pellagra (p. 222)
Drugs	Thiazides, phenothiazines, tetracyclines, amiodarone
Plant phototoxins	Phytophotodermatitis (photosensitivity induced by contact of the skin with certain plants, e.g. celery, hogweed, rue, line, fig tree)
Skin disease	Rosacea
	Rarely atopic eczema, psoriasis, lichen planus (these normally improve in sunlight)
Idiopathic photodermatoses	
Polymorphic light eruption	
Chronic actinic dermatitis	
Solar urticaria	

CDLE, chronic discoid lupus erythematosus; SCLE, subacute cutaneous lupus erythematosus; SLE, systemic lupus erythematosus

Phototherapy and photoprotection

Phototherapy

UVB and UVA both have a suppressive effect on cutaneous inflammation and there is increasing evidence that they can suppress systemic immunoreactivity to some degree. However, both types can cause skin ageing and predispose to skin malignancy if excessive doses are used. This is more of a problem in white-skinned individuals. Unaffected regions of skin or high-risk areas like the scrotum can be screened during phototherapy.

UVB is less carcinogenic than UVA. Narrow-band UVB (311 nm) is used more than broad-band UVB because it is more effective and less likely to burn. It is used in the treatment of eczema and psoriasis (especially in children) and is usually given three times per week for 6–10 weeks. Eye protection needs to be worn during therapy.

UVA is relatively ineffective on its own so is used in conjunction with a photosensitizer ('psoralen'), hence the term 'PUVA'. The psoralen can be given by mouth or applied to the skin in bath water. PUVA is given twice a week and eye protection must been worn for the whole of the treatment day as the psoralen sensitizes the retina. It is more effective than broad-band UVB but is limited by its carcinogenic potential (especially induction of squamous cell carcinoma). A maximum dose is given over a lifetime depending on skin type (1000 joules or 200 sessions approximately). It is used for many conditions including psoriasis, eczema, cutaneous T-cell lymphoma, some photosensitive dermatoses and vitiligo.

Sunbeds are used for tanning and consist of predominantly UVA light and are therefore rarely effective in treating skin disease. If used frequently there is an increased risk of skin cancer and premature ageing.

Photoprotection

There are two broad classes of sunblock cream: they either absorb UV light (e.g. aminobenzoic acid or methoxycinnamate) or reflect it (e.g. titanium dioxide). Most modern creams protect against UVA and UVB to varying degrees. UVB protection is graded by the 'sun protection factor' (SPF): an SPF of 15 implies you can spend 15 times as long in the sun before burning providing it is applied correctly. SPFs above 15 confer little extra protection. There is no standardized way of assessing efficiency against UVA. Some sunscreens (especially aminobenzoates) may rarely cause photosensitization. This can be proven by photopatch testing.

IDIOPATHIC PHOTODERMATOSES

Polymorphic light eruption (PLE)

This is the most common photosensitive eruption in temperate regions, affecting up to 10–20% of the population. It is most common in young women. In many it is mild and often goes undiagnosed. An itchy rash appears some hours after sun exposure which is strictly confined to the exposed sites. Lesions may be papules, vesicles or plaques. They can last for several hours or several days. The condition starts in spring and often improves during the summer because of skin 'hardening'.

Treatment

Avoidance of sunlight and the use of sunblocks are helpful in mild cases. Topical steroids help treat an attack. In those individuals who only get PLE after very intense sun exposure (e.g. on sunny holidays) a short course of oral prednisolone (30 mg daily for 7–10 days) will often prevent or treat an attack. For resistant cases 'desensitization' with low-dose PUVA (or narrow-band UVB) in the springtime may be required but patients will need to 'top up' their sun exposure from natural sunlight during the summer to keep their skin desensitized.

Chronic actinic dermatitis (photosensitive eczema, actinic reticuloid)

This is a relatively rare type of eczema occurring in a photosensitive distribution over the face, neck and hands. It typically affects middle-aged or elderly males. There may be a pre-existing eczema so the subsequent development of photosensitivity is often missed. This is further confounded by the fact that the eczema usually will spread to affect skin not exposed to sunlight, and the patient can become erythrodermic. The skin has typical features of eczema but there is often marked skin thickening. Histology is often atypical and can look almost lymphoma-like. The diagnosis can be confirmed by specialist monochromator light-testing. The most severe cases can be exacerbated by even artificial lighting as these patients can become exquisitely photosensitive.

Treatment

This consists of strict avoidance of sunlight including high-factor sunblocks and screening of house and car windows. Topical steroids and emollients are useful in milder cases. Oral prednisolone may be needed and azathioprine (1–2 mg/kg daily) is used for long-term suppression. Low-dose PUVA under steroid cover may help with 'desensitization'.

Solar urticaria

This is extremely rare. Itchy urticarial lesions occur within minutes of sun exposure and characteristically settle within 1–2 hours. Sun avoidance, sunblocks, H_1 antihistamines and low-dose PUVA are all used in treatment.

ERYTHRODERMA

Erythroderma, meaning 'red skin', refers to the clinical state of inflammation or redness of all (or nearly all) of the skin. It is sometimes called exfoliative dermatitis, but dermatitis is not always present.

Aetiology

There are a number of underlying causes (Table 23.7); previous skin disease and drugs are the most common.

Table 23.7 Causes of erythroderma

Common
Atopic eczema
Psoriasis
Drugs (e.g. sulphonamides, gold, sulfonylureas, penicillin, allopurinol, captopril)
Seborrhoeic eczema
Idiopathic
Rare
Chronic actinic dermatitis
Cutaneous T-cell lymphoma (Sézary syndrome) (p. 1260)
Malignancy (especially leukaemias)
Pemphigus foliaceus
Pityriasis rubra pilaris (a hereditary disorder of keratinization)
HIV infection
Toxic shock syndrome (p. 125)

Clinical features

It is commoner in males and later in life. Patients often complain of their skin feeling 'tight' as well as itchy. Long-standing erythroderma is often associated with hair loss, ectropion of the eyelids and even nail shedding. Systemic symptoms are common such as malaise, pyrexia, widespread lymphadenopathy and other complications (see below). Erythroderma can occasionally lead to death so it should be regarded as a medical 'emergency'.

Examination should specifically look for pustules and nail changes suggestive of psoriasis.

A skin biopsy may further help to elucidate the cause, especially of cutaneous lymphoma. T cell receptor gene rearrangement studies (looking for evidence of clonal T cell expansion in the skin and blood) are also useful in the diagnosis of lymphoma.

A number of cases defy an exact diagnosis. Lymph node biopsy should be considered in lymphoma. In non-malignant disease lymph nodes normally show non-specific, reactive (dermatopathic) changes.

Complications

The skin is one of the largest organs of the body, so perhaps it is no surprise that inflammation of the whole organ can cause metabolic and haemodynamic problems. Examples are:

- high-output cardiac failure from increased blood flow
- hypothermia from heat loss
- fluid loss by transpiration
- hypoalbuminaemia
- increased basal metabolic rate
- 'capillary leak syndrome'.

Capillary leak syndrome is the most severe complication and has been responsible for a fatal outcome in some cases of psoriasis, although this is extremely rare. It is thought that the inflamed skin releases large quantities of cytokines that cause a generalized vascular leakage. This can cause cutaneous oedema and acute lung injury (p. 917).

Treatment

Treatment of erythroderma is best initiated in hospital. Patients must be kept very warm (with space blankets and heaters). Their vital signs should be monitored regularly, particularly fluid balance. Changes in serum electrolytes and albumin, as well as circulatory status, should be corrected. Swabs should be taken to detect any secondary skin infection.

The skin condition is treated with bed rest and either a bland emollient or a mild topical steroid. All non-essential

FURTHER READING

Yashar SS, Lim HW. Classification and evaluation of photodermatoses. *Dermatologic Therapy* 2003; **16**: 1–7.

drugs should be stopped. Where known, the underlying cause should be treated. The blanket use of systemic steroid therapy for erythroderma remains controversial in view of possible side-effects.

Capillary leak syndrome will often require specialized haemodynamic management in an intensive care unit.

CUTANEOUS SIGNS OF SYSTEMIC DISEASE

Some dermatoses are associated with a variety of underlying systemic diseases. Furthermore, some medical conditions may present with cutaneous features.

Erythema nodosum

Erythema nodosum has a number of underlying causes (Table 23.8). It presents as painful or tender dusky blue-red nodules, commonly over the shins or lower limbs, which fade over 2–3 weeks leaving a bruised appearance (see Fig. 23.39). It is most common in young adults, especially females. It is often unilateral if associated with pregnancy. It may be associated with arthralgia, malaise and fever. Inflammation occurs in the dermis and the subcutaneous layer (panniculitis).

Treatment is of the symptoms with non-steroidal anti-inflammatory drugs (avoid in pregnancy), light compression bandaging and bed rest, as the condition resolves spontaneously. The underlying cause should be treated. In very persistent cases dapsone (100 mg daily), colchicine (500 mg twice daily) or prednisolone (up to 30 mg daily) can be useful.

Erythema multiforme

Erythema multiforme (EM) is a hypersensitivity rash of acute onset frequently caused by infection or drugs. A cell-mediated cutaneous lymphocytotoxic response is present.

In 50% of cases the cause is not found but the following should be considered:

- herpes simplex virus (the most common identifiable cause)
- other viral infections (e.g. EBV, orf)
- drugs (e.g. sulphonamide, anticonvulsants)
- mycoplasma infection
- connective tissue disease (e.g. SLE, polyarteritis nodosa)
- HIV infection
- Wegener's granulomatosis
- carcinoma, lymphoma.

Clinically the lesions can be erythematous, polycyclic, annular or show concentric rings ('target lesions') (Fig. 23.22). Frank blistering is not uncommon. The rash tends to be symmetrical and commonly affects the limbs, especially the hands and feet where palms and soles may be involved. Occasionally there is severe mucosal involvement leading to necrotic ulcers of the mouth and genitalia, and a conjunctivitis ('EM major' – not to be confused with Stevens–Johnson syndrome – see Table 23.9).

The term 'EM minor' is used for cases without mucosal involvement.

Erythema multiforme usually resolves in 2–4 weeks.

Rarely, recurrent erythema multiforme can occur and this is triggered by herpes simplex infection in 80% of cases.

Treatment

This is symptomatic and involves treating the underlying cause. Some advocate the use of oral steroids in severe mucosal disease but this remains controversial.

Recurrent erythema multiforme can be treated with prophylactic oral aciclovir (200 mg twice daily) even if no cause has been found, as 80% of cases appear to be driven by herpes simplex virus. In resistant cases, azathioprine (1–2 mg/kg daily) is used.

Pyoderma gangrenosum

Pyoderma gangrenosum is a condition of unknown aetiology that presents with erythematous nodules or pustules which frequently ulcerate (Fig. 23.23). The ulcers can be large and grow at an alarming speed. The ulcer has a typical bluish black ('gangrenous') undermined edge and a purulent surface ('pyoderma'). There may be an associated pyrexia and malaise. Biopsy through the ulcer edge shows an intense neutrophilic infiltrate and occasionally a vasculitis but the diagnosis depends mostly on the clinical appearance. The main causes are:

Table 23.8 Causes of erythema nodosum

Streptococcal infection*
Drugs (e.g. sulphonamides, oral contraceptive)
Sarcoidosis*
Idiopathic*
Bacterial gastroenteritides, e.g. *Salmonella, Shigella, Yersinia*
Fungal infection (histoplasmosis, blastomycosis)
Tuberculosis
Leprosy
Inflammatory bowel disease
Chlamydia infection

*Common causes

(a) (b)

Fig. 23.22 **Erythema multiforme major. (a)** Target lesions of the palm. **(b)** Mucosal involvement around the mouth.

Table 23.9 Clinical spectra of erythema multiforme, Stevens–Johnson syndrome and toxic epidermal necrolysis

Diagnosis	Skin lesions	Mucosal lesions	Other signs/symptoms
Erythema multiforme	Three-ring target lesions often hands and feet	EM major only	Recent infection (herpes simplex, mycoplasma)
Stevens–Johnson syndrome (SJS)	Scattered macules/ blisters scattered over face, trunk proximal limbs (<10% body surface area) Recent drug exposure Occasional two-ring target lesion	Always	Fever Skin tenderness
Toxic epidermal necrolysis	As for SJS but >30% body surface area involved	Always	As for SJS Respiratory and gastrointestinal lesions Hypotension Decreased consciousness
SJS/TEN overlap	As for SJS (10–30% body surface area involved)	Always	

Fig. 23.23 **Pyoderma gangrenosum.**

- inflammatory bowel disease
- rheumatoid arthritis
- myeloma, monoclonal gammopathy, leukaemia, lymphoma
- liver disease (primary biliary cirrhosis)
- idiopathic (>20% in some series).

Treatment

This is with very potent topical steroids or 0.1% tacrolimus ointment. High-dose oral steroids may be needed to prevent rapidly progressive ulceration. Oral dapsone and minocycline may help. Other immunosuppressants, such as ciclosporin, are useful in resistant cases. The underlying cause should be treated.

Acanthosis nigricans

Acanthosis nigricans presents as thickened, hyperpigmented skin predominantly of the flexures (Fig. 23.24). It can appear warty or velvety when advanced. In early life it is seen in obese individuals who have very high levels of insulin owing to insulin resistance (and this is sometimes termed 'pseudo-acanthosis nigricans'). In older people it normally reflects an underlying malignancy (especially gastrointestinal tumours). Rarely it is associated with hyperandrogenism in females.

Treatment

Topical or oral retinoids (0.5 mg/kg/day) may help (p. 1247), and weight loss is advised in the obese. Any underlying malignancy should be treated.

Fig. 23.24 **Acanthosis nigricans.**

Dermatomyositis (see also p. 547)

The rash is distinctive and often photosensitive. Facial erythema and a magenta-coloured rash around the eyes with associated oedema are usually present. Bluish red nodules or plaques may be present over the knuckles (Gottron's papules) and extensor surfaces. The nail folds are frequently ragged with dilated capillaries. The diagnosis is made from the clinical appearance, muscle biopsy, electromyography (EMG) and a serum creatine phosphokinase. Skin biopsy is not diagnostic.

There is a childhood form which usually occurs before the age of 10 and which eventually resolves. This type is often associated with calcinosis in the skin and can cause significant long-term functional problems with weak muscles and contractures. Life-threatening bowel infarction can also occur in the childhood form. The adult form usually occurs after the age of 40. Some cases are associated with an underlying malignancy whereas some are associated with other autoimmune rheumatic disease. This latter group may overlap with scleroderma and lupus erythematosus.

Treatment

Skin disease may respond to sunscreens and hydroxychloroquine (200 mg twice daily) as well as immunosuppressants, e.g. azathioprine or ciclosporin.

Scleroderma (see also p. 545)

The term scleroderma refers to a thickening or hardening of the skin owing to abnormal dermal collagen. It is not a diagnostic entity in itself. Systemic sclerosis and morphoea both show sclerodermatous changes but are separate conditions.

Systemic sclerosis (often called scleroderma) has cutaneous and systemic features and is discussed fully on page 546.

Morphoea is confined to the skin and usually presents in children or young adults. It is commoner in females and the cause is unknown. Lesions are usually on the trunk and appear as bluish red plaques which progress to induration and then central white atrophy. A linear variant exists in childhood which is more severe as it can cause atrophy of underlying deep tissues and thus cause unequal limb growth or scarring alopecia.

Rarely sclerodermatous skin changes may be seen in Lyme disease (acrodermatitis chronica atrophicans), chronic graft-versus-host disease, polyvinyl chloride disease, eosinophilic myalgia syndrome (due to tryptophan therapy) and bleomycin therapy.

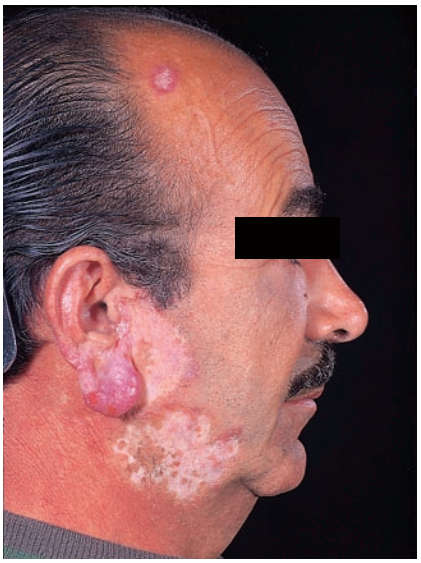


Fig. 23.25 **Chronic discoid lupus erythematosus,** showing scaling, atrophy and hypopigmentation.

Lupus erythematosus (LE)

There are three clinical variants to this disease but some patients may show features of more than one type.

- Chronic discoid lupus erythematosus (CDLE).
- Subacute cutaneous lupus erythematosus (SCLE).
- Systemic lupus erythematosus (SLE).

The aetiology is unknown but is presumably a reflection of some abnormality in immune function as variable autoantibodies may be found in all types. Very rarely it can be induced by certain drugs such as phenothiazines, hydralazine, methyldopa, isoniazid, tetracycline, mesalazine and penicillin.

Chronic discoid lupus erythematosus (CDLE)

CDLE is the most common type of LE seen by dermatologists and more frequently affects females. Clinically it presents with fixed erythematous, scaly, atrophic plaques with telangiectasia, especially on the face or other sun-exposed sites (Fig. 23.25). Hypopigmentation is common and follicular plugging may be apparent. Scalp involvement may lead to a scarring alopecia. Oral involvement (erythematous patches or ulceration) occurs in 25% of cases.

CDLE may be triggered and exacerbated by UV exposure. A few patients may also suffer with Raynaud's phenomenon or unusual chilblain-like lesions (chilblain lupus). Only 5% of cases will go on to develop SLE but this is more common in children. Serum anti-nuclear factor (ANF) is positive in 30% of cases.

Skin biopsy shows a dense patchy, dermal cellular infiltrate (mostly T cells) which often is centred on appendages. Epidermal basal layer damage, follicular plugging and hyperkeratosis may be present. Direct immunofluorescence studies of lesional skin may show the presence of IgM and C3 in a granular band at the dermoepidermal junction ('lupus band').

Treatment

First-line therapy is with sunscreens and potent topical steroids. Certain oral antimalarials (hydroxychloroquine 100–200 mg twice daily and mepacrine 100 mg daily) can prove very useful and are generally safe for long-term intermittent use. Oral prednisolone is beneficial but its use is limited by its side-effect profile. Azathioprine, retinoids, ciclosporin and thalidomide can be useful in resistant cases.

Prognosis

The disease is usually chronic although it may fluctuate in severity. CDLE remains confined to the skin in most patients and it will eventually go into remission in up to 50% of cases (after many years).

Subacute lupus erythematosus (SCLE)

SCLE is a rare cutaneous variant of LE. It presents with widespread indurated, sometimes urticated erythematous lesions, often on the upper trunk. The lesions can also be annular. Photosensitivity is often a prominent feature. Complications, such as arthralgia and mouth ulceration, are seen but significant organ involvement is rare. ANF and extractable nuclear antibodies (anti-Ro and anti-La) are usually positive (see p. 544).

Treatment is with oral dapsone, antimalarials or systemic immunosuppression (prednisolone and ciclosporin).

Systemic lupus erythematosus (SLE)

(see also p. 541)

The cutaneous involvement of SLE is one of the minor problems of the disease but it may be the presenting feature.

Features include macular erythema over the cheeks, nose and forehead ('butterfly rash' – Fig. 23.26). Palmar erythema, dilated nail fold capillaries, splinter haemorrhages and digital infarcts of the finger tips may also be seen but are not always noticed by the patient. Joint swellings, livedo reticularis and purpura are occasionally seen. Rarely SLE can be complicated by an atypical erythema multiforme-like rash ('Rowell's syndrome').

Treatment (p. 544) is usually managed by a rheumatologist.

Fig. 23.26 **Systemic lupus erythematosus** – macular erythema on the cheeks ('butterfly rash').

Table 23.10 Medical conditions associated with pruritus

Iron deficiency anaemia
Internal malignancy (especially lymphoma)
Diabetes mellitus
Chronic kidney disease
Chronic liver disease (especially primary biliary cirrhosis)
Thyroid disease
HIV infection
Polycythaemia vera

Pruritus

The pathophysiology of pruritus (itch) is poorly understood but may be due to peripheral mechanisms (as in skin disease), central or neuropathic mechanisms (as in multiple sclerosis), neurogenic (as in cholestasis/μ opioid receptor stimulation) or psychogenic mechanisms (e.g. parasitophobia). Evidence suggests that low stimulation of unmyelinated C-fibres in the skin is associated with the sensation of itch (high stimulation produces pain). Histamine, tachykinins (e.g. substance P) and cytokines (e.g. interleukin-2) may also play a role peripherally in the skin. The major nerve pathways for itch and the influence of the central nervous system are not well characterized but opioid μ-receptor-dependent processes can regulate the perception and intensity of itch.

Pruritus (see p. 1240, lichen simplex, nodular prurigo/neurodermatitis) in the absence of a demonstrable rash can be caused by a number of different medical problems (Table 23.10).

Asteatotic eczema and cholinergic urticaria are common causes of pruritus where the rash is often missed. The term idiopathic pruritus or 'senile' pruritus probably overlaps with asteatotic eczema and this is common in the elderly.

Treatment involves avoiding soaps and symptomatic measures (as for asteatotic eczema). Phototherapy, low-dose amitryptiline or gabapentin may help intractable cases. Underlying medical problems should be treated.

Sarcoidosis (see also p. 868)

Sarcoidosis is a multisystem granulomatous disorder of unknown aetiology. It may present as reddish brown dermal papules and nodules, especially around the eyelid margins and the rim of the nostrils. More polymorphic lesions (papules, nodules and plaques) may appear on the body. It is most common in black Africans where it is often accompanied by hypo- or hyperpigmentation. Rarely it can present with a bluish red infiltrate or swelling, especially of the nose or ears, called *lupus pernio*. Both these types of lesion can be seen anywhere on the body but are common on the face. Erythema nodosum (p. 1250) of the shins is sometimes seen in acute-onset sarcoidosis. Erythema nodosum is an immunological reaction and not due to sarcoid tissue infiltration. Swollen fingers from a dactylitis may also be present. Whilst sarcoidosis may be confined to the skin, all patients should be investigated for evidence of systemic disease (p. 869).

Treatment of cutaneous lesions includes very potent topical steroids (0.05% clobetasol propionate), intralesional steroids, oral steroids and occasionally methotrexate or antimalarials.

Neurofibromatosis type 1 (von Recklinghausen's disease)

(see also p. 1170)

Type 1 neurofibromatosis is an autosomal dominant condition with high levels of penetrance. It often presents in childhood with a variety of cutaneous features. Many cases are new mutations in the *NF1* gene. Genetic testing is available but clinical evaluation is more helpful. Early signs include café-au-lait spots (brown macules, greater than 2.5 cm in diameter and more than five lesions) and axillary freckling. Lisch nodules (hyperpigmented iris hamartomas) may be seen in the eyes by slit lamp examination. Learning difficulties and skeletal dysplasia occur. Later on, fleshy skin tags and deeper soft tumours (neurofibromas) appear and they can progress to completely cover the skin causing significant cosmetic disability. A number of endocrine disorders are rarely associated including phaeochromocytoma, acromegaly and Addison's disease.

Tuberous sclerosis (epiloia)

The tuberous sclerosis complex (TSC) is an autosomal dominant condition of variable severity which may not present until later childhood. It is characterized by a variety of hamartomatous growths. The three cardinal features are (a) mental retardation, (b) epilepsy, and (c) cutaneous abnormalities – but not all have to be present. In most cases it is due to a mutation in either the *TSC1* gene (encodes hamartin) or the *TSC2* gene (encodes tuberin). Genetic testing is available but the diagnosis still remains clinical, requiring two major features or one major and two minor features. The *skin signs* include:

- adenoma sebaceum (reddish papules/fibromas around the nose)
- periungual fibroma (nodules arising from the nail bed)
- shagreen patches (firm, flesh-coloured plaques on the trunk)
- ash-leaf hypopigmentation (pale macules best seen with UV light)

- forehead plaque (indurated flesh-coloured patch)
- café-au-lait patches
- pitting of dental enamel.

Internal hamartomas can arise in the heart, lung, kidney, retina and CNS. Parents of a suspected case should be carefully examined (under UV light) as they may have a 'forme fruste' of the condition which can manifest just as hypopigmented patches. This and gonadal mosaicism can have genetic implications for future offspring. A large contiguous gene defect may involve *TSC2* and *PKD1* genes causing tuberous sclerosis and polycystic kidney disease together in the same patient.

Diabetes mellitus (see also p. 1057)

Diabetes mellitus can have a number of cutaneous features. Complications of diabetes itself include:

- fungal infection (e.g. candidiasis)
- bacterial infections (e.g. recurrent boils)
- xanthomas
- arterial disease (ulcers, gangrene)
- neuropathic ulcers.

Specific dermatoses of diabetes include:

- necrobiosis lipoidica (a patch of spreading erythema over the shin which becomes yellowish and atrophic in the centre and may ulcerate)
- diffuse granuloma annulare (p. 1254)
- diabetic dermopathy (red-brown flat-topped papules)
- blisters (usually on the feet or hands)
- diabetic stiff skin (tight waxy skin over the fingers with limitation of joint movement owing to thickened collagen – also called cheiroarthropathy).

Chronic liver disease (see also p. 328)

Chronic liver disease may present with jaundice, palmar erythema, spider naevi, white nails, hyperpigmentation and pruritus.

Porphyria cutanea tarda (PCT, p. 1074) is a rare genetic disorder associated with liver disease usually due to hepatic damage from excessive alcohol consumption or hepatitis C infection. Sevnty five per cent of cases are sporadic, 25% familial. Overall, 20% of cases have underlying hereditary haemochromatosis (p. 356). PCT presents clinically on exposed skin with sun-induced blisters, skin fragility, scarring, milia and hypertrichosis. Treatment of the cutaneous features is with repeated venesection and/or very low-dose chloroquine plus avoidance of alcohol. There is anecdotal evidence that specific treatment of hepatitis C (p. 341) will also help the skin, presumably by improving liver function. All patients with PCT are at risk of hepatic carcinoma.

Chronic kidney disease (CKD)
(see also p. 629)

Chronic kidney disease is commonly associated with intractable pruritus. Pallor, hyperpigmentation and ecchymoses are commonly seen. Rarely it is associated with non-inflammatory blisters, pseudo-porphyria cutanea tarda and cutaneous calcification. Long-standing renal transplant patients often suffer with recurrent viral warts and squamous cell carcinomas due to the immunosuppression.

'Nephrogenic fibrosing dermopathy' (also known as *nephrogenic systemic fibrosis*) is a newly described severe scleroderma-like skin disease in a subset of patients with CKD (usually on dialysis). The disease is rapid in onset (days to weeks) with skin discoloration and thickening, joint contractures, muscle weakness and generalized pain. Widespread tissue fibrosis may ensue, causing severe morbidity. Patients may rapidly become wheelchair-bound. The condition is strongly associated with the contrast medium gadolinium used in MRI scans, and such contrast agents are best avoided in patients with low GFRs. There is no accepted treatment but some advocate PUVA and extracorporeal photophoresis. No spontaneous remissions have been recorded. Rapid correction of renal function generally stops the condition progressing.

Calciphylaxis is discussed on page 630.

Thyroid disease (see also p. 985)

Hypothyroidism may cause dry firm gelatinous (myxoedematous) skin with diffuse hair thinning and loss of the outer third of the eyebrows. Hyperthyroidism may be associated with warm sweaty skin and a diffuse alopecia. Graves' disease is rarely associated with thyroid acropachy ('clubbing' with underlying bone changes) and pretibial myxoedema (a red-brown mucinous infiltration of the shins which can become lumpy and tender).

Cushing's syndrome (see also p. 1012)

Cushing's syndrome may cause hirsutism, a moon face, a buffalo hump, stretchmarks (striae) and a pustular folliculitis (often called steroid acne) of the skin.

Hyperlipidaemias (see also p. 1063)

Hyperlipidaemias can present with xanthomas, which are abnormal collections of lipid in the skin. All patients with xanthomas should be investigated for hyperlipidaemia although the most common type, called xanthelasmas (yellow plaques around the eyes), are usually associated with normal lipids. There are a number of other clinical variants of xanthomas such as (i) tuberous xanthoma (firm orange-yellow nodules and plaques on extensor surfaces), (ii) tendon xanthoma (firm subcutaneous swellings attached to tendons), (iii) plane xanthoma (orange-yellow macules often affecting palmar creases), (iv) eruptive xanthoma (numerous small yellowish papules commonly on the buttocks).

Amyloidosis

Macular amyloid is a common purely cutaneous variant seen in Asians. It is characterized by itchy brown rippled macules on the upper back.

Systemic amyloid may be associated with reddish brown papules, nodules or plaques especially around the eyes, the flexural areas and mucosal surfaces. Distinctive periorbital bruising and macroglossia may also be present.

FURTHER READING

Various authors. *Dermatologic Therapy* 2005; **18**: 288–365.

Badgwell C, Rosen T. Cutaneous sarcoidosis therapy updated. *Journal of the American Academy of Dermatology* 2007; **56**: 69–83.

Chung VQ, Moschella SL, Zembowizc A et al. Clinical and pathological findings of paraneoplastic dermatoses. *Journal of the American Academy of Dermatology* 2006; **54**: 745–762.

Curatolo P, Bombardieri R, Jozwiak S. Tuberous sclerosis. *Lancet* 2008; **372**: 657–668.

High WA, Ayers RA, Cowper SE. Gadolinium is quantifiable within the tissue of patients with nephrogenic systemic fibrosis. *Journal of the American Academy of Dermatology* 2007; **56**: 710–712.

Rees J, Murray CS. Itching for progress. *Clinical and Experimental Dermatology* 2005; **30**: 471–473.

Systemic malignant disease

Certain rashes may be a non-metastatic manifestation of an underlying malignancy: paraneoplastic dermatoses (Table 23.11). Rarely tumours can metastasize to the skin where they normally present as papules or nodules which may proceed to ulcerate.

BULLOUS DISEASE

Primary blistering diseases of the skin are rare. A variety of skin proteins hold the skin together. Inherited abnormalities or immune damage of these proteins cause abnormal cell separation, inflammation, fluid accumulation and blistering (Fig. 23.27). The level of blistering determines the clinical picture as well as the prognosis. Therefore skin biopsy for light and electron microscopy together with immunofluorescence (IMF) studies is paramount in diagnosis. However, remember that the commonest causes of skin blistering are chickenpox, herpes, impetigo, pompholyx eczema and insect bite reactions, although these are often localized.

Table 23.11 Non-metastatic cutaneous manifestations of underlying malignancy

Dermatosis	Tumour
Dermatomyositis	Lung, GI tract, GU tract
Acanthosis nigricans	GI tract, lung, liver
Paget's disease (localized patch of eczema around the nipple)	Ductal breast carcinoma
Erythroderma	Lymphoma/leukaemia
Tylosis (thickened palms/soles)	Oesophageal carcinoma
Ichthyosis (dry flaking of skin)	Lymphoma
Erythema gyratum repens (concentric rings of erythema which change rapidly)	Lung, breast
Necrolytic migratory erythema (burning, geographic and spreading annular areas of erythema)	Glucagonoma

IMMUNOBULLOUS DISEASE

Pemphigus vulgaris

Pemphigus vulgaris is a potentially fatal blistering disease occurring in all races but commoner in Ashkenazi Jews and possibly in people from the Indian subcontinent. Onset is usually in middle age and both sexes are affected equally. The development of autoantibodies against the desmosomal protein desmoglein 3 is pathogenic in this disease and the autoantibodies can be measured experimentally as markers of disease activity. Desmoglein 3, an adhesion molecule, is expressed in skin and mucosal surfaces, so both will blister. Rarely the disease can be drug induced (e.g. penicillamine or captopril).

Skin biopsy shows a superficial intraepidermal split just above the basal layer with acantholysis (separation of individual cells). In the rarer variant, pemphigus foliaceus (characterized by anti-desmoglein 1 autoantibodies), the split is higher in the upper epidermis. As desmoglein 1 is not expressed in oral mucosa, only cutaneous blisters are seen in this subtype. Both direct IMF of skin (perilesional) and indirect IMF using patients' serum show intercellular staining of IgG within the epidermis.

Clinical features

Mucosal involvement (especially oral ulceration) is common and is the presenting sign in up to 50% of cases. This may then be followed by the appearance of flaccid blisters, particularly involving the trunk. They tend to be sore rather than itchy. Blistering usually becomes widespread but they rapidly

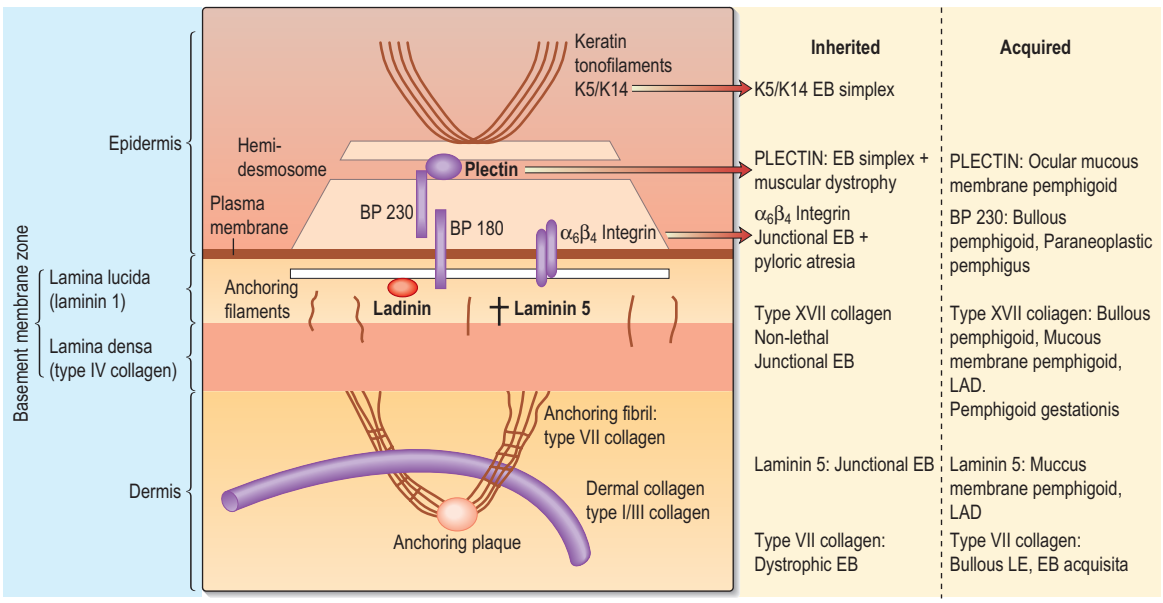


Fig. 23.27 **Section of the basement membrane zone**, showing the structural sites of damage in bullous disorders. LAD, linear IgA disease; EB, epidermolysis bullosa; K, Keratin.

denude; thus pemphigus often presents with erythematous, weeping erosions. Blisters can be extended with gentle sliding pressure (Nikolsky's sign). Flexural lesions often have a vegetative appearance. In pemphigus foliaceus the blisters and erosions often start in a seborrhoeic distribution (scalp, face and upper chest) before becoming more widespread.

Treatment

This is with very high-dose oral prednisolone (60–100 mg daily) or pulsed methylprednisolone and may need to be lifelong. Therefore other immunosuppressants such as azathioprine or mycophenolate mofetil (or occasionally cyclophosphamide or ciclosporin) are used as steroid-sparing agents. Intravenous immunoglobulin infusions may help gain quick control whilst waiting for these other drugs to work. The anti-B cell drug rituximab (anti-CD20 monoclonal antibody) helps multidrug-resistant cases and has been used in combination with immunoglobulin. It is expensive and has been associated with severe side-effects including septicaemia and death in some studies.

Whilst treatment is normally effective, perhaps up to 10% of patients may die either due to complications of the disease or more commonly from side-effects of the treatment.

Bullous pemphigoid

Bullous pemphigoid is more common than pemphigus. It presents in later life (usually over 60) and mucosal involvement is rarer. Autoantibodies against a 230 kDa or 180 kDa hemidesmosomal protein ('bullous pemphigoid antigen 1' and 'type XVII collagen') play an aetiological role.

Skin biopsy shows a deeper blister (than in pemphigus) due to a subepidermal split through the basement membrane. Direct and indirect IMF studies show linear staining of IgG along the basement membrane.

Clinical features

Large tense bullae appear anywhere on the skin (Fig. 23.28) but often involve limbs, hands and feet. They may be centred on an erythematous or urticated background and they can be haemorrhagic. Pemphigoid can be very itchy. Mucosal ulceration is uncommon but a variant of pemphigoid exists which predominantly affects mucosal surfaces with scarring (mucous membrane pemphigoid).

Treatment

This is with high-dose oral prednisolone (30–60 mg daily) and steroid-sparing agents such as azathioprine or mycophenolate mofetil. Weekly methotrexate is also occasionally used. In general, disease control is easier than with pemphigus. Often treatment can be withdrawn after 2–3 years. Treatment often causes side-effects, especially as most patients are elderly. Occasionally localized or mild disease can be controlled with superpotent topical steroids, oral dapsone or high-dose oral minocycline.

Fig. 23.28 **Bullous pemphigoid.**

Dermatitis herpetiformis (see also p. 279)

Dermatitis herpetiformis (DH) is a rare blistering disorder associated with gluten-sensitive enteropathy (coeliac disease). DH and coeliac disease are associated with other organ-specific autoimmune disorders. The HLA associations (B8, DR3, DQ2 in 80–90% of cases) and immunological findings (endomysial, tissue transglutaminase, reticulin and gliadin autoantibodies present in serum) are similar to coeliac disease.

Skin biopsy shows a subepidermal blister with neutrophil microabscesses in the dermal papillae. Direct immunofluorescent studies of uninvolved skin show IgA in the dermal papillae and patchy granular IgA along the basement membrane. The jejunal mucosa shows a partial villous atrophy.

Clinical features

Dermatitis herpetiformis is commoner in males and can present at any age but is most likely to appear for the first time in young adult life. It presents with small, intensely itchy blisters of the skin. The lesions have a predilection for the elbows, extensor forearms, scalp and buttocks. The tops of the blisters are usually scratched off thus crusted erosions are often seen at presentation. Remissions and exacerbations are common.

Treatment

This should always be with a gluten-free diet (GFD, p. 279). Control of the skin disease can be obtained with oral dapsone (50–200 mg daily) or sulphonamides. If a strict GFD is adhered to, oral medication can often be withdrawn after 2 years.

Use of dapsone. Dapsone frequently causes a mild dose-related haemolytic anaemia (which is usually well tolerated) but the haemolysis can be devastating if there is G6PD deficiency. Liver damage, peripheral neuropathy and aplastic anaemia can also rarely occur so regular monitoring of a blood count and liver function is needed.

Linear IgA disease (chronic bullous dermatosis of childhood)

Linear IgA disease is a subepidermal blistering disorder of adults and children. Pathogenic IgA autoantibodies bind to a variety of basement membrane proteins including type XVII collagen and laminin 5 (see Fig. 23.27). It is the most common immunobullous disease seen in children. Rarely it is induced by vancomycin.

Clinical features

Linear IgA disease can present with circular clusters of large blisters, a pemphigoid type of blistering or a dermatitis herpetiformis picture. Mucosal involvement of the mouth, vulva and eyes is not uncommon and can cause scarring. Direct

FURTHER READING

Fassihi H, Wong T, Wessagowit V et al. Target proteins in inherited and acquired blistering skin disorders. *Clinical and Experimental Dermatology* 2006; **31**: 252–259.

Joly P, Mouquet H, Roujeau J-C et al. A single cycle of rituximab for the treatment of severe pemphigus. *New England Journal of Medicine* 2007; **357**: 545–552.

Stanley JR, Amagai M. Pemphigus, bullous impetigo and the staphylococcal scalded-skin syndrome. *New England Journal of Medicine* 2006; **355**: 1800–1810.

IMF studies of skin show linear IgA deposition along the basement membrane.

Treatment

This is with oral dapsone (50–200 mg daily) or sulphonamides. Occasionally immunosuppression is needed. Many patients show spontaneous resolution after 3–6 years.

MECHANOBULLOUS DISEASE (EPIDERMOLYSIS BULLOSA, 'EB')

These are due to inherited abnormalities in structural skin proteins which lead to 'skin fragility'. The resultant blistering tends to arise secondary to trauma and often appears at or shortly after birth. These conditions can be a mild inconvenience, severely disabling or fatal but fortunately are very rare. There are three groups of disorders in which the fundamental gene/protein abnormalities have been characterized.

Epidermolysis bullosa simplex is a group of autosomal dominant genodermatoses characterized by 'superficial' blistering owing to mutations of cytoskeleton proteins within the basal layer of the epidermis, e.g. keratin 5 (chromosome 12q) or keratin 14 (chromosome 17q). Most forms of EB simplex show mild disease with intermittent blistering of the hands and feet, especially in hot weather. The teeth and nails are normal and scarring is absent.

Epidermolysis bullosa dystrophica is a group of genodermatoses characterized by 'deeper' blistering associated with scarring and milia formation. The level of split is deep within the basement membrane and is due to a mutation in the *COL-7A1* gene (locus at chromosome 3p21.1) which causes a loss of collagen VII in the anchoring fibrils. Nails, mucosae and even the larynx are often involved. The autosomal dominant variety is milder but the autosomal recessive type produces severe disease with disabling scarring, fusion of digits, joint contractures and dysphagia. Life expectancy is significantly reduced. Repeated scarring results in the development of multiple squamous cell carcinomas and most die from this complication in early adult life. The average life expectancy after the appearance of the first squamous cell carcinoma is 5 years.

Junctional epidermolysis bullosa is the most severe form. It is characterized by a split in the lamina lucida of the basement membrane and is due to mutations in various proteins, mainly laminin 5 but also $\alpha_6\beta_4$ integrin or the 180 kDa bullous pemphigoid-2 antigen. It presents at birth with widespread blistering and areas of absent skin. Erosions of the central face and hoarseness from laryngeal involvement are common. Nail and teeth abnormalities are also common. Both a lethal and a rarer non-lethal form of junctional EB exist and they show an autosomal recessive inheritance. The lethal form causes death in infancy or early childhood.

Investigation and treatment

Investigation and treatment of EB should be carried out in a specialist centre. Diagnosis at birth on clinical grounds is difficult and should be avoided. Exact diagnosis depends on ultrastructural analysis of induced blisters in the skin and immunohistochemistry, and these investigations can be further confirmed with genetic testing. Only then can prognosis and genetic counselling be given accurately to parents. Prenatal diagnosis is available for the more severe forms of EB.

SKIN TUMOURS

BENIGN CUTANEOUS TUMOURS

Melanocytic naevi (moles)

Moles are a benign overgrowth of melanocytes that are common in white-skinned people. They appear in childhood and increase in number and size during adolescence and early adult life. They often start as flat brown macules with proliferation of melanocytes at the dermoepidermal junction (junctional naevi). The melanocytes continue to proliferate and grow down into the dermis (compound naevi) which causes an elevation of the mole above the skin surface. The pigmentation is usually even and the border regular. They eventually mature into a dermal naevus (cellular naevus), often with a loss of pigment.

Blue naevus is an acquired asymptomatic blue-looking mole. It is due to a proliferation of melanocytes deep in the mid-dermis.

Basal cell papilloma (seborrhoeic wart)

This is a common benign overgrowth of the basal cell layer of the epidermis. The lesion can be flesh coloured, brown or even black and often has a greasy appearance. The surface is irregular and warty and the lesions appear very superficial as though stuck on to the skin (Fig. 23.29). Tiny keratin cysts may be seen on the surface. They can be treated with cryotherapy or curettage.

Dermatofibroma (histiocytoma)

Dermatofibromas appear as firm, elevated pigmented nodules which may feel like a button in the skin. A peripheral ring of pigmentation is sometimes seen. They are often found on the leg and are commoner in females. There may be a preceding history of trauma or insect bite. The lesion consists of histiocytes, blood vessels and varying degrees of fibrosis. If symptomatic, excision is required.

Epidermoid cyst (previously 'sebaceous cyst')

Epidermoid cysts present as cystic swellings of the skin with a central punctum. They contain 'cheesy' keratin rather than sebum; thus the old term 'sebaceous cyst' should be avoided. These cysts occasionally rupture causing significant dermal inflammation which is not infected.

Pilar cyst (trichilemmal cyst)

Pilar cysts are smooth cysts without a punctum usually found on the scalp. They may be multiple and familial.

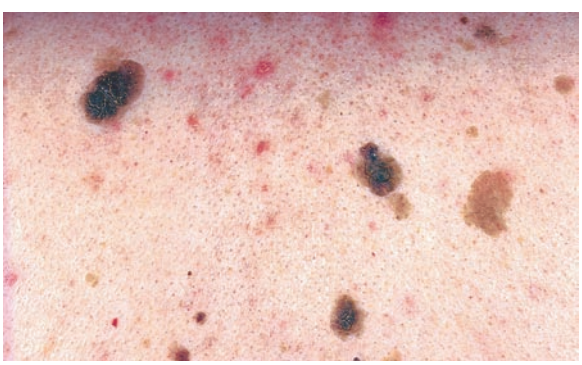


Fig. 23.29 **Seborrhoeic warts** (basal cell papillomas).

Fig. 23.30 **Keratoacanthoma.**

Keratoacanthoma

Keratoacanthomas are rapidly growing epidermal tumours which develop central necrosis and ulceration (Fig. 23.30). They occur on sun-exposed skin in later life and can grow up to 2–3 cm across. Whilst they may resolve spontaneously over a few months, they are best excised both to exclude a squamous cell carcinoma (which they can mimic) and to improve the cosmetic outcome.

Pyogenic granuloma (granuloma telangiectaticum)

Pyogenic granulomas are a benign overgrowth of blood vessels. They present as rapidly growing pinkish red nodules which are friable and readily bleed. They may follow trauma and are often found on the fingers and lips. They are best excised to exclude an amelanotic malignant melanoma.

Cherry angioma (Campbell de Morgan spots)

These are benign angiokeratomas that appear as tiny pinpoint red papules, especially on the trunk, and increase with age. No treatment is required.

POTENTIALLY PRE-MALIGNANT CUTANEOUS TUMOURS

Solar keratoses (actinic keratoses)

These frequently develop later in life in white-skinned people who have had significant sun exposure. They appear on exposed skin as erythematous silver-scaly papules or patches with a conical surface and a red base (Fig. 23.31). The background skin is often inelastic, wrinkled and may show flat brown macules ('liver spots' or solar lentigos) reflecting diffuse solar damage. A small proportion of these keratoses can transform into squamous cell carcinoma but only after many years.

Treatment of lesions is with cryotherapy, topical 5-fluorouracil cream or 5% imiquimod cream.

Bowen's disease

This is a form of intraepidermal carcinoma-in-situ which rarely can become invasive. It is thought to be due to long-term sun exposure. It presents on exposed skin, most commonly women's legs, as an isolated scaly red patch or plaque looking rather like psoriasis although it has a rather irregular edge. The lesions do not clear but slowly increase in size with time.

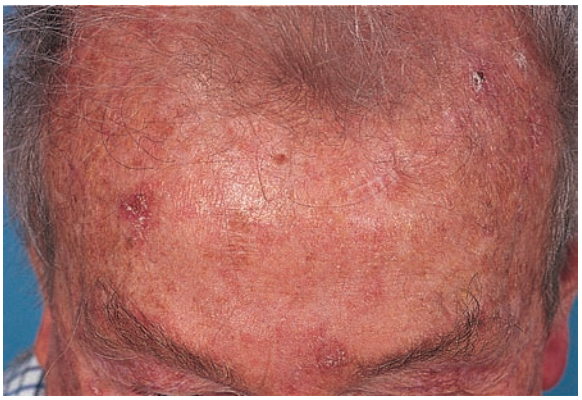


Fig. 23.31 **Solar keratoses with background actinic damage.**

A variant which can show partial or full-thickness dysplasia can involve the epidermis of the mucosa or neighbouring skin. This can affect the vulva, the glans penis and perianal skin and is termed vulval (penile, or anal) intraepithelial dysplasia. Clinically it can present as non-specific erythema or as a warty thickening. These diseases have a stronger link with HPV and probably have a higher pre-malignant potential than Bowen's disease. They are commoner in immunosuppressed individuals. The anal form is increasingly reported in HIV-positive patients (as indeed is anal carcinoma) and extension into the rectum may occur.

Treatment is with topical 5-fluorouracil, 5% imiquimod cream, cryotherapy, curettage, photodynamic therapy or a tissue-destructive laser.

Atypical mole syndrome (dysplastic naevus syndrome)

This is often familial. A large number of melanocytic naevi begin to appear in childhood even on unexposed sites. Individual lesions may be large with irregular pigmentation and border, and histologically they may show cytological and architectural atypia but no frank malignant change. Individuals with this condition have an increased risk of developing malignant melanoma. They should have their moles photographed and be regularly reviewed. Suspicious lesions should be excised.

Giant congenital melanocytic naevi

These are very large moles present at birth. Only the very large lesions (>20 cm across) show an increased risk of developing malignant melanoma and the risk has recently been downgraded: each lesion probably has a <5% risk of malignancy. Excision may be considered but is rarely possible without multiple operations and marked disfigurement so regular monitoring is advised. There have been a number of recent reports showing that a few of these lesions improve spontaneously and partially resolve during childhood.

Lentigo maligna

This is a slow-growing macular area of pigmentation seen in elderly people, commonly on the face. The border and pigmentation are often irregular. Some people regard this lesion as a melanoma-in-situ. There is an increased risk of developing invasive malignant melanoma. Treatment is by excision if possible but 5% imiquimod cream is currently

being tried in the very large lesions where surgery would be disfiguring.

MALIGNANT CUTANEOUS TUMOURS

Basal cell carcinoma (rodent ulcer)

Basal cell carcinomas are the most common malignant skin tumour and most relate to excessive sun exposure. They are common later in life on exposed sites although rare on the ear. They are 10–20 times more common in the chronically immunosuppressed solid organ transplant population. They can present as a slow-growing papule or nodule (or rarely be cystic) which may go on to ulcerate (Fig. 23.32). Telangiectasia over the tumour or a skin-coloured jelly-like 'pearly edge' may be seen. A flat, diffuse superficial form exists as an ill-defined 'morphoeic' variant. Basal cell carcinomas will slowly grow and erode structures if untreated but these tumours almost never metastasize.

Treatment

Treatment is usually with surgical excision with a 3–5 mm border. Radiotherapy, photodynamic therapy, cryotherapy or 5% imiquimod cream can be useful for large superficial forms but follow-up for recurrence is required. Curettage may occasionally be used in older patients although not for central facial lesions as they often recur. Recurrent tumour or morphoeic basal cell carcinoma is best treated with Mohs' micrographic surgery to ensure adequate clearance.

Squamous cell carcinoma

Squamous cell carcinoma (SCC) is a more aggressive tumour than basal cell carcinoma as it can metastasize if left untreated. Most relate to sun exposure, and daily application of sun cream has been shown to reduce the incidence in Australia. SCC can arise in pre-existing solar keratoses or Bowen's disease or be due to chronic inflammation such as in lupus vulgaris. Rarely multiple tumours may arise due to arsenic ingestion in early life. Multiple tumours also occur in people who have had prolonged periods of immunosuppression such as renal transplant patients where certain human papilloma virus subtypes may be involved in malignant transformation.

Clinically the lesions are often keratotic, rather ill-defined nodules which may ulcerate (Fig. 23.33). They can grow very rapidly. Examination of regional lymph nodes is essential. They are most common on sun-exposed sites in later life. Ulcerated lesions on the lower lip or ear are often more aggressive.

Treatment is with excision or occasionally radiotherapy. Curettage should be avoided.

Malignant melanoma

Malignant melanoma is the most serious form of skin cancer as metastasis can occur early and it causes a number of deaths even in young people. As with other types of skin cancer the incidence is continuing to increase, probably due to excessive exposure to sunlight. The history of childhood sun exposure and intermittent intense sun exposure appears to be necessary for the development of malignant melanoma. Other risk factors include pale skin, multiple melanocytic naevi (>50), sun sensitivity, immunosuppression, atypical mole syndrome, giant congenital melanocytic naevi, lentigo maligna and a positive family history of malignant melanoma. Malignant melanoma is commoner in later life but many young adults are also affected. A number of oncogenes and tumour suppressor proteins (CDK4, CDKN2a, B-Raf, PTEN, Ras, Rb, p53, p16) have been implicated in the pathogenesis of melanoma.

Diagnosis of melanoma is not always easy but the clinical signs listed in Table 23.12 help distinguish malignant from benign moles. Examination with a dermatoscope (a hand-held polarized light source with magnification) can further help in detecting malignant lesions.

Four clinical types exist:

- *Lentigo maligna melanoma* is where a patch of lentigo maligna develops a papule or nodule signalling invasive tumour.
- *Superficial spreading malignant melanoma* is a large flat irregularly pigmented lesion which grows laterally before vertical invasion develops.

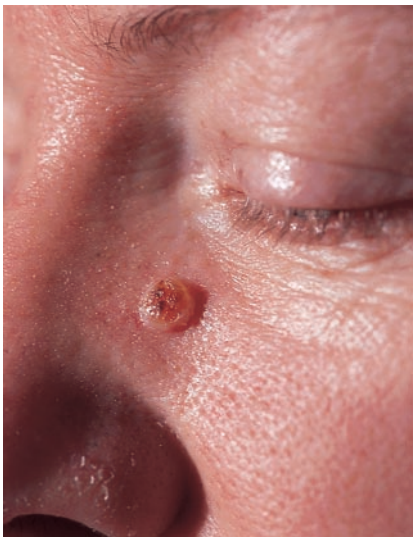


Fig. 23.32 **Ulcerating basal cell carcinoma.**

Fig. 23.33 **Squamous cell carcinoma.**

Table 23.12 Clinical criteria for the diagnosis of malignant melanoma

ABCDE criteria (USA)	
Asymmetry of mole	
Border irregularity	
Colour variegation	
Diameter >6 mm	
Elevation	
The Glasgow 7-point checklist	
Major criteria	Change in size Change in shape Change in colour
Minor criteria	Diameter more than 6 mm Inflammation Oozing or bleeding Mild itch or altered sensation

- *Nodular malignant melanoma* (Fig. 23.34) is the most aggressive type. It presents as a rapidly growing pigmented nodule which bleeds or ulcerates. Rarely they are amelanotic (non-pigmented) and can mimic pyogenic granuloma.
- *Acral lentiginous malignant melanomas* arise as pigmented lesions on the palm, sole or under the nail and they usually present late. They may not be related to sun exposure.

Treatment

In the UK all patients with melanoma >1 mm thick should be referred to their regional multidisciplinary team (dermatology/plastic surgery/oncology/melanoma nurse) for expert management. Surgery is the only curative treatment and entails urgent wide excision: 1 cm margin for thin melanomas (<1 mm) increasing to 3 cm margin for thicker melanomas (>2 mm). Histological analysis will determine the depth of invasion ('Clark's level') and the thickness of the tumour ('Breslow thickness'). Patients with melanomas greater than 1 mm are staged according to the American AJCC criteria which look at tumour thickness, metastasis and lymph node status and this helps predict prognosis and 5-year survival rates. Sentinel node biopsy for patients with thicker lesions is currently under assessment as a tool for predicting prognosis and it may prove to be useful in determining therapy (e.g. lymph node dissection) and improving survival but this is only available in a few centres and it remains a research tool. Metastatic disease can involve surgery to lymph nodes, isolated limb perfusion, radiotherapy, immunotherapy and chemotherapy but none of these is currently very effective. Initial optimism for high-dose alpha-interferon therapy in advanced disease has recently been challenged with a systematic review suggesting no clear benefit. Other areas of therapeutic research involve immunization with the ganglioside GM_2 (a molecule frequently expressed on melanoma cell surfaces), VEGF receptor tyrosine kinase inhibitors and VEGF inhibitors, but none of these is yet of proven benefit.

The role of governments and medical personnel in public health education to discourage sunbathing and encourage the use of sunscreens is of the utmost importance in skin cancer prevention.

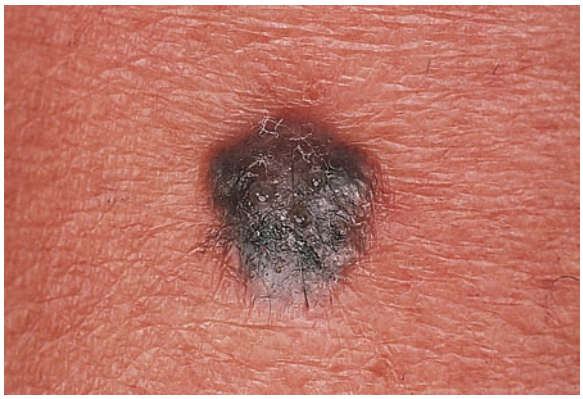


Fig. 23.34 **Nodular malignant melanoma.**

Cutaneous T cell lymphoma (mycosis fungoides)

This is a rare type of skin tumour which often follows a relatively benign course. It presents insidiously with scaly patches and plaques which can look eczematous or psoriasiform. Lesions often appear initially on the buttocks. These lesions may come and go or remain persistent over many years. Patients may well die of unrelated causes. Skin biopsy confirms the diagnosis, showing invasion by atypical lymphocytes. T cell receptor gene rearrangement studies show that there is often a monoclonal expansion of lymphocytes in the skin.

Occasionally the disease can progress to a cutaneous nodular or tumour stage, which may be accompanied by systemic organ involvement. In elderly males the disease may progress rarely to an erythrodermic variant accompanied by lymphadenopathy and peripheral blood involvement ('Sézary syndrome', 483).

All patients should be staged at the time of diagnosis to assess for any systemic involvement.

Treatment

Early cutaneous disease can be left untreated or treated with topical steroids or PUVA. More advanced disease of the skin, or systemic involvement, may require radiotherapy, chemotherapy (e.g. methotrexate), immunotherapy or electron beam therapy. Bexarotene, a retinoid X receptor agonist, is a new therapy for more advanced disease.

Kaposi's sarcoma

This is a tumour of vascular and lymphatic endothelium that presents as purplish nodules and plaques. There are three types:

- *The 'classic' or 'sporadic' form* (as described by Kaposi) occurs in elderly males, especially Jewish people from Eastern Europe. It presents as slow-growing purple tumours in the foot and lower leg which rarely cause any significant problem.
- *The 'endemic' form* occurs in males from central Africa and shows more widespread cutaneous involvement as well as lymph node (or occasionally systemic) involvement. Oedema is a prominent feature.
- *The immunosuppression-related form* is more severe and is most common in homosexual patients with HIV. Lesions are widespread and often affect the skin, bowel, oral cavity and lungs.

FURTHER READING

Braun RP, Rabinovitz HS, Oliviero M et al. Dermoscopy of pigmented skin lesions. *Journal of the American Academy of Dermatology* 2005; **52**: 109–121.

DiLorenzo G, Konstantinopoulos PA, Pantanowitz L et al. Management of AIDS-related Kaposi's sarcoma. *Lancet Oncology* 2007; **8**: 167–176.

Miller AJ, Mihm MC Jr. Melanoma. *New England Journal of Medicine* 2006; **355**: 51–65.

Rubin AI, Chen EH, Ratner D. Basal-cell carcinoma. *New England Journal of Medicine* 2005; **353**: 2262–2269.

All three types have a strong association with herpes virus type 8 but other factors must be involved as herpes type 8 seroprevalence is up to 10% in the USA and 50% in some African countries. HAART (p. 191) has significantly reduced the incidence of Kaposi's sarcoma in HIV although the prevalence has started to increase again over the last few years for as yet unexplained reasons.

Treatment

Treatment of advanced Kaposi's sarcoma is with radiotherapy, immunotherapy or chemotherapy.

DISORDERS OF BLOOD VESSELS/LYMPHATICS

LEG ULCERS

Leg ulcers are common and can have many causes (Table 23.13). Venous ulcers are the most common type in developed countries.

Venous ulcers

Venous ulcers are the result of sustained venous hypertension in the superficial veins, owing to incompetent valves in the deep or perforating veins or to previous deep vein thrombosis. The increased pressure causes extravasation of fibrinogen through the capillary walls, giving rise to perivascular fibrin deposition, which leads to poor oxygenation of the surrounding skin.

Venous ulcers are common in later life and cause a significant drain on health-care budgets as they are often chronic and recurrent; they affect 1% of the population over the age of 70 years. They are most commonly found on the lower leg in a triangle above the ankles (Fig. 23.35), and may be associated with:

- oedema of the lower legs
- venous eczema (p. 1239)
- brown pigmentation from haemosiderin, which is derived from red blood cells that have extravasated through the capillary wall
- varicose veins, but not invariably
- lipodermatosclerosis (the combination of induration, reddish brown pigmentation and inflammation) – a fibrosing panniculitis of the subcutaneous tissue
- scarring white atrophy with telangiectasia (atrophie blanche).

Table 23.13 Causes of leg ulceration

Venous hypertension
Arterial disease (e.g. atherosclerosis)
Neuropathic (e.g. diabetes, leprosy)
Neoplastic (e.g. squamous or basal cell carcinoma)
Vasculitis (e.g. rheumatoid arthritis, SLE, pyoderma gangrenosum)
Infection (e.g. ecthyma, tuberculosis, deep mycoses, Buruli ulcer, syphilis, yaws)
Haematological (e.g. sickle cell disease, spherocytosis)
Drug (e.g. hydroxycarbamide (hydroxyurea))
Other (e.g. necrobiosis lipoidica, trauma, artefact)

Treatment is with high-compression bandaging (e.g. Unna boot or four-layer bandaging) and leg elevation to try to decrease the venous hypertension. Doppler studies should always be done before bandaging to exclude arterial disease. This treatment is best delivered in the community by appropriately trained nurses. 'Four-layer bandaging' is increasingly popular as this provides high levels of graduated compression (with pressures decreasing up the leg). Ulcer dressings are used to keep the ulcer moist and free of slough and exudates. Up to 80% of ulcers can be healed within 26 weeks. Slower healing rates occur in patients with decreased mobility and if the ulcers are very large, present for longer than 6 months or bilateral. Diuretics are sometimes helpful to reduce the oedema. Antibiotics are necessary only for overt bacterial infection. Unusual fungal infection ('tinea incognito') is increasingly reported under compression bandaging.

Venous leg ulcers can be very painful so adequate analgesia should be given, including opiates if required. Split-thickness skin grafting is used in resistant cases. Support stockings (individually fitted) should be worn lifelong after healing as this lessens recurrence.

Underlying venous disease is best investigated with duplex ultrasound or plethysmography. Surgery for purely superficial venous disease can occasionally be useful for ulcer healing but, in general, venous surgery is unhelpful.

Support stockings should be used for prophylaxis against ulcers.

Arterial ulcers

Arterial ulcers present as punched-out, painful ulcers higher up the leg or on the feet. There may be a history of claudication, hypertension, angina or smoking. Clinically the leg is cold and pale. Absent peripheral pulses, arterial bruits and loss of hair may be present. Doppler ultrasound studies will confirm arterial disease.

Treatment depends on keeping the ulcer clean and covered, adequate analgesia and vascular reconstruction if appropriate. Compression bandaging must not be used.

Neuropathic ulcers

Neuropathic ulcers tend to be seen over pressure areas of the feet, such as the metatarsal heads, owing to repeated trauma. These are most commonly seen in diabetics due to peripheral neuropathy. In some developing countries leprosy is a common cause.

Fig. 23.35 **Venous leg ulcer.**

FURTHER READING

NICE UK guidelines on pressure sores:

http://guidance.nice.org.uk/CGB *(Risk assessment and prevention)*

http://guidance.nice.org.uk/CG29 *(Management)*

Bergan JJ, Schmid-Schönbein GW, Smith PD et al. Chronic venous disease. *New England Journal of Medicine* 2006; **355**: 488–498.

Carlson JA, Cavaliere LF, Grant-Kels JM. Cutaneous vasculitis: diagnosis and management. *Clinics in Dermatology* 2006; **24**: 414–429.

Reddy M, Gill SS, Rochon PA. Preventing pressure ulcers: a systematic review. *Journal of the American Medical Association* 2006; **296**: 974–984.

Ask the authors
www.studentconsult.com

Treatment depends on keeping the ulcer clean and removing pressure or trauma from the affected area. Diabetics should pay particular attention to foot care and correctly fitting shoes with the help of a specialist podiatrist (p. 1056).

PRESSURE SORES (DECUBITUS ULCERS, BEDSORES)

These occur in the elderly, immobile, unconscious or paralysed patients. They are due to skin ischaemia from sustained pressure over a bony prominence, most commonly the heel and sacrum. Normal individuals feel the pain of continued pressure, and even during sleep movement takes place to change position continually. Pressure sores may be graded:

- Stage I: non-blanchable erythema of intact skin
- Stage II: partial-thickness skin loss of epidermis/dermis (blister or shallow ulcer)
- Stage III: full-thickness skin loss involving subcutaneous tissue but not fascia
- Stage IV: full-thickness skin loss with involvement of muscle/bone/tendon/joint capsule

There are numerous risk factors for development of pressure sores (Table 23.14).

The majority of pressure sores occur in hospital. Seventy per cent appear in the first 2 weeks of hospitalization, and 70% are in orthopaedic patients, especially those on traction. Between 20% and 30% of pressure sores occur in the community.

Eighty per cent of patients with deep ulcers involving the subcutaneous tissue die in the first 4 months.

The early sign of red/blue discoloration of the skin can lead rapidly to ulcers in 1–2 hours. Leaving patients on hard emergency room trolleys, or sitting them in chairs for prolonged periods, must be avoided.

Management

Prevention

Prevention is better than cure. Specialist 'tissue-viability nurses' help identify at-risk patients and train other clinical staff. Several risk assessment tools have been devised for the immobile patient based on the known risk factors. The 'Norton scale' and Waterlow Pressure Sore Risk Assessment (Box 23.3) are two such validated systems which produce a

Table 23.14 Risk factors for the development of pressure sores

Prolonged immobility:
- paraplegia
- arthritis
- severe physical disease
- apathy
- operation and postoperative states
- plaster casts
- intensive care

Decreased sensation:
- coma, neurological disease, diabetes mellitus
- drug-induced sleep

Vascular disease:
- atherosclerosis, diabetes mellitus, scleroderma, vasculitis

Poor nutrition:
- anaemia
- hypoalbuminaemia
- vitamin C or zinc deficiency

Box 23.3 Pressure sore risk-assessment tools

Norton scale for pressure sores

Physical	Neurology	Activity	Mobility	Incontinence
4 Good	4 Alert	4 Ambulant	4 Full	4 None
3 Fair	3 Apathetic	3 Walks with help	3 Slightly	3 Occasionally
2 Poor	2 Confused	2 Not bound	2 Limited*	2 Usually
1 Very poor	1 Stupor	1 Bedfast	1 Very limited; Immobile	1 Double

*Low scores carry a high risk

Waterlow Pressure Sore Risk Assessment

Build/weight for height		Visual skin type		Continence		Mobility		Sex Age		Appetite	
Average	0	Healthy	0	Complete	0	Fully mobile	0	Male	1	Average	0
Above average	2	Tissue paper	1	Occasionally incontinent	1	Restricted/ difficult	1	Female	2	Poor	1
Below average	3	Dry	1					14–18	1	Anorectic	2
		Oedematous	1	Catheter/ incontinent of faeces	2	Restless/fidgety	2	50–64	2		
		Clammy	1					65–75	3		
		Discoloured	2			Apathetic	3	75–80	4		
		Broken/spot	3	Doubly incontinent	3	Inert/traction	4	81+	5		

Special risk factors		Assessment value	
1. Poor nutrition; e.g. terminal cachexia	8	At risk	10
2. Sensory deprivation, e.g. diabetes, paraplegia, cerebrovascular disease	6	High risk	15
3. High-dose anti-inflammatory or steroids in use	3	Very high risk	20
4. Smoking 10+ per day	1		
5. Orthopaedic surgery/fracture below waist	3		

numerical score, enabling staff to identify those at most risk.

Treatment

- bed rest with pillows and fleeces to keep pressure off bony areas (e.g. sacrum and heels) and prevent friction
- air-filled cushions for patients in wheelchairs
- special pressure-relieving mattresses and beds
- regular turning but avoid pressure on hips
- ensure adequate nutrition
- non-irritant occlusive moist dressings (e.g. hydrocolloid)
- adequate analgesia (may need opiates)
- plastic surgery (debridement and grafting in selected cases)
- treatment of underlying condition

VASCULITIS (see also p. 548)

Vasculitis is the term applied to an inflammatory disorder of blood vessels which causes endothelial damage. Cutaneous vasculitis (confirmed by skin biopsy) may be an isolated problem but occasionally is associated with vasculitis in other organs. The most commonly used classification is based on the size of blood vessel involved (see Tables 10.18 and 10.19).

The cutaneous features are of haemorrhagic papules, pustules, nodules or plaques which may erode and ulcerate. These purpuric lesions do not blanch with pressure from a glass slide ('diascopy'). Occasionally a fixed livedo reticularis pattern may appear which does not disappear on warming. Pyrexia and arthralgia are common associations even in the absence of significant systemic involvement. Other clinical features depend on the underlying cause.

Leucocytoclastic vasculitis (LCV) or angiitis is the most common cutaneous vasculitis affecting small vessels. This usually appears on the lower legs as a symmetrical palpable purpura. It is rarely associated with systemic involvement. It can be caused by drugs (15%), infection (15%), inflammatory disease (10%) or malignant disease (<5%) but often no cause is found (55–60%). Investigations are only necessary with persistent lesions or associated signs and symptoms. Whilst LCV often settles spontaneously, treatment with analgesia, support stockings, dapsone or prednisolone may be needed to control the pain and to heal up any ulceration. *Urticarial vasculitis* is discussed on page 1243.

Calciphylaxis (calcific uraemic arteriopathy) is described on page 630.

LYMPHATICS

Lymphoedema

Lymphoedema refers to a chronic non-pitting oedema due to lymphatic insufficiency. It is most commonly seen affecting the legs and tends to progress with age. The legs can become enormous and prevent wearing of normal shoes. Chronic disease may cause a secondary 'cobblestone' thickening of the skin. Lymphoedema can be primary (and present early in life) due to an inherited deficiency of lymphatic vessels (e.g. Milroy's disease) or can be secondary due to obstruction of lymphatic vessels (e.g. filarial infection or malignant disease).

Treatment is with compression stockings and physical massage. If there is recurrent cellulitis, long-term antibiotics are advisable as each episode of cellulitis will further damage the lymph vessels. Surgery should be avoided.

Lymphangioma circumscriptum

This is a rare hamartoma of lymphatic tissue. It usually presents in childhood with multiple small vesicles in the skin which weep lymphatic fluid and sometimes blood. They reflect deeper vessel involvement so surgery should be avoided. Cryotherapy or CO_2 laser treatment may help the superficial lesions.

DISORDERS OF COLLAGEN AND ELASTIC TISSUE

Ehlers–Danlos syndrome (see also p. 568)

Ehlers–Danlos syndrome (EDS) can be subdivided into at least 10 variants. They are all inherited disorders causing abnormalities in collagen of the skin, joints and blood vessels. Clinically this causes increased elasticity of the skin, hypermobile joints and fragile blood vessels causing easy bruising or in some cases internal haemorrhage. The skin is hyperextensible but recoils normally after stretching. It is easily injured and heals slowly with scarring like tissue paper. Pseudotumours may occur at the sites of scarring (such as elbows and knees) consisting mainly of fat, but calcification can occur.

Pseudoxanthoma elasticum

Pseudoxanthoma elasticum is a rare group of disorders characterized by abnormalities in collagen and elastic tissue affecting the skin, eye and blood vessels. The skin may be loose, lax and wrinkled. It can look yellowish and papular ('plucked chicken skin') and tends to loose its elastic recoil when stretched. Skin changes are best seen in the flexures, especially the sides of the neck. Non-cutaneous features are not always present but they include recurrent gastrointestinal bleeding, early myocardial infarction, claudication and angioid streaks on the retina reflecting disruption of vascular elastic tissue.

Striae

Striae are visible linear scars due to dermal collagen damage and stretching. Histologically a thinned epidermis overlies parallel bundles of fine collagen. They occur commonly over the abdomen and breasts in pregnancy but also occur on the thighs and trunk in rapidly growing adolescents as well as in some obese individuals. They are also seen in Cushing's syndrome and corticosteroid therapy. They are also rarely seen in Marfan's syndrome (p. 780). Striae are initially reddish blue but fade to white atrophic marks. Puberty-related striae normally disappear completely.

Keloid scars

Keloid scars are characterized by smooth hard nodules (Fig. 23.36) due to excessive collagen production. They may occur spontaneously or follow skin trauma/surgery and they are often itchy. They tend to affect young adults and are much

FURTHER READING

Robles DT, Berg D. Abnormal wound healing: keloids. *Clinics in Dermatology* 2007; **25**: 26–32.

Uitto J, Jiang Q. Pseudoxanthoma elasticum-like phenotypes: more diseases than one. *Journal of Investigative Dermatology* 2007; **127**: 507–510.

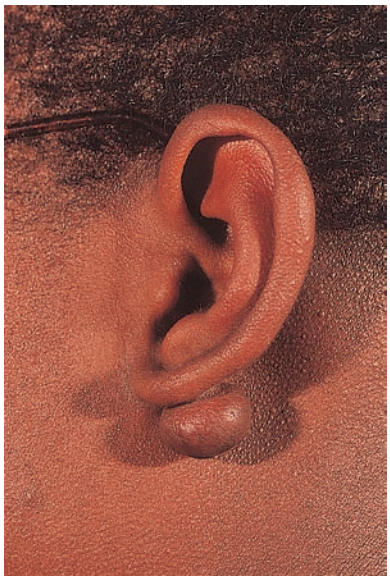


Fig. 23.36 Keloid scar of the lobe of the ear.

commoner in black Africans. Sites of predilection include the shoulders, upper back and chest, earlobes and chin. Unlike hypertrophic scars (which fade within 12 months) keloids are persistent and may continue to enlarge.

Treatment is with triamcinolone injection, compression with silica gels or surgery, but surgery must be followed by steroid injection or superficial radiotherapy or it may make the problem worse.

DISORDERS OF PIGMENTATION

HYPOPIGMENTATION

Vitiligo

Vitiligo is a common autoimmune disorder of depigmentation due to areas of melanocyte loss. Sufferers often have relatives with other organ-specific autoimmune disorders. It presents in childhood or early adult life with well-demarcated macules of complete pigment loss. There is no history of preceding inflammation. Patients are very susceptible to sunburn. Lesions are often symmetrical and frequently involve the face, hands and genitalia (Fig. 23.37). The hair can also depigment. Trauma may induce new lesions. Spontaneous repigmentation can occur and often starts around hair follicles, giving a speckled appearance. However repigmentation is rare if a lesion has persisted for more than 1 year or if the hair is depigmented. The psychological consequences of vitiligo can be devastating, especially in Asian or black African people.

Treatment is often unsatisfactory and has no impact on the long-term outcome. Sunblocks should be used to prevent burning. Potent topical steroids or phototherapy may help some individuals. Tacrolimus ointment (0.1%) is being assessed. If vitiligo is almost universal and fixed, depigmentation with monobenzone can be used. Finally, referral to a specialist camouflage clinic is often the most helpful 'treatment'.

Fig. 23.37 Vitiligo of the hands showing areas of depigmentation.

Post-inflammatory hypopigmentation

This is one of the most common causes of pale skin. It is much more common in people with pigmented skin. It may be seen as a consequence of eczema, acne or psoriasis and may even be the reason for individuals presenting to a doctor. Providing the skin disease is controlled the pigmentation will recover slowly after many months. Post-inflammatory hyperpigmentation can also occur.

Oculocutaneous albinism

This is a group of rare autosomal recessive disorders affecting the pigmentation of skin, hair and eyes. It can affect all races. Melanocytes are normal in number but have abnormal function. Clinically it presents with universal pale skin, white or yellow hair and a pinkish iris. Photophobia, nystagmus and a squint are also present in most cases.

Treatment involves obsessive protection against sunlight to avoid sun burning and development of skin cancer.

Idiopathic guttate hypomelanosis

This occurs most commonly in black African people and is of unknown aetiology. It presents with small (2–4 mm) asymptomatic porcelain-white macules, often on skin exposed to sunlight. The borders are often sharply defined and angular. There is no effective treatment.

Leprosy (see also p. 140)

Both tuberculoid leprosy and indeterminate leprosy can present with anaesthetic patches of depigmentation and should always be considered in people from endemic regions. Loss of hair and decreased sweating may also be present in the lesions.

HYPERPIGMENTATION

Freckles (ephelides)

These appear in childhood as small brown macules after sun exposure. They fade in the winter months.

Lentigos

These are a more permanent macule of pigmentation similar to freckles but they tend to persist in the winter. Solar lentigos (also called 'liver spots') occur in older people on exposed skin due to actinic damage.

Chloasma (melasma)

These are brown macules often seen symmetrically over the cheeks and forehead and are most common in women. Chloasma can occur spontaneously but it is also associated with pregnancy and the oral contraceptive pill.

Metabolic/endocrine effects

A generalized skin darkening can occur with chronic liver disease, especially haemochromatosis. It also is seen sometimes in Cushing's syndrome, Addison's disease (more marked in palmar creases and buccal mucosa) and Nelson's syndrome.

Peutz–Jeghers syndrome (p. 285)

This is an autosomal dominant condition which presents with brown macules of the lips and perioral region. It is associated with gastrointestinal polyposis.

Urticaria pigmentosa (cutaneous mastocytosis)

This presents most commonly with multiple pigmented macules in children. These lesions tend to become red, itchy and urticated if they are rubbed (Darier's sign). Occasionally lesions may blister, and in the rare congenital, diffuse form of the disease the skin may become thickened and leathery. Occasionally systemic symptoms are present such as wheeze, flushing, syncope or diarrhea, reflecting extensive mast cell degranulation from the skin. Anaphylaxis occurs very rarely and may be precipitated by mast cell degranulators such as aspirin or opiates. The condition spontaneously resolves after some years in children but is persistent in adults.

Skin biopsy shows an excess of mast cells in the skin. A mutation in the proto-oncogene *c-kit* often underlies the condition resulting in mast cell proliferation and mast cell apoptosis. Monoclonal antibodies against mast cell markers (tryptase and CD117) help confirm the cell type.

Rarely there may be infiltration of internal organs with mast cells (*systemic mastocytosis*), especially in adult disease or neonatal disease. This can involve any organ but especially the bone (where it can cause severe pain), gastrointestinal tract, liver and spleen. There is a small risk of developing leukaemia if the bone marrow is heavily infiltrated.

Treatment of the skin, if required, is with antihistamines, sodium cromoglicate or PUVA.

Other conditions with pigmentation

Café-au-lait macules are seen in neurofibromatosis types 1 and 2, tuberous sclerosis, ataxia telangiectasia, Fanconi's anaemia, multiple endocrine neoplasia type 1 and McCune–Albright syndrome.

Multiple lentigines are seen in Peutz–Jeghers syndrome, xeroderma pigmentosum, LEOPARD syndrome (**L**entigines, **E**lectrocardiographic conduction abnormalities, **O**cular hypertelorism, **P**ulmonary stenosis, **A**bnormalities of genitalia, **R**etardation of growth, **D**eafness) and Carney complex (lentiginosis, myxomas of heart and skin, endocrine overactivity.)

Acquired melanocytic naevi are seen in Turner's syndrome (p. 1007) and atypical mole syndrome (dysplastic naevus syndrome, p. 1258).

DRUG-INDUCED RASHES

Drugs can be toxic and teratogenic but they can also cause problems through allergic reactions. This frequently presents in the skin where just about any type of skin rash can arise (Table 23.15) although a widespread symmetrical maculopapular rash is the most common type (Fig. 23.38). 'Fixed drug eruptions' may occur where a rash evolves and resolves at a specific site. The rash is reproduced at exactly the same site after a repeated exposure.

The history is of great value in assessing drug reactions as a drug cause is very common. The use of prick testing and patch-testing is rarely helpful and not without risk. Drug allergy can only be proven by rechallenging but this is rarely justified as it carries some risks. Rechallenging is occasionally justified for antituberculosis drugs or antiretroviral drugs but this should be carried out under medical supervision as there is a risk of anaphylaxis. Certain individuals (e.g. those with HIV infections) are more susceptible to drug rashes (Fig. 23.39).

Table 23.15 Morphological types of drug rashes and some common causes

Maculopapular	Penicillin/amoxicillin
Urticaria	Penicillin, aspirin
Vasculitis	Gold, hydralazine
Fixed drug rash	Phenolphthalein in laxatives, tetracyclines, paracetamol
Pigmentation	Minocycline (black), amiodarone (slate grey)
Lupus erythematosus	Penicillamine, isoniazid
Photosensitivity	Thiazides, chlorpromazine, sulphonamide, amiodarone
Pustular	Carbamazepine
Erythema nodosum	Sulphonamide, oral contraceptive
Erythema multiforme	Barbiturates
Acneiform	Corticosteroids
Lichenoid	Chloroquine, thiazides, gold, allopurinol
Psoriasiform	Methyldopa, gold, lithium, beta-blockers
Toxic epidermal necrolysis	Penicillin, co-trimoxazole, carbamazepine, NSAIDs
Pemphigus	Penicillamine, ACE inhibitors
Erythroderma	Gold, sulfonylureas, allopurinol

Fig. 23.38 **Morbilliform drug rash due to penicillin allergy.**

FURTHER READING

Jin Y, Mailloux CM, Gowan K et al. NALP1 in vitiligo-associated multiple autoimmune disease. *New England Journal of Medicine* 2007; **356**: 1216–1225.

Taïcb A, Picardo M. Vitiligo. *New England Journal of Medicine* 2009; **360**: 160–169.

Whitton ME, Ashcroft DM, Barrett CW et al. Interventions for vitiligo. *Cochrane Database of Systematic Reviews* 2006; **25**: CD003263.

Fig. 23.39 **Erythema nodosum in a patient with HIV on co-trimoxazole.**

Most rashes will settle spontaneously once the offending agent is removed. The commonest culprits in a hospital setting are antibiotics and chemotherapy agents. The three most serious types of drug rashes are:

- erythroderma (p. 1249)
- toxic epidermal necrolysis
- drug-induced hypersensitivity syndrome.

Toxic epidermal necrolysis is characterized by widespread subepidermal blistering and sloughing of more than 30% of the skin and a high mortality (30–50%). A prodrome of cough, myalgia and anorexia may precede skin signs by 2–3 days. The skin may be itchy but typically takes on a burning quality. Fever and mucosal involvement are common. The internal epithelial surfaces (lung, bladder, gastrointestinal tract) are also involved. Multiorgan failure and sepsis often occur. Toxic epidermal necrolysis can be fatal even after drug withdrawal and intensive care support. Patients should be managed in intensive care or a specialized burns unit. All drugs where possible should be withdrawn. Occlusive cutaneous dressings significantly reduce the pain, and an ophthalmological assessment and oral hygiene are necessary. Specific medical treatment with steroids or ciclosporin is controversial and rarely used. Intravenous immunoglobulin may be beneficial if given early in the disease and this is more commonly used.

A variant exists called *Stevens–Johnson syndrome* where the damage is restricted to the mucosal surfaces with milder bullous involvement of the skin (<10%). Both SJS and TEN are commoner in slow acetylators and very much commoner in those with HIV infection.

Drug-induced hypersensitivity syndrome (DHS, also called '*anticonvulsant hypersensitivity syndrome*', 'drug reaction with eosinophilia and systemic symptoms') is a serious adverse systemic reaction to a drug typically occurring 3–6 weeks after initial exposure. It is characterized by a generalized mucocutaneous rash, fever and lymphadenopathy with variable arthralgia, pharyngitis, periorbital oedema and hepatosplenomegaly. Rarely pustulation of the skin and conjunctivitis are present. The blood may show a peripheral eosinophilia, lymphocytosis with atypical lymphocytes, and a hepatitic picture. It can progress to multiorgan failure.

The cause of DHS is unknown but it may in part be due to genetic deficiency of epoxide hydrolase, a hepatic enzyme that detoxifies the arene oxide metabolites of antiepileptic drugs. There have been a number of reports about reactivation of various herpes family viruses (HHV 6 and 7, CMV, EBV) associating with DHS, the significance of which remains unclear. DHS can occur with any drug but a common culprit is one of the aromatic anticonvulsants (carbamazepine, phenytoin, phenobarbitone, primidone and clonazepam) and as they can cross-react all these drugs must be avoided in the future. Sodium valproate is a suitable alternative. There is also potential for cross-reaction with the newer anticonvulsants vigabatrin and lamotrigine. Treatment is with drug withdrawal, systemic steroids and supportive care.

FURTHER READING

Hussain W, Craven NM. Toxic epidermal necrolysis and Stevens-Johnson syndrome. *Clinical Medicine* 2005; **5**: 555–558.

Wolf R, Orion E, Marcos B et al. Life-threatening acute adverse cutaneous drug reactions. *Clinics in Dermatology* 2005; **23**: 171–181.

DISORDERS OF NAILS

Psoriasis and fungal nail infection are the commonest causes of nail dystrophy and are discussed on pages 1240 and 1232.

- *Nail pitting* can be caused by psoriasis, alopecia areata and atopic eczema. A few pits can be present due to trauma.
- *Onycholysis* (distal nail plate separation) is caused by psoriasis, thyrotoxicosis, following trauma and rarely due to a photosensitive reaction to drugs such as tetracyclines.
- *Koilonychia* (thin spoon-shaped nails) can be caused by iron deficiency anaemia or rarely be congenital.
- *Leuconychia* (white nails) is seen in hypoalbuminaemia. A striate congenital leuconychia exists.
- *Beau's lines* (transverse lines) appear as solitary depressions which grow out slowly over many months. They arise due to a severe illness or shock which causes a temporary arrest in nail growth.
- *Yellow-nail syndrome* is a rare disorder of lymphatic drainage. It presents with thickened, slow-growing, yellow nails which may be associated with pleural effusions, bronchiectasis and lymphoedema of the legs.
- *Onychogryphosis* is a gross thickening of the nail which is seen in later life, especially in the big toenail. There is often a history of preceding trauma. Both psoriasis and fungal infection can also cause nail thickening.
- *Nail–patella syndrome* is an autosomal dominant condition which presents with triangular rather than half-moon-shaped lunulae, especially of the thumb and forefingers. The nail plates may be small or dystrophic. The patellae are hypoplastic or absent. Other skeletal anomalies may be present and renal impairment (glomerulonephritis) occurs in up to 30% of individuals.
- *Melanonychia* (longitudinal brown streaks) may be seen as a normal variant in black-skinned patients. In a white patient it may reflect an underlying subungual melanoma, especially if the pigmentation progresses proximally onto the nail fold ('Hutchinson's sign').
- *Clubbing* is discussed on page 820.

FURTHER READING

Numerous authors and articles on nail disease. *Dermatologic Clinics* 2006; **24**: 297–399.

DISORDERS OF HAIR

HAIR LOSS

Hair loss can be due to a disorder of the hair follicle in which the scalp skin looks normal (non-scarring alopecia) or due to a disorder within the scalp skin that causes permanent loss

Table 23.16 Causes of alopecia

Scarring alopecia	Non-scarring alopecia
Discoid lupus erythematosus	Androgenic alopecia
Kerion (tinea capitis)	Telogen effluvium
Lichen planus	Alopecia areata
Dissecting cellulitis	Trichotillomania (self-induced hair-pulling)
X-irradiation	Tinea capitis
Idiopathic ('pseudopelade')	Traction alopecia
	Metabolic (iron deficiency, hypothyroidism)
	Drug (e.g. heparin, isotretinoin, chemotherapy)

of the follicle (scarring or cicatricial alopecia). This latter form causes shiny atrophic bald areas in the scalp which are devoid of follicular openings. There are many causes of alopecia (Table 23.16).

Androgenic alopecia

Androgenic alopecia (male pattern baldness) is the most common type of non-scarring hair loss and depends on genetic factors and an abnormal sensitivity to androgens. It presents in young men with frontal receding followed by thinning of the crown, and there is often a positive family history. It also occurs in females but tends to occur at a later age, be milder and show little in the way of frontal recession. If acne and menstrual disturbance are also present, the cause may be polycystic ovary syndrome or another endocrine disorder of androgens.

Treatment may not be required. Topical 5% minoxidil lotion or oral finasteride (1 mg daily) can help arrest progression and may cause a small amount of regrowth, providing it is used early in disease, but the treatment needs to be continued possibly lifelong. Approximately one-third of patients will not respond to either therapy. Finasteride is a selective inhibitor of 5α-reductase type II and it can cause side-effects in 1% of patients such as loss of libido. It should not be used in females as it can affect the sexual development of a male fetus. However, antiandrogen therapy (e.g. cyproterone acetate or spironolactone) may help some women.

Alopecia areata

Alopecia areata may be regarded as an immune-mediated type of hair loss associated with other organ-specific autoimmune diseases. It presents in children or young adults with patches of baldness. These may regrow to be followed by new patches of hair loss. The presence of broken exclamation mark hairs (narrow at the scalp and wider and more pigmented at the tip) at the edge of a bald area is diagnostic. Regrowth may initially be with white hairs and often occurs slowly over months. Occasionally all of the scalp hair is lost (alopecia totalis) and rarely all body hair is lost (alopecia universalis). The nails may be pitted or roughened.

Treatment has no effect on the long-term progression. Potent topical or injected steroids may be of limited use. Topical immunotherapy with diphencyprone, PUVA or topical 5% minoxidil is occasionally tried but often does not help. Wigs can be provided for severe cases and patient support groups are often beneficial.

Other disorders

Traction alopecia

This refers to the 'mechanical damage' type of hair loss that arises from pulling the hair back into a bun or tight plaiting. It is more common in black Africans.

Telogen effluvium

Telogen effluvium refers to the pattern of diffuse hair loss that occurs some 3 months after pregnancy or a severe illness. It occurs because 'stress' puts all the hairs into the telogen phase of hair shedding at the same time. The hair fully recovers and the normal staggered hair growth/hair shedding cycle resumes within a few months.

Dissecting cellulitis

This is a chronic folliculitis affecting predominantly young black males. It presents with papules and pustules over the occipital region of the scalp with hair loss. If severe, the back of the scalp becomes a boggy swelling (discharging pus) with areas of scarring alopecia. It can be complicated by keloid scar formation ('acne keloidalis nuchae').

Treatment is difficult but prolonged courses of low-dose antibiotics are worth trying in early disease. Prolonged courses of isotretinoin can help a few individuals and deep surgical excision can be used in recalcitrant cases.

INCREASED HAIR GROWTH

Hirsutism (p. 1003)

Hirsutism refers to the male pattern of hair growth seen in females. The racial variation in hair growth must be considered. Certain races (e.g. Mediterranean and Asian) have more male pattern hair growth than northern European females. This is not due to excess androgens but may reflect a genetically determined altered sensitivity to them. If virilizing features (deep voice, clitoromegaly, dysmenorrhoea, acne) are present, a full endocrine assessment is necessary. Hirsutism can cause severe psychological distress in some individuals.

Treatment involves physical methods such as bleaching, waxing, electrolysis, photoepilation and laser therapy. Antiandrogen therapy is occasionally helpful.

Hypertrichosis

Hypertrichosis refers to the state of excessive hair growth at any site and occurs in both sexes. It can be seen in anorexia nervosa, porphyria cutanea tarda and underlying malignancy, and is caused by certain drugs (e.g. ciclosporin, minoxidil).

BIRTH MARKS/NEONATAL RASHES

Strawberry naevus (cavernous haemangioma)

Strawberry naevus affects up to 1% of infants. It presents at, or shortly after birth as a single red lumpy nodule (Fig. 23.40) that grows rapidly for the first few months. Multiple lesions

FURTHER READING

Haedersdal M, Gøtzsche PC. Laser and photoepilation for unwanted hair growth. *Cochrane Database of Systematic Reviews* 2006; **18**: CD004684.

Wasserman D, Guzman-Sanchez DA, Scott K et al. Alopecia areata. *International Journal of Dermatology* 2007; **46**: 121–131.

Fig. 23.40 **Strawberry naevus (cavernous haemangioma).**

can be present. They will spontaneously resolve with good cosmesis but may take up to 7 years for complete resolution. Occasionally plastic surgery is needed after resolution to remove residual slack skin. Reassurance of parents is usually all that is required.

Treatment is indicated if:

- the lesion interferes with feeding or vision
- the lesion ulcerates or bleeds frequently
- the lesion is associated with high-output cardiac failure from shunting of large volumes of blood
- the lesion consumes platelets and/or clotting factors causing potentially life-threatening haemorrhage ('Kasabach–Merritt syndrome').

The last two complications are very rare and only tend to occur in large lesions with significant deep vessel involvement.

Treatment modalities include intralesional or oral corticosteroids, surgery (for selected lesions), and tunable dye laser (for treating ulceration). Alpha interferon injections, vincristine or embolization are only used for life-threatening events.

Port-wine stain (naevus flammeus)

Port-wine stain is also called a capillary haemangioma but strictly speaking it is not a haemangioma, just an abnormal dilation of dermal capillaries. It presents at birth as a flat red macular area and is commonly found on the face. It does not improve spontaneously and it may become thickened with time. If the lesion is found in the distribution of the first division of the trigeminal nerve it may be associated with ipsilateral meningeal vascular anomalies that can cause epilepsy and even hemiplegia (Sturge–Weber syndrome, p. 1169). If a port-wine stain involves the skin near the eye, glaucoma is a risk and ophthalmic assessment is mandatory.

Treatment of port-wine stains is ideally with the tunable dye (pulsed dye) laser. Facial lesions respond best but lesions can redarken after some years.

Other conditions in the newborn

Milia

'Milk spots' are small follicular epidermal cysts. They are small pinhead white papules commonly found on the face of infants. They resolve spontaneously, unlike in adults.

Mongolian blue spot

This appears in infants as a deep blue-grey bruise-like area, usually over the sacrum or back, and is occasionally mistaken as a sign of child abuse. It is due to deep dermal melanocytes. It is very common in Oriental children, less common in black Africans and rare in Caucasians. It has usually disappeared by the age of 7 years.

Toxic erythema of the newborn (erythema neonatorum)

Toxic erythema of the newborn is a term used to describe a common transient blotchy maculopapular rash in newborns. The rash is occasionally pustular but the child is not toxic or unwell. It disappears spontaneously within a few days.

Nappy rash ('diaper dermatitis')

This is an irritant eczema caused by occlusion of faeces and urine against the skin. It is almost universal in babies. The flexures are usually spared, which is a useful differentiating feature from seborrhoeic and atopic eczema. If satellite lesions are present around the edge it may indicate a superimposed *Candida* infection. This rash can also occur in the elderly incontinent.

Treatment involves frequent changing of the nappy and regular application of a barrier cream.

Acrodermatitis enteropathica (p. 225)

This is due to a rare inherited deficiency of zinc absorption. It presents 4–6 weeks after weaning or earlier in bottle-fed babies. There is an erythematous, sometimes blistering, rash around the perineum, mouth, hands and feet. It may be associated with photophobia, diarrhoea and alopecia.

Treatment is with lifelong oral zinc, which seems to override the poor absorption. The response is rapid.

HUMAN IMMUNODEFICIENCY VIRUS AND THE SKIN (p. 188)

HIV infection commonly causes significant dermatological problems. A rash may even be the presenting feature of underlying HIV infection. It is estimated that 90% of HIV-positive patients will suffer with a mucocutaneous disorder during the illness and that up to 30% of people with AIDS will suffer from three different dermatoses. These rashes can often be clinically atypical and difficult to diagnose, and skin biopsy and skin culture is sometimes required for diagnosis. Many of the skin problems are resistant to standard treatments, but HAART (p. 199) has decreased the prevalence.

Cutaneous infection and opportunistic infection are increased due to HIV-induced immune deficiency. *Molluscum contagiosum* is particularly common, especially on the face. Lesions are often multiple and of a 'giant' size measuring over 1 cm across. Molluscum is rarely seen in immunocompetent adults but can be the presenting feature of HIV. Other viral infections such as extensive ulcerative herpes or widespread viral warts may be seen. Bacterial infections (e.g. staphylococcal boils) and fungal infections (e.g. ringworm and *Candida*) are

FURTHER READING

Harper J, Oranje A, Prose N (eds). *Textbook of Pediatric Dermatology,* 2nd edn. Oxford: Blackwell Scientific, 2005.

also common. Recalcitrant and recurrent oropharyngeal candidiasis is a particular problem.

Opportunistic infections such as cutaneous cytomegalovirus (pustules or necrotic ulcers), sporotrichosis (linear nodules) or *Cryptococcus* (red papules, psoriasiform or molluscum-like lesions) can pose diagnostic difficulties, stressing the need for skin biopsy and culture.

Inflammatory dermatoses show an increased incidence with HIV infection, probably due to an immune dysfunction or imbalance rather than as a consequence of immune suppression. Severe, *extensive seborrhoeic eczema* is very common and may be a presenting sign of HIV. Other types of eczema, psoriasis, ichthyosis (dry scaly skin), nodular prurigo and pruritus are all common in HIV infection and can be very severe. *Granuloma annulare* and *lichen planus* are probably increased in incidence. The treatment of these conditions can be difficult in patients with low CD4 counts (<200 mm^2), and oral immunosuppressive therapies (e.g. prednisolone, ciclosporin) are best avoided. Topical therapies and phototherapy seem relatively safe and oral retinoids are a safe option for psoriasis.

'Autoimmune dermatoses' such as bullous pemphigoid, thrombocytopenic purpura and vitiligo seem to be increased in incidence. The aetiology is related to polyclonal stimulation of B lymphocytes by HIV with a resulting abnormal antibody production. *Erythroderma* is sometimes seen in HIV disease where skin biopsy suggests a 'graft-versus-host disease' mechanism. This presumably reflects a severe underlying immune dysfunction of T lymphocyte control.

Adverse drug rashes are much commoner in HIV patients. Reactions to co-trimoxazole, dapsone (used in *Pneumocystis* prophylaxis) and antiretroviral drugs appear particularly common. Drug rashes may be severe (especially with nevirapine and efavirenz), resulting in erythroderma or toxic epidermal necrolysis. Other unusual rashes include a striking nail/mucosal pigmentation from zidovudine, paronychia from indinavir and facial lipodystrophy mostly from protease inhibitors.

Cutaneous tumours, e.g. Kaposi are sarcoma (p. 204), are much commoner in homosexuals with HIV than other groups. Basal and squamous cell carcinomas and benign melanocytic naevi are also a little increased in incidence, presumably reflecting a loss of immune surveillance.

'Specific' HIV dermatoses

'Itchy folliculitis' of HIV

Itchy folliculitis (also called papular pruritic eruption) is common in HIV as CD4 counts decline. The previously described staphylococcal folliculitis, eosinophilic folliculitis, pityrosporum folliculitis, and demodex mite folliculitis are probably all part of a spectrum and the term 'itchy folliculitis' is useful to encompass these. It presents with intensely itchy papules centred on hair follicles and occurring most commonly over the upper trunk and upper arms. The face is more commonly involved in black patients. Individual lesions frequently have the top scratched off, leaving a crateriform appearance. The aetiology is unknown but may reflect a hypersensitivity reaction as high IgE and eosinophil counts may be present.

Treatment with oral minocycline, potent topical steroids and emollients may help. Phototherapy or oral isotretinoin is useful in resistant cases.

Oral hairy leucoplakia

This is characterized by white plaques with vertical ridging on the sides of the tongue. Unlike with oral *Candida* the lesions cannot be peeled off to leave raw areas underneath. It was first recognized in HIV disease but can rarely occur in other forms of immunosuppression. It is thought to be due to co-infection with Epstein–Barr virus.

Treatment with aciclovir, ganciclovir or foscarnet may help.

DERMATOSES OF PREGNANCY

There are a number of minor skin changes during pregnancy. There is an increase in spider naevi, melanocytic naevi, skin tags and chloasma. The abdomen shows mid-line pigmentation (linea nigra) and striae (stretch marks). There are four less common skin problems associated with pregnancy.

Polymorphic eruption of pregnancy (PEP, also called 'pruritic urticarial papules and plaques of pregnancy')

This rash tends to appear in the last trimester of a first pregnancy in 1 in 160 cases. It is of unknown aetiology and recurs only rarely in subsequent pregnancies. It presents with very itchy urticated papules and plaques and occasionally small vesicles. Lesions usually start on the abdomen and striae but may spread to the upper arms and thighs. The umbilicus may be spared. PEP is commoner in twin pregnancies. The rash is not associated with any maternal or fetal risk. PEP has recently been shown to be associated with low maternal serum cortisol levels.

Treatment is with reassurance, bland emollients and mild topical steroids. The rash disappears after childbirth.

Prurigo of pregnancy

Prurigo of pregnancy affects 1 in 300 pregnancies. It usually starts on the abdomen in the third trimester but may persist for some months after delivery. Clustered excoriated papules (prurigo-like lesions) occur on the abdomen and extensor surfaces of the limbs. The cause is unknown but pregnancy related itch (pruritus gravidarum) may be due to cholestasis (p. 362). Rarely liver biochemical tests are abnormal and urinary HCG levels may be elevated. It can recur in subsequent pregnancies. Some authors believe the condition is associated with an increase in fetal mortality but this remains controversial.

Treatment is with topical steroids and oral antihistamines.

Pruritic folliculitis of pregnancy

This occurs in the second or third trimester of pregnancy and is characterized by an itchy folliculitis which looks similar to steroid-induced 'acne'. It is not associated with any increased maternal or fetal risk.

Treatment with topical benzoyl peroxide and hydrocortisone cream helps to relieve symptoms.

FURTHER READING

Colven R (ed.). Update on HIV/AIDS. *Dermatologic Clinics* 2006; **24**: 407–578.

Pemphigoid gestationis (herpes gestationis)

This is the rarest of the pregnancy-related rashes (1 in 60 000). The immune changes of pregnancy appear to set off bullous pemphigoid. It is characterized by an itchy blistering urticated eruption that starts on the abdomen but may become widespread. Large bullae may be present. Unlike PEP it can occur early, starting in the second or even first trimester of pregnancy, and the umbilicus is often involved. It tends to recur in subsequent pregnancies and at an earlier stage. Diagnosis is confirmed by immunofluorescence studies.

A transient bullous eruption occurs in 5% of infants, presumably due to transplacental passage of the offending antibody. There is no increase in fetal mortality but there is an increased incidence of prematurity and low birth weight which is probably due to the autoantibody causing placental insufficiency.

Treatment of mild cases is with potent topical steroids but most cases will require oral corticosteroids. The steroid dose may need to be increased after delivery as there is often a postpartum flare-up of the disease. The rash can be set off again by the oral contraceptive pill and this should be avoided.

FURTHER READING

Ambros-Rudolph CM, Müllegger RR, Vaughan-Jones SA et al. The specific dermatoses of pregnancy revisited and reclassified: results of a retrospective two-center study on 505 pregnant patients. *Journal of the American Academy of Dermatology* 2006; **54**: 395–404.

PRINCIPLES OF TOPICAL THERAPY

Dermatology is unique in having such direct accessibility to the affected organ. This allows the use of topical treatments which can avoid certain systemic side-effects.

A topical therapy consists of an *active ingredient*, an appropriate *vehicle* or *base* to deliver this, and often a *preservative* or *stabilizer* to maintain the product's shelf-life. A cosmetically acceptable product is necessary and must be accompanied by instructions to the patients about correct usage as without this compliance tends to be poor. Perfumed or scented products should be avoided.

Bases and their uses

- *Creams.* A cream is a semisolid mixture of oil and water held together by an emulsifying agent. They need to have added preservatives such as parabens. They are 'lighter' and rub in more easily than ointments. They have a high cosmetic acceptability and are useful for topical treatments of the face and hands. Aqueous cream is a particularly useful as a soap substitute.
- *Ointments.* These are semisolid and contain no water, being based usually on oils or greases such as polyethylene glycol (water soluble) or paraffin (fatty). They feel greasy or sticky to the touch. They are the best treatment for dry, flaky skin disorders as they are good at hydrating the stratum corneum and they deliver an active ingredient (e.g. a steroid) more effectively. If patients dislike the greasy nature of ointments a cream is better than no treatment at all, but creams are less effective and do have to be used more frequently. A compromise may be to use a cream on the face and an ointment elsewhere (Table 23.17).

Table 23.17 Emollients commonly used in the UK

Greasy emollients	Lighter creams
Diprobase ointment*	E45 cream*
Oily cream	Diprobase cream*
Unguentum Merck*	Aveeno cream*
50:50 white soft paraffin/liquid paraffin	Aqueous cream

*Trade names

- *Lotions.* These are based on a liquid vehicle such as water or alcohol. They are usually volatile and rapid evaporation promotes a cooling effect on the skin. They are useful for weeping skin conditions and are ideal for use on hairy skin (e.g. the scalp). The cooling effect can be a useful antipruritic. Alcohol-based lotions should be avoided on broken skin as they cause stinging.
- *Gels.* These are semisolid preparations of high molecular weight polymers. They are non-greasy and liquefy on contact with the skin. They are useful for treating hairy skin (e.g. the scalp).
- *Pastes.* Pastes contain a high percentage (>40%) of powder in an ointment base. They are thick and stiff and difficult to remove from the skin. They are useful when a treatment needs to be applied precisely to a skin lesion without it smearing on to surrounding normal skin. An example would be dithranol in Lassar's paste (used on plaques of psoriasis) as dithranol will burn the surrounding normal skin.

Safety of topical steroids

Providing that preparations of appropriate strength are used for the body site being treated, these compounds can be used safely on a long-term intermittent basis (p. 1237). If potent steroids are misused they will cause skin atrophy manifest as striae, wrinkling, fragility and telangiectasia.

Problems with topical therapies

- *Systemic absorption* may occur but only if very large areas of inflamed skin are treated topically and especially if the treatment is occluded with bandages or polyurethane films. Neonates are particularly susceptible to this owing to the relative increase in body surface area to volume.
- *Contact allergy* to topical preparations is not uncommon and may be suspected by unusually resistant disease or by apparent worsening of a condition after application of a substance. It is more common with creams as it is often the result of allergy to the preservative or emulsifying agent. Allergy can also be due to the active ingredient itself (e.g. neomycin or hydrocortisone).
- *Folliculitis* can occur due to blockage of hair follicles. Creams and ointments should be applied to the skin in the same direction as hair growth to try and prevent this blockage. It is a particular problem with the use of ointments in hot weather (especially if under occlusive bandages) and a lighter cream may be more appropriate at this time.

CHAPTER BIBLIOGRAPHY

Burns T, Griffiths C, Breathnach S (eds) et al. *Rook's Textbook of Dermatology,* 7th edn. Oxford: Blackwell Scientific, 2004.

Weedon D (ed.). *Skin Pathology,* 2nd edn. Edinburgh: Churchill Livingstone, 2002.

SIGNIFICANT WEBSITES

http://www.bad.org.uk
British Association of Dermatologists

http://guidance.nice.org.uk/
Dermatology treatment guidelines in UK

http://www.cochrane.co.uk/en/index.htm
The Cochrane Database provides systematic reviews of treatment

http://www.library.nhs.uk/skin
Clinical guidelines, systematic reviews, evidence-based synopses, image database

http://www.aad.org
American Academy of Dermatology

http://tray.dermatology.uiowa.edu/Dermimag.htm
Dermatology images (atlas)

http://www.usc.edu/hsc/nml/index.html
Dermatology images (atlas)

http://dermnetnz.org/
Dermatology information and images

http://www.eczema.org
National Eczema Society

http://www.psoriasis-association.org.uk
Psoriasis Association

http://www.vitiligosociety.org.uk
Vitiligo society

http://www.skin-camouflage.net
British Association of Skin Camouflage

http://www.debra.org.uk
Dystrophic Epidermolysis Bullosa Association

http://www.hairlineinternational.co.uk
Hairline International

This page intentionally left blank

Index

Page numbers in **bold** represent major sections of text. Those in *italics* represent figures, tables and boxes with specific information.

Abacavir, *192*, *195*
  allergic reaction to, 198
Abacteriuric frequency, 602
Abatacept
  Crohn's disease, *290*
  rheumatoid arthritis, *530*, 531
ABC transporter superfamily, *22*
Abciximab
  acute coronary syndromes, *754*
  angina pectoris, 750
  arterial thrombosis, *443*
  myocardial infarction, 758
Abdomen
  acute conditions
    appendicitis *see* Appendicitis
    peritonitis *see* Peritonitis
  clinical examination *see* Abdominal examination
Abdominal CT, 244, *245*
Abdominal distension (bloating), 313
Abdominal examination, 243, *243*
  acute abdomen, 314
  auscultation, 243
  inspection, 243
  palpation, 243, *243*
  percussion, 243
  see also Abdominal X-ray
Abdominal MRI, 244
Abdominal pain, 242–3, 275, 314
  abdominal wall, 243
  acute abdomen, 314
  autosomal-dominant polycystic kidney disease, 642
  biliary, 242
  epigastric *see* Epigastric pain
  gynaecological, 242
  hypercalcaemia, 1020
  lower abdomen, 242–3
  non-specific, *314*
  right hypochondrial, 242
  tricuspid stenosis, 769
  upper abdomen, 242
Abdominal PET, 244
Abdominal ultrasound, 244
Abdominal X-ray, 244
  Addison's disease, 1011
  liver disease, 327
  pancreas, 376
  ulcerative colitis, 291
  urinary tract obstruction, 615
Abducens nerve, 1106
Abetalipoproteinaemia, 284
Abiraterone, 463
  prostate cancer, 492, 646
Abnormal stature, 977–9
  short, 978–9, *979*
  tall, 979
  see also Growth
ABO blood group, 424, *424*
  incompatibility, 417, 427
  sugar chains, *424*
Abreaction, 1197
Abscess
  amoebic, 361
  anorectal, 300
  brain, *86*, 1141, 1159, *1159*
  Brodie's, 541
  hepatomegaly, *331*
  intracranial epidural, 1159
  liver, 360–1
  lung, *86*, 863
  pyogenic, *91*, 360–1
  spinal epidural, 1159–60
  subphrenic, 317
  visceral, with glomerulonephritis, 593
Absidia, 152
ABT-874 (J695), Crohn's disease, *290*
Abuse, 1201
  battery, unlawful, 4
  bruising/bleeding, *see* Trauma
  childhood, 1210
  sexual, 1201
Acalculia, *1099*
Acamprosate, alcohol dependence, 1214
Acanthocytosis, 284
Acanthosis nigricans, *455*, 1251, *1251*
  and diabetes, *1031*
  malignancy-related, *1255*
Acarbose, 1039
Acarus siro, 877
Accelerated junctional tachycardias, *721*
Accessory cells, *see* Antigen-presenting cells (APC)
Accessory nerve, 1110
  lesions, 1110
Acclimatization
  altitude, 956
  heat, 954
ACE inhibitors, 628, 1023
  acute coronary syndromes, *754*
  diabetes mellitus, 1050
  genetic polymorphism in metabolism, *927*
  heart failure, 738, *739*
  hyperaldosteronism, 1024
  hypertension, 804
  Marfan's syndrome, 781
  myocarditis, 788
  nephrotic syndrome, 585
  side-effects, *928*
    cough, 845–6
    endocrine disease, *964*
    rash, *1265*
    respiratory, *877*
  teratogenesis, *924*
Acebutalol, 730
Acetaldehyde, 238, 358
Acetaminophen *see* Paracetamol
Acetazolamide, *657*
  epilepsy, *1143*
  glaucoma, 1091
  hyperuricaemic nephropathy, 606
  mountain sickness, 956
  side-effects, renal calculi, 611
N-Acetylgalactosamine, 424
Acetyl CoA, formation of, 221
N-Acetyl transferase, genetic polymorphism, *926*
Acetylcholine, 1098
Acetylcholine receptor antibodies, measurement, *69*
N-Acetylcysteine, *937*, 948
  contrast nephropathy, 625
  cystic fibrosis, 845
  regimen for, *948*
N-Acetylglucosamine, 122
N-Acetylmuramic acid, 122
Achalasia, 253–4
  clinical features, 253
  complications, 254
  investigations, 253–4
  pathogenesis, 253
  treatment, 254
Achilles tendonitis/bursitis, *515*, 516
Achilles tendonosis, *515*, 516
Achlorhydria, 399, 400
Acholuric jaundice, 329
Achondroplasia, 568
Aciclovir, 100–1, *101*, 107
  encephalitis, 1156
  erythema multiforme, 1250
  genital herpes, 181
  herpes simplex keratitis, 1090
  oral hairy leukoplakia, 1269
  shingles, 1231
  structure, *101*
Acid excretion, 674, *674*
Acid-base balance, 575–6, 672–5, *672*
  acid excretion, 674, *674*
  disturbances, **672–5**
    causes, 674
    diagnosis, 674–5, *675*
    see also Acidosis; Alkalosis
  disturbances in, 893–4, *893*
  Flenley acid-base nomogram, *675*
Acidosis
  correction, 632
  lactic *see* Lactic acidosis
  metabolic *see* Metabolic acidosis
  mixed metabolic, 30
  renal tubular, 676–7, *676*
  respiratory *see* Respiratory acidosis
  uraemic, 677
Acidotic respiration *see* Kussmaul respiration
Acinar cells, 375, *375*
Acinetobacter, *123*, 128
  ventilator-associated pneumonia, 914
Acinus, 320, *320*, 812
Acipimox, hyperlipidaemia, *1066*
Acitretin, acne vulgaris, 1247
Acne fulminans, 1246
Acne vulgaris, 1225, 1245–7
  clinical features, 1246, *1246*
  pathophysiology, *1246*
  polycystic ovary syndrome, 1003
  treatment, 1246–7
Acneiform rash, *1265*
Acoustic neuroma, 1079, 1106, *1114*
Acquired immune response *see* Immune system
Acquired immunodeficiency syndrome *see* AIDS
Acrocyanosis, 417
Acrodermatitis enteropathica, 225, 1268
Acromegaly, 972, *972*, 979–80
  clinical features, 979, *980*
    diabetes, *1031*
  investigations, 979
  rheumatological problems, 556
  treatment, 979–80
Acromioclavicular joint, rheumatoid arthritis, 527
ACTH, 384, *971*, 1098
  Cushing's syndrome, 1013
  hypersecretion, 982
  release, 1008
    in shock, 895
  secretion
    circadian rhythm, 967
    ectopic, *455*
    excess, 972
  stimulation test, *968*, *969*
ACTH stimulation tests *see* Synacthen test
Actin, 404, 501, 682
Actinic keratoses, 1258, *1258*
Actinic reticuloid, 1249
Actinomyces, *1179*
  A. israelii, pneumonia, *858*
Actinomycetes infection, *123*
  joints, 540
Actinomycosis, 149, *1179*
Action potential, *1098*
Activated charcoal *see* Charcoal, activated
Activated partial thromboplastin time (APTT), 433, 435
  DIC, 441
  haemophilia A, *439*
  vitamin K deficiency, *439*
  von Willebrand disease, *439*
Activator protein-1, 65
Active transport, 271
Activin, *970*, *971*
Activity, physical *see* Exercise
Acupuncture, 1118
  osteoarthritis, 522
Acute abdomen, *314*, **314–17**
  causes, *314*
  history, 314
    pain, 314
    vomiting, 314
  investigations, 315
  medical causes, *315*
  physical examination, 314–15
  see also Abdomen, acute conditions
Acute chest syndrome, 410
Acute coronary syndromes (ACS), **752–5**
  biochemical markers, 752
  chest pain, 688
  clinical presentation, 752
  coronary intervention, 754
  electrocardiogram, 752
  investigation and treatment, 753–4, *754*
    antiplatelet agents, 753–4, *754*
    antithrombins, 754–5, *754*
    plaque stabilization/remodelling, 755
  pathophysiology, 752
  prevention, 755
  risk stratification, 752–3, *753*
  see also Angina pectoris, unstable
Acute cortical necrosis, 624
Acute disseminated encephalitis (ADEM), 1157
Acute fatty liver of pregnancy (AFLP), 394

Acute inflammatory demyelinating polyradiculoneuropathy *see* Guillain-Barré syndrome
Acute inflammatory neuropathy *see* Guillain-Barré syndrome
Acute intermittent porphyria, acute abdomen, *315*
Acute kidney injury (AKI) *see* Acute renal failure (ARF)
Acute lung injury (ALI), 917–19
  clinical presentation, 918
  definition and causes, 917, *917*
  management, 918–19, *919*
  pathogenesis and pathophysiology, 917–18
  prognosis, 919
  see also Adult respiratory distress syndrome (ARDS)
Acute lymphoblastic leukaemia (ALL) *see* under Leukaemia
Acute lysis syndrome, 458
Acute myelogenous leukaemia (AML) *see* under Leukaemia
Acute organic reaction *see* Delirium
Acute phase reactants, 898
Acute renal failure (ARF), **619–25**, 910
  classification, 619
  contrast nephropathy, 624–5
  definition, 619
  epidemiology, 619
  hepatorenal syndrome, 625
  hospital-acquired, 619
  postrenal, 620
  prerenal, 619–20
  recovery phase, 624
  renal parenchymal disease, 620–1
  RIFLE criteria, 619, *619*
  treatment, 910
  tubular necrosis *see* Acute tubular necrosis (ATN)
Acute respiratory distress syndrome *see* Adult respiratory distress syndrome (ARDS)
Acute tubular necrosis (ATN)
  causes, *620*, 621
  clinical and biochemical features, 622
  course, 621–2
  investigations, 622
  management, 622–4
  pathogenesis, 621
  and renal transplantation, 640
Acyclovir *see* Aciclovir
Adalimumab
  Crohn's disease, *289*, 290
  rheumatoid arthritis, *530*, 531
ADAM33, in asthma, 847, 852
Adamantanes, 101, *101*
ADAMs, 500
ADAMTS 13, thrombotic thrombocytic purpura, 437
Adaptive immunity, 53, *55*, **62–6**, 818–19
  antigen receptors, 63–4, *63*
  T cell development/activation, 64–6
  see also Immune system
Addison's disease, 247, 663, *965*, **1010–11**
  autoantigens, *74*
  clinical features, 1010, *1011*
  incidence, 1010
  investigations, 1010–11
  pathophysiology and causes, 1010, *1010*
  patient advice, 1011–12
  skin manifestations, 1265
  tests for, 968, *969*
  treatment, 1011–12, *1012*
Adefovir, 101, *101*
  hepatitis B, 339
Adenine, 29, *30*
Adenocarcinoma
  bronchial, 880–1
  gall bladder, 374
  gastric, 262, 263–6
  oesophagus, 256
  small intestine, 284
  of unknown primary, 493, *493*
Adenoma
  adrenal, 1014
  colonic, 300–1
    risk of malignant change, *301*
  hepatic, 364
  mucinous cyst, 384
  serous cyst, 384
Adenoma sebaceum, 1253
Adenosine, *731*
  side-effects, 723
    respiratory, *877*
  supraventricular tachycardia, 722–3
Adenosine deaminase deficiency, 46
Adenosine triphosphate *see* ATP
Adenovirus, 104, *104*
  conjunctivitis, 1088
  pharyngitis, 833
Adenylyl cyclase agonists, autosomal-dominant polycystic kidney disease, 643
Adherent junctions, *25*, 26
Adhesins, 85, 599
  FimH, 599
  PapG, 599
Adhesion, 67
Adhesion molecules, 65, 686, 687
  shock response, 896
Adhesive capsulitis *see* Frozen shoulder
Adipostat, 228
Adjustment disorder, 1209
Adrenal antibodies, 1011
Adrenal gland
  anatomy and function, *971*, 1008
  biochemistry, 1008
  investigations, 1009–10, *1010*
  medulla *see* Adrenal medulla
  physiology, 1008–9, *1009*
  tumours, 1014
Adrenal medulla, 1024–5
  phaeochromocytoma, 222, 1014, 1025
Adrenal steroids, 1008, 1015–16
see also Androgens; Corticosteroids; Glucocorticoids; Mineralocorticoids
Adrenaline (epinephrine), 907, *908*, 1030, 1098
  calcium channel blocker overdose, 941
  cardiac contraction, 684, *684*
  laryngitis, 833
Adrenarche, premature, 996–7
Adrenergic receptors, β-receptors *see* Beta(β)-Adrenoceptors
α-Adrenergic antagonists (blockers) *see* Alpha(α)-blockers
β-Adrenergic antagonists (blockers) *see* Beta(β)-blockers
β-Adrenergic receptors, antagonists/blockers *see* Beta(β)-blockers
Adrenocorticotrophic hormone *see* ACTH
Adult respiratory distress syndrome (ARDS), 917–19, *918*
  clinical presentation, 918
  definition and causes, 917, *917*
  disorders associated with, *917*
  management, 918–19, *919*
  multiorgan failure, 624
  pathogenesis and pathophysiology, 917–18
  prognosis, 919
Advance directive, 6
Advanced cardiac life support, 711, *712*
Adverse drug reactions, *927*, **927–9**
  classification, 928, *928*
  diagnosis, 928–9
  elderly patients, *924*
  management, 929 *see also* specific drugs
Aedes
  A. aegypti, 114
  A. africanus, 114
  A. albopictus, 115
  A. caballus, 116
Aerobactin, 599
Aeromonas hydrophila, skin and soft tissue infections, *124*
Aerophagia, 242, 311
Aerophobia, 120
Affect, 1186–7
  blunting, 1187
Affective disorders *see* Mood (affective) disorders
Affinity maturation, 64
AFP *see* Alpha(α)-fetoprotein (AFP)
African histoplasmosis, 151
African sleeping sickness *see* African trypanosomiasis
African tick typhus, *147*
African trypanosomiasis, *82*, 157–8
  clinical features, 157
  control, 158
  diagnosis, 157–8
  epidemiology, 157
  parasitology, 157
  treatment, 158
After depolarizations, 716
Afterload, 892, *892*
  increased, 734
Age, and coronary artery disease, 745
Age-related macular degeneration, 1093–4, *1094*
  treatment, 1094
Ageing
  cellular theories, 226–7
  early origins of health and disease, 227
  evolutionary theories, 227
  molecular theories, 226
  and nutrition, 226–8, *227*
  postural hypotension, *659*
  process, 226–7
  system theories, 227
  see also Elderly
Aggrecan, 500
Agoraphobia, 1208
Agouti-related protein, 229, *229*
Agraphaesthesia, *1099*
Agraphia, 1099, *1099*
AIDS
  defining conditions, *187*
  definition, 187
  gastrointestinal problems, 309
  liver involvement, 362
  musculoskeletal involvement, 541
  see also HIV infection
AIDS-dementia complex, 188, 1158
Air, pollution *see* Air pollution
Air hunger *see* Kussmaul respiration
Air pollution, 830, *830*
  and asthma, 849
Air quality, 830
Air travel, 955
  COPD, 840
  heart failure, 738
Airborne spread of infection, 82
AIRE gene, 74
Airport malaria, 153
Airway hyperresponsiveness, 846, 848
Airways, 709, 815–17
  airflow, 815–16, *816*
    in asthma, 846
  bronchomotor tone, 815
  epithelium, 851
  flow-volume loops, 816–17, *816*
  innervation, 852
  resistance, 815, *824*
  smooth muscle, 852
Akathisia, neuroleptic-induced, 1148
Akinesia, 1112
  Parkinson's disease, 1146
Akinetic-rigid syndromes, 1112, 1145–9
  in children, 1148
ALA synthase, 390, 1073
Alabaster skin, 974
Alanine, 210
Alanine:glyoxylate aminotransferase deficiency, *1070*
Alanine aminotransferase (ALT), 324, 332
Albendazole, *166*
  cysticercosis, 171
  filariasis, *163*
  giardiasis, 162
  hydatid disease, 362
  intestinal helminths, *163*
  loiasis, 165
  strongyloidiasis, 167
  trichinosis, 168
Albinism, 1070
Albumin, 321
  in ascites, *352*
  clinical features, 352
  functions, 321
  half-life, 321
  human albumin solutions, 429
  serum, 324
    myeloma, 485
  small intestinal disease, 276
  see also Hypoalbuminaemia
Albuminuria, diabetes mellitus, 1054
Albuterol *see* Salbutamol
Alclometasone dipropionate, *1237*
Alcohol, **238–9**
  behavioural effects, *1212*
  calorie content, 238
  and cancer, 449
  consumption in pregnancy, 239
  and coronary artery disease, 747
  enzyme induction by, 358
  gastric bleeding, *267*
  gastritis, 263
  and hypertension, 798
  and hypoglycaemia, 1060
  metabolism, 238, *238*, 358
  polyneuropathy, 1174, *1175*
  reducing consumption, 1065
  sensible drinking, *239*
  unit measures, *238*
Alcohol dehydrogenase, 238
Alcohol misuse/dependence, 238–9, *1114*, 1211–14
  alcohol dependence syndrome, 1212–13, *1212*
  causes, 1213
  clinical consequences
    chronic pancreatitis, 380, 381
    cirrhosis *see* Alcoholic cirrhosis
    megaloblastic anaemia, 402
  detection, 1212, *1212*
  diagnostic markers, 1212
  epidemiology, 1212, *1212*
  and epilepsy, 1141
  heavy drinkers, 1211
  physical effects, *239*
  problem drinkers, 1211
  symptoms, *1213*
  treatment, 1213–14
    drugs, 1213–14

outcome, 1214
psychological, 1213
Alcohol poisoning, 943, *943*
clinical features, 943, *943*
metabolic acidosis, *939*
treatment, 943
Alcohol withdrawal, 1213–14
see also Delirium tremens (DTs)
Alcoholic cirrhosis, 238, 358
clinical features, 359
investigations, 359
management, 359
Alcoholic hepatitis, 358
clinical features, 359
investigations, 359
management, 359
Alcoholic liver disease, 238, 239, 358–9
clinical features, 358–9
investigations, 359
management, 359
pathology, 358
prognosis, 359
transplantation, 348
see also Alcoholic cirrhosis
Alcoholism *see* Alcohol misuse/dependence
Aldosterone, 573, *651*, 652, 1008
Aldosterone:renin ratio, 1024
Aldosterone
Addison's disease, 1011
escape from, 653
hydrogen excretion, *673*
oedematous conditions, 653
potency, *1008*
secretion disorders, 1023–4
see also Conn's syndrome; Hypoaldosteronism
Aldosterone antagonists, *657*
clinical uses, 657
heart failure, 739, *739*
myocardial infarction, 759
Alefacept, psoriasis, 1242
Alemtuzumab, *464*
acute lymphocytic leukaemia, 475
Alendronate, osteoporosis, 563, *564*
Alexia, 1099, *1099*
Alfacalcidol
parathormone suppression, 633
rickets/osteomalacia, 567
Alginates, 252
Alicaforsen
Crohn's disease, *290*
ulcerative colitis, 293
Aliskerin, hypertension, 805, 1023
Alkaline phosphatase (ALP), 324, 332
bone disorders, *559*
bone malignancy, 555
bone-specific, 560
chronic myeloid leukaemia, 473
common bile duct stones, 370
gallstones, 368
hypercalcaemia, 1020
Paget's disease, 565
primary biliary cirrhosis, 355
pyogenic abscess, 361
rickets/osteomalacia, 567
schistosomiasis, 361
Alkalosis
metabolic *see* Metabolic alkalosis
respiratory, *675*
Alkaptonuria, 1070
Alkylating agents, 76, 459–60
All-*trans*-retinoic acid, 471
Allergen avoidance, 832–3
Allergen immunotherapy, 72–3, *73*
Allergen provocation tests, 853
Allergens, 68, 72
and asthma, 850, *850*
provocation tests, skin prick tests *see* Skin-prick tests
Allergic bronchopulmonary aspergillosis, 875–6, *875*
Allergic conjunctivitis, 1089
Allergic disease (type I hypersensitivity), **72–3**
Allergic rhinitis, 72, 819
Allergy/allergic reactions
antiretroviral drugs, 198
and asthma, 847
diagnosis, *69*
eosinophilia, *431*
food *see* Food allergy
restoration of tolerance, 76
skin-prick tests, 829
see also Allergens; Anaphylactic reaction/anaphylaxis; individual diseases
Alloantibodies, 414
Alloantigens, 75
human platelet, 437
Allocation of resources, health authorities, 7–8
Allogeneic stem cell transplantation, 462–3
acute myeloid leukaemia, 470
chronic myeloid leukaemia, 473
follicular lymphoma, 482
non-myeloablative, 463
Alloimmune haemolytic anaemia, 417–18
haemolytic disease of newborn, 417–18
Alloimmunization
pregnancy, 423–4
transfusion-related, 426–7
Allopurinol
cytochrome P450 inhibition, *925*
gout, 538
hyperuricaemic nephropathy, 606
polycythaemia vera, 420
side-effects
acute tubulointerstitial nephritis, *604*
hepatotoxicity, *366*
rash, *1265*
Allylamines, 100, *100*
Almotriptan, migraine, 1137
Alopecia, 1266–7
androgenic, 1267
causes, *1267*
systemic lupus erythematosus, 543
traction, 1267
Alopecia areata, 1267
Alpha(α)-blockers
hypertension, 805
poisoning, antidotes, *937*
side-effects
elderly patients, *924*
postural hypotension, *659*
Alpha$_2$(α_2)-blockers, 1203
Alpha(α)-fetoprotein (AFP), 321, 325, *493*
hepatocellular carcinoma, 363
as tumour marker, *456*
Alphaviruses, *110*, 114
Alport's syndrome, 568, 594
management, 594
Alprostadil, erectile dysfunction, 1000, 1056
Alström syndrome, *1035*
Alteplase
arterial thrombosis, *443*, 444
stroke, *1131*
Altitude, 955–6, *955*
cerebral oedema, 956
hypoxia, 956, *956*
mountain sickness, 956
Aluminium, poisoning, antidotes, *937*
Aluminium hydroxide, and phosphate deficiency, 226
Alveolar echinococcosis, 362
Alveolar haemorrhage, diffuse, 874
Alveolar macrophages, 58, 812
Alveolar oxygen tension, 892, *892*
Alveolar pressure, 816
Alveolar sacs, *812*
Alveolar-arterial oxygen difference, 893
Alveolar-capillary interface, *817*
Alveoli, 812, *812*
perfusion, 817
stability, 817–18
ventilation, 817
Alzheimer's disease, 1073, 1167–8
aetiology, 1168
diagnosis, 48
neurochemical changes, 1168
neuropathology, 1168
Amalgam tattoo, 247
Amanita
A. muscaria, 952
A. pantherina, 952
A. phalloides, 952
hepatotoxicity, *366*
Amantidine, 101, *101*
multiple sclerosis, 1153
Parkinson's disease, 1148
side-effects, inclusion bodies, 597
Amatoxins, *133*
Amaurosis fugax, 1093, 1128, *1128*
Amblyopia ex anopsia, 1105
Ambulatory blood pressure monitoring, 801–2, *801*
Amenorrhoea, 1001–2
causes, 1001–2
chronic renal failure, 630
hypothyroidism, 1001
polycystic ovary syndrome *see* Polycystic ovary syndrome (PCOS)
Turner's syndrome, 1002
weight-related, 1001
clinical assessment, *1001*
differential diagnosis, *1002*
hypothalamic, 1001
investigations, 1002, *1002*
primary, *995*
secondary, *995*
treatment, 1002
American trypanosomiasis *see* Chagas' disease
Amfetamines
misuse, 1179, 1214–15
overdose, *935*, 939–40
clinical features, 939–40
treatment, 940
side-effects, endocrine disease, *964*
Amikacin, 96, *96*
resistance, 866
tuberculosis, 866
Amiloride, 656, *657*
clinical uses, 657
Conn's syndrome, 1024
hypertension, 804
nephrotic syndrome, 585
Amino acids, 210
degradation, 321
essential, 210
inborn errors of metabolism, 1069–71, *1070*
non-essential, 210
transport defects, 1069
see also Proteins; specific types
Aminoacidurias, 1069–71, *1071*
cystinuria *see* Cystinuria
Fanconi syndrome, 567, 575, 599, 1069–70
Hartnup's disease, 222, 1071
Lowe's syndrome, 1070
Aminobenzoic acid, 1248
Aminoglutethimide
Cushing's syndrome, 1014
side-effects, endocrine disease, *964*
Aminoglycosides, 96, *96*
δ-Aminolaevulinic acid (ALA), 397
Aminosalicylates
Crohn's disease, *289*
ulcerative colitis, 291–2
5-Aminosalicylic acid, ulcerative colitis, 291
Aminotransferases
common bile duct stones, 370
gallstones, 368
pancreatitis, *378*
polymyositis/dermatomyositis, 547
see also individual aminotransferases
Amiodarone, 730, *731*
arrhythmogenic right ventricular cardiomyopathy, 791
side-effects, *928*
endocrine disease, *964*
hepatotoxicity, *366*
long QT syndrome, *727*
neuropathy, *1174*
rash, *1265*
respiratory, *877*
thyroid, 984, 988
supraventricular tachycardia, 723
teratogenesis, *924*
ventricular tachycardia, 726
Amitriptyline
chronic pain syndromes, 1194
chronic widespread pain, 516
genetic polymorphism in metabolism, *926*
migraine, 1137
pain relief, *495*
prostadynia, 603
side-effects, long QT syndrome, *727*
Amlodipine, 941
hypertension, 805
Ammonia, 321, 674
in portosystemic encephalopathy, 353
Ammonium, 674
Amnesia, 1100, 1188
causes, *1100*
transient global, 1128
see also Memory
Amoebiasis, 161
clinical features, 161
management, 161
prevention, 161
returning travellers, *90*
Amoebic abscess, 361
Amorolfine, 100, *100*
ringworm, 1233
Amoxicillin, *94*, 95
bronchitis, 835
COPD, 839
fixed dose prescribing, *925*
gonorrhoea, 176
H. pylori eradication, 261
infective endocarditis, *773*
Lyme arthritis, 540
pneumonia, 862
side-effects, *928*
hepatotoxicity, *366*
rash, *1265*
tonsilitis, 833
urinary tract infections, 602
AMP, 1098
Amphotericin, 100, *100*
cryptococcosis, 151, 200
leishmaniasis, 160
nephrotoxicity, 189
side-effects
acute tubular necrosis, 617
renal tubular acidosis, *676*
sporotrichosis, 152
Ampicillin, 95
gastroenteritis, *132*
infective endocarditis, *773*
side-effects, hepatotoxicity, *366*
tonsilitis, 833
Ampulla of Vater, 320, 375
carcinoma, 383, *383*
Amsacrine, 460
Amylase, 272, 375
serum, 315
acute pancreatitis, 377, *378*
urinary, 377
Amylin, 589, 1034

Amyloid (deposits), 121, 283, 1073
  hepatomegaly, *331*
  see also Amyloidosis; specific proteins
Amyloidosis, 599, 1072–3
  cerebral, 1073
  diagnosis, 1073
  dialysis-related, 638–9, 1073
  familial (transthyretin-associated), 589, 1072
  immunoglobulin-light-chain-related, 589, 1072
  local, 1073
  neuropathy, 1174
  reactive systemic, 1072–3
  renal effects, 528, 588–9
    diagnosis and treatment, 589
    pathology, 588, *589*
    tubular acidosis, *676*
  and rheumatoid arthritis, 526, 528
  secondary, 588
  skin manifestations, 1254
  treatment, 1073
Amyotrophic lateral sclerosis, 1166
Anabolic steroids, side-effects, hepatotoxicity, *366*
Anacetrapib, 1065, *1066*
Anaemia, **392–406**
  Addison's disease, 1011
  aplastic *see* Aplastic anaemia
  autoimmune haemolytic *see* Autoimmune haemolytic anaemias
  bone marrow, 393, *393*
  chronic renal failure, 628
    treatment, 633–4
  classification, *392*
  clinical features, 392–3
  Cooley's, 408
  Crohn's disease, 288
  Fanconi's, 453, *453*
  haemolytic *see* Haemolytic anaemia
  HIV-related, 189
  hypothyroidism, 986
  investigations, 393
  leucoerythroblastic, 430
  macrocytic, *392*, 397–402
    megaloblastic *see* Megaloblastic anaemia
    pernicious *see* Pernicious anaemia (PA)
    without megaloblastic changes, 401–2
  in malaria, 154–5, **154**
  microcytic, *392*, 393–7
    of chronic disease, *396*, 397
    differential diagnosis, *396*
    iron deficiency *see* Iron deficiency anaemia
    sideroblastic *see* Sideroblastic anaemia
  normocytic, *392*, 397
    pyogenic abscess, 361
  pernicious *see* Pernicious anaemia (PA)
  protein-energy malnutrition, 216
  rheumatoid arthritis, 528
  sickle cell *see* Sickle cell anaemia
  sideroblastic, 945
  signs, 392–3
  symptoms, 392
  Wilson's disease, 357
Anaesthetics, volatile, side-effects, *928*
Anagrelide, thrombocythaemia, 421
Anakinra, 76
  familial Mediterranean fever, 556
  rheumatoid arthritis, *530*, 531
  side-effects, respiratory, *877*
Anal fissures, 300
Analgesia
  drugs used *see* Analgesics
  WHO analgesic ladder, 494–5
  see also Analgesics; Pain
Analgesic nephropathy, 605–6
Analgesics, 1118
  acute coronary syndromes, *754*
  adjuvant, 495, *495*
  morphine *see* Morphine
  musculoskeletal problems, 517–18
  NSAIDs *see* NSAIDs
  pain management
    pancreatitis, 379, 381
    see also Analgesia; Pain; specific drugs
Anaphylactic reaction/anaphylaxis *see also* Allergy/allergic reactions
Anaphylactic shock, 897
  clinical features, 899
Anaphylatoxins, complement-derived, 57
Anaphylaxis, 73, *73*
  haemodynamic changes, *899*
  plasma protein incompatibility, 427
  treatment, *73*
Anaplasma phagocytophilum, *148*
Anaplastic lymphoma, *480*
Anarthria, 1146
Anastrozole, 463, 488
Anatomy
  adrenal gland, *971*, 1008
  bladder, 1119
  ear, 1077, *1077*
  extrapyramidal system, 1111, *1111*, 1112–13
  eye, 1085–6, *1085*, *1086*
  heart, 681–8
  hypothalamus, 969–70, *969*
  nose, 1080, *1080*
  pituitary gland, 969–70, *969*
  respiratory tract, 811–14
  stomach, 258
  throat, 1082
  thyroid gland, 982–3
ANCA, 325
  glomerular disease, *585*
  measurement, *69*
  MPO, 596
  musculoskeletal problems, 504
  PR3, 596
  primary sclerosing cholangitis, 373
  vasculitides, 870
ANCA-positive vasculitis, 596–7, *596*
  pathogenesis, 596, *596*
  treatment, 597
Ancylostoma
  A. braziliense, 168
  A. caninum, 168
  A. duodenale, *163*, 167
Anderson disease, *1069*
Androgen(s), 1008
  deficiency, *999*
  effects of, *999*
  excess, *1002*
Androgen deprivation, 463
Androgen receptor blockers, 463
Androgen replacement therapy, *999*
Androstenedione, PCOS, 1004
Anencephaly, 45, 1169
Aneurysms, 807–9
  aortic *see* Aortic aneurysms
  Charcot-Bouchard, 1133–4
  dilatation, 758
  saccular (berry), 1134, *1134*
  thoraco-abdominal, 808, *808*
  see also Aortic dissection
Angelman's syndrome, 44
Angiitis
Angina pectoris, 688, **748–52**
  aortic regurgitation, 767
  aortic stenosis, 765
  cardiac syndrome X, 749
  classical/exertional, 748
  decubitus, 748
  diagnosis, 748, 749
  examination, 749
  intractable, 752
  investigations, 749
  nocturnal, 749
  stable, plaque thrombosis, *745*
  treatment, 749–52
    coronary artery bypass grafting, 751–2, *751*
    percutaneous coronary intervention, 750–1, *751*
    prognostic therapies, 750
    symptomatic, 750
  unstable, 749, 752
    inter-aortic balloon pumping, 714–15
    plaque thrombosis, *745*
    see also Acute coronary syndromes
  variant (Prinzmetal's), 749
Angio-oedema, 1243
  hereditary, 1244
Angioedema, 72
Angiogenesis, 451–2
  regulation of, 687
Angiography
  coeliac axis and mesenteric, 270
  coronary, 706, *706*, *708*
  digital cerebral, 1120
  digital subtraction, 709
  lowe limb ischaemia, *806*
  specific conditions
    liver disease, 327–8
    polyarteritis nodosa, 551
  spine, 1120
Angioneurotic oedema, 237
Angiopoietin-1, 50
Angiopoietin-2, 452
Angiostatin, 50
Angiotensin, 686
Angiotensin I, 576, 1023
Angiotensin II, 576, 895, 1023
  chronic renal failure, 631
  inhibition, 632, *632*
Angiotensin II receptor antagonists, 1023
  heart failure, 738, *739*
  hypertension, 804, 805, 1023
  nephrotic syndrome, 585
Angiotensin-converting enzyme, 41, 576, 1023
  sarcoidosis, 870
  vasomotor control, *686*, 687
Angiotensin-converting enzyme inhibitors *see* ACE inhibitors
Angiotensinogen, 1023
Angular stomatitis, iron deficiency anaemia, 395
Anhedonia, 1186
Anidulafungin, 100, *100*
Animal models
  gene therapy, 47–8
Anion gap, 675, 1049
  high, 677
  normal, 675–7, *675*
  urinary, 677
Anismus, 296
Anisocoria, 1103
Anisocytosis, *394*, 396
Anistreplase, myocardial infarction, 757
Ankle/brachial pressure index, 806
Ankylosing spondylitis, 501, 532–4, 884
  clinical features, 533
  epidemiology, 533
  HLA association, *60*, 532–3
  inheritance, *44*
  investigations, 533–4, *533*
  low back pain, 512
  pathogenesis, 533
  prognosis, 534
  treatment, 534
Ankylostoma, eosinophilia, 874
Ankyrin, 404, 405
Annulus fibrosus, 682
Anogenital warts, 181–2, 1231
Anorchia, *1007*
Anorectal abscess, 300
Anorectal disorders, 299–300, *301*
  anal fissures, 300
  anorectal abscess, 300
  fistula in ano, 300, *300*
  haemorrhoids, 299
  intussusception, 300
  pruritus ani, 299
  rectal prolapse, 300
  solitary rectal ulcer syndrome, 300
Anorexia, 214, 241, 243
  cancer-related, 495
Anorexia nervosa, 1219–20
  aetiology, 1219–20
  prevalence, 1219
  prognosis, 1220
  treatment, 1220
Anosmia, 1082, 1100
Anserine bursitis, 513
Antacids
  side-effects, renal calculi, 611
Antegrade pyelography, 580–1, *580*
  urinary tract obstruction, 615
Antenatal diagnosis *see* Prenatal diagnosis
Anterior descending arteries, 685
Anterior knee pain syndrome, 513
Anterior spinal artery occlusion, 1165
Anterior tibial syndrome, 516
Anterior uveitis *see* Uveitis, anterior
Anthracyclines, 460
Anthraquinones, *297*
Anthrax, *80*, *83*, 142
  as biological agent, *959*, 960
  clinical features, 141
  control, 141
  diagnosis, 141
  gastrointestinal, 141
  inhalational, 141
  management, 141
  rash, *88*
Anti-β_2-glycoprotein I test, 544
Anti-c antibodies, 418
Anti-cardiolipin antibody, 1122
Anti-citrullinated cyclic peptide, rheumatoid arthritis, 523, 524–5
Anti-cyclic citrullinated peptide antibodies, measurement, *69*
Anti-D, 418
  thrombocytic purpura, 436
Anti-D antibodies, 418
Anti-DNA antibodies, *69*
Anti-double-stranded DNA antibodies, 503
Anti-extractable nuclear antigen antibodies, 503
Anti-GBM antibodies
  glomerular disease, *585*
  glomerulonephritis, 595–6, *595*
Anti-inflammatory drugs, 55
  asthma, 855
  bronchiectasis, 844
  musculoskeletal problems, 517–18
  rhinitis, 833
  see also Immunosuppressive agents; specific types
Anti-ischaemia agents, *754*, 755
Anti-K antibodies, 418
Anti-mitochondrial antibody, *324*, 325
  primary biliary cirrhosis, 355
Anti-muscarinics
  asthma, 855
  poisoning, *935*
  side-effects, elderly patients, *924*
Anti-neuronal antibodies, 1122
Anti-neutrophil cytoplasmic antibodies *see* ANCA

Anti-oestrogens, 463
Anti-TNF agents *see* Tumour necrosis factor (TNF) blockers
Antiarrhythmics, 729–31
  class I, 730
  class II, 730
  class III, 730–1
  class IV, 731
  classification, *731*
  mechanisms, *731*
  risk factors, *731*
  see also individual drugs
Antibacterial drugs *see also* Antibiotics
Antibiotics, **94–100**
  aminoglycosides, 96, *96*
  asthma, 856
  bronchiectasis, 843–4
  carbapenems, 95–6
  cephalosporins and cephamycins, 95, *95*, *96*
  chloramphenicol, 97–8, *97*
  COPD, 839
  cytotoxic, 460
  fusidic acid, 98
  glycopeptides, 99
  macrolides, 97
  monobactams, 95
  nitroimidazoles, 99, *99*
  oxazolidinones, 99, *99*, 124
  penicillins, 94–5, *94*
  prophylactic, 92, *93*
  quinolones, 98–9
  specific conditions
    bronchitis, 835
    C. difficile, *132*
    Campylobacter, *132*
    cholera, *132*
    diarrhoea, *132*
    dysentery, *132*
    gastroenteritis, *132*
    myocarditis, 788
    Salmonella, *132*
    Shigella, *132*
    travellers' diarrhoea, *132*
  streptogramins, 124
  sulphonamides and trimethoprim, 98
  tetracyclines, 96–7
  see also individual drugs
Antibodies
  acetylcholine receptor, *69*
  adrenal, 1011
  ANCA *see* ANCA
  anti-c, 418
  anti-cyclic citrullinated peptide, *69*
  anti-D, 418
  anti-DNA, *69*
  anti-GBM, *585*, 595–6, *595*
  anti-K, 418
  anti-mitochondrial, *324*, 325, 355
  antinuclear, 503, *504*
  antiphospholipid, 504
  antithyroid, 984
  blood, 424–5
  congenital deficiency in production, 70
  drug-induced, 414
  endomysial, 279
  hepatitis A, 335
  hepatitis B, *337*
  IgA, 64
  IgG, *73*, 185–6, 415, 425
  IgM, 416, 425
  leucocyte, 427
  lupus anticoagulant, 441
  myositis-specific, 547
  parathyroid, 1022
  parietal cell, 399
  platelet, 436
  red cell, 425
  therapeutic
  see also Immunoglobulin(s); Monoclonal antibodies
Antibody-dependent cellular cytotoxicity, 66
Anticancer drugs, *see* Cancer treatment; Chemotherapy
Anticardiolipin test, 544
Anticholinesterases, myasthenia gravis, 1180
Anticoagulants/anticoagulation, 444–5, *445*
  bleeding, 446
  direct thrombin inhibitors, 445
  fondaparinux, 445
  heparin, 444–5
  length of, 447
  longer-term, 447
  oral, 445–6
    contraindications, 445–6
    indications, *446*
    new drugs, 446
    vitamin K deficiency, 220
    warfarin *see* Warfarin
  outpatient, 447
  side-effects, *928*
  specific conditions, atrial fibrillation, 725
  stroke, 1133, *1133*
Anticonvulsants, *1114*
  drug levels, 1142
  overdose, 940
  side-effects, 1142
    elderly patients, *924*
    endocrine disease, *964*
  teratogenesis, *924*
  see also individual drugs
Anticytokines, 76
Antidepressants, 1201–2
  tricyclic *see* Tricyclic antidepressants (TCAs)
  see also individual types
Antidiabetics, overdose, 940
Antidiuretic hormone (ADH; vasopressin), 573, 652, 970, *971*, 1098
  deficiency *see* Diabetes insipidus
  heart failure, 736
  osmolality response, *1016*
  plasma osmolality, 654
  receptor antagonists, *657*, 658, 663
  receptors, 653
  release
    factors affecting, *1017*
    in shock, 895
  secretion, 654–5
  synthetic *see* Desmopressin
  treatment of shock, 908
  water homeostasis, 653
Antidotes, *937*
Antiepileptic drugs *see* Anticonvulsants
Antifungal drugs, 100, *100*
  allylamines, 100, *100*
  azoles, 100, *101*
  echinocandins, 100, *100*
  polyenes, 100, *100*
Antigen(s)
  degradation, 323
Antigen presentation, 58, 62, *62*
  cross presentation, 62, *62*
  endogenous route, 62, *62*
  exogenous route, 62, *62*
Antigen presenting cells, 62
  functional defects, 71
  interactions with T-cells, *69*
Antigen receptors, 63–4, *63*
antigen-specific (peptide) immunotherapy, 76
Antigenic drift, 117
Antigenic shift, 116
Antihistamines, 72
  atopic eczema, 1238
  see also H_2-receptor antagonists
Antihypertensive drugs, 802–5, *802*
  ACE inhibitors, 804
  alpha(α)-blockers, 805
  angiotensin II receptor antagonists, 804
  beta(β)-blockers, 804
  calcium-channel blockers, 804–5
  centrally acting, 805
  diuretics, 803–4
  renin inhibitors, 805
  stroke, 1132 *see also* specific drug types; Vasodilators
Antileukotriene agents, COPD, 840
Antilymphocyte globulin, 403
  paroxysmal nocturnal haemoglobinuria, 419
Antimalarials, overdose, 940–1
Antimetabolites, 460
Antimicrobial chemotherapy, **91–102**
  antibiotics, **94–100**
    chemoprophylaxis, 92
  antifungal drugs, 100, *100*
  antituberculous drugs, 100
  antiviral drugs, 100–2
  combination therapy, 92
  dose and duration of therapy, 92
  empirical 'blind' therapy, 92
  mechanism of action, 92–4
  pharmacokinetics, 92
  principles of use, 91–2
  renal and hepatic insufficiency, 92
  resistance, 92–4, *93*, *94*
  therapeutic drug monitoring, 92
  see also Antibiotics
Antimucolytic agents, COPD, 839
Antinuclear antibodies, 503, *504*
  glomerular disease, *585*
Antioxidant nutrients, 224
Antioxidants, 223–4
  dietary intake, 224
  randomized controlled trials, 224
  respiratory tract, 818
Antiparkinsonian drugs, 1147–8
  see also specific drugs
Antiphospholipid antibodies, musculoskeletal problems, 504
Antiphospholipid syndrome, 544–6, 599
  clinical features, 545
  pathogenesis, 544–5
  treatment, 545
α_2-Antiplasmin, 433, 434
Antiplatelet drugs
  acute coronary syndromes, 753–4, *754*
  arterial thrombosis, 443, *443*
  stroke, 1132
  see also Anticoagulants/anticoagulation; specific drugs
Antipsychotics *see* Neuroleptics
Antipyretics, 87
Antiretroviral drugs (ARVs), *192–3*, **192–9**
  adherence, 197
  changing therapy, 197
  choice of drugs, 195–6
  co-receptor blockers, *193*, 194
  complications, 197–8
    allergic reactions, 198
    bone metabolism, 198
    IRIS, 198
    lipodystrophy and metabolic syndrome, 198
    mitochondrial toxicity and lactic acidosis, 198
  drug interactions, 197
  drug resistance, 196–7, *196*
  fusion inhibitors, *193*, 194
  integrase inhibitors, *193*, 194
  prescribing, *194*
  protease inhibitors, 192, 194, 196
  reverse transcriptase inhibitors, 192, 195
  side-effects, neuropathy, *1174*
  starting therapy, 194–5, *195*
  stopping therapy, 197
  structured treatment interruptions, 197
  treatment failure, 197
  see also Highly Active Antiretroviral Therapy (HAART); specific drugs
Antistreptolysin-O titre, *585*
  chronic renal failure, 627
Antisynthetase syndrome, 547
Antithrombin agents, acute coronary syndromes, 754–5, *754*
Antithrombin (AT), 433, 897
  deficiency, 442–3
Antithyroid antibodies, 984
Antithyroid drugs, 989, *989*
  see also Carbimazole; Propylthiouracil
α_1-Antitrypsin, 321, 325, 433
  deficiency, 358
    COPD, 837
  replacement, 840
  respiratory tract, 818
Antituberculous chemotherapy, 100
Antitubulin agents, 460
Antivenoms, 951
Antiviral drugs, 100–2
  adamantanes, 101, *101*
  antiretroviral drugs *see* Antiretroviral drugs (ARVs)
  interferon-alpha, *101*, 102
  neuraminidase inhibitors, 101–2, *101*
  nucleoside analogues, 100–1, *101*
  nucleotide analogues, 101, *101*
  palivizumab, *101*, 102
  pyrophosphate analogues, 101, *101*
  ribavirin, *101*, 102
Anton's syndrome, 1102, 1130
Antral gastritis, 259
Anuria, 614
Anus, role in defecation, 295
Anxiety, 1187, *1207*
  communication difficulties for patients, 11
Anxiety disorders, *1190*, 1206–10
  differential diagnosis, *1207*
  generalized, 1206–7, *1207*
  mixed anxiety and depression, 1207
  treatment, 1208–9
    drugs, 1209
    psychological, 1208–9
    see also individual conditions
Aorta
  aneurysms *see* Aortic aneurysms
  ascending
    dilatation/enlargement, 693
    dissection, 693
  calcification, 693
  coarctation *see* Coarctation of aorta
Aortic aneurysms, 807–8
  echocardiography, 703–4
  endovascular stent, 808
  investigations, 807
  laparoscopic surgery, 808
  management, 807–8
  prognosis, 808
  repair, 808
  rupture, 807
  signs, 807
  symptoms, 807
Aortic dissection, 688, 693, 808–9
  echocardiography, 703–4
  investigations, 808
  magnetic resonance imaging, 707
  management, 808–9
  signs, 808
  symptoms, 808
Aortic regurgitation, *765*, 766–7
  causes, 766–7, *767*
  investigations, 767
  magnetic resonance imaging, 707
  pathophysiology, *767*
  signs, 767
  symptoms, 767
  treatment, 767

Aortic stenosis, **765–6**
auscultation, 766
congenital, 765, *774*
echocardiography, 703
investigations, 766
left bundle branch block, *720*
pathophysiology, 765
percutaneous valve replacement, 766
precordial palpation, 766
pulse, 765
symptoms, 765
treatment, 766
valvular, 765, *765*
Aortic valve, 684
Aortography, 581, *581*, 708
APACHE scoring system, 378, *379*, 920
Apaf1, *28*
Apathetic thyrotoxicosis, 988
Apex beat
double pulsation, 692
heaving, 692
left parasternal heave, 692
palpable thrill, 692
tapping, 692
vigorous, 692
Aphasia, 1098, 1099, *1128*
Broca's, 1099, *1099*
global (central), 1100
nominal (anomic, amnestic), 1099
primary progressive, 1168
Wernicke's, 1099, *1099*
Aphthous ulcers/ulceration, 246–7
causes, 247, *247*
major, 247
minor, 247
treatment, 247
Apixaban, 446
arterial thrombosis, *443*
atrial fibrillation, 725
Aplasia
bone marrow, 402
pure red cell, 403
Aplastic anaemia, 325, 402–3
causes, 402, *402*
clinical features, 402
differential diagnosis, 402
investigations, 402, *402*
prognosis, 403
treatment, 403
Aplastic crisis, 109
Apocrine sweat glands, 1227
Apomorphine
erectile dysfunction, 1000
Parkinson's disease, 1148
Apoplexy, pituitary, 975
Apoprotein(s), 1061
see also individual types
Apoprotein A-I, 1062
Apoprotein B, familial defective, *1064*
Apoprotein B-100, 1061, 1062
gene mutations, 1064, *1064*
Apoprotein C, 1062
Apoprotein C-II, 1061
deficiency, 1063
Apoprotein E, 1061, 1062
Apoprotein E2, 1065
Apoptosis, 27–9, *28*, 49, 50, 451
and ageing, 227
disorders, *see* Cancer
heart failure, 735
Appearance, 1186
Appendicectomy, and inflammatory bowel disease, 286
Appendicitis, *314*, 315
clinical features and management, 315
differential diagnosis, 315
treatment, 315
Appendix
infection/inflammation *see* Appendicitis
Appetite
control in obesity, 229–30, *229*
see also Anorexia
Apraxia
constructional, 1100
dressing, *1099*, 1100
gait, *1096*, 1097
limb, *1099*
Aprepitant, 461
Aprotinin, 428
APUD cells, 271, 811
carcinoid tumours, 284–5
Aquaporins, 322, 575, 653–4, *654*
Aquaretics, *657*
clinical uses, 658
Aqueduct of Sylvius, stenosis, 1162
Aqueous cream, *1270*
Aqueous humour, 1085, *1085*
Arabinosides, 460
Arachidonic acid, *210*
metabolism
inhibition of, 849, *849*
products of, 895
Arachis oil, enema, *297*
Arbovirus, 113–14, *114*
arthritis in, 541
Arcuate arteries, 571
Arcuate nucleus, 229, *229*
ARDS *see* Adult respiratory distress syndrome (ARDS)
Arenaviruses, *80*, *110*, *114*, 120–1
Argentinian haemorrhagic fever, *114*
Argyll Robertson pupil, 1104
Aristolochic acid, 606
Arnold-Chiari malformation, *1114*, 1162, 1165, 1169
Aromatase inhibitors, 463, 488
side-effects, 488, *488*
Arrhythmogenic right ventricular cardiomyopathy, *789*, 791
Arsenic
poisoning, antidotes, *937*
side-effects, hepatotoxicity, *366*
Arsenic trioxide, 471
Arsenicals, 158
Arterial blood gases, *675*, *824*, 826, *893*
COPD, 838
idiopathic pulmonary fibrosis, 873
pancreatitis, 379
respiratory failure, 911
Arterial disease, *see* Atherosclerosis; Ischaemia
Arterial inflow occlusion, 283
Arterial oxygen content, 814, 890
Arterial oxygen saturation, 826
Arterial pH, 814
Arterial pulse *see* Pulse
Arterial stenosis, *1128*
Arterial thrombosis, 442
prevention and treatment, 443–4, *443*
antiplatelet drugs, 443, *443*
thrombolytic therapy, 443–4, *443*
risk factors, 442
Arterial ulcers, 1261
Arteriosclerosis
Arteriovenous malformation, 1134
Arteritis
giant-cell *see* Giant-cell arteritis
Takayasu's, *549*, 550
Artesunate-mefloquine, *156*
Arthralgia
infective endocarditis, *770*
sarcoidosis, 869
see also Joint(s), pain
Arthritis, *86*
children, 553–4
crystal, 504, 522, **536–9**
differential diagnosis
distal interphalangeal, 534
elderly patients, *552*
enteropathic, 536
enthesitis-related, 553
gonococcal, 540
infective, *see* Septic arthritis
inflammatory, 522–3
hip, 513
joint involvement, *523*
rheumatoid *see* Rheumatoid arthritis (RA)
wrist and hand, 509
juvenile idiopathic *see* Juvenile idiopathic arthritis
Lyme, 522, 540
meningococcal, 540
mutilans, 534, *534*
osteoarthritis *see* Osteoarthritis
polyarticular *see* Polyarthritis
postviral, 522
psoriatic, 534–5, 1242
reactive, 535–6
rheumatoid *see* Rheumatoid arthritis
septic, 503, 504, 526
syphilitic, 540
tuberculous, 540
viral-associated, 540–1
see also Polyarthritis; specific types/causes of arthritis
Arthropathies
Jaccoud's, 542
pyrophosphate *see* Pseudogout
see also individual conditions
Arthropod disease vectors, *82*
Arthropod ectoparasites, 171–2
local hypersensitivity reactions, 171
resident infections, 171–2
jiggers, 171–2
myiasis, 173
systemic envenoming, 172
Arthropod-borne diseases, 11235
Arthroscopy, 505
osteoarthritis, 522
Artificial respiration, 709–11, *710*
see also Mechanical ventilation
Asbestosis, 879–80, *879*
Ascariasis, 166
Ascaris lumbricoides, *163*, 166
eosinophilia, 874
Aschoff nodules, 137
Ascites, 243, 318, 331, **351–3**, 655
Budd-Chiari syndrome, 360
causes, *352*
chronic pancreatitis, 381
chylous, *352*
haemorrhagic, *352*
investigations, 352
management, 352–3
paracentesis, 353
portal hypertension, 351
serum-ascites albumin gradient, *352*
sodium retention, 351
spontaneous bacterial peritonitis, 353
straw-coloured, *352*
tricuspid stenosis, 768
water retention, 351
Ascorbic acid *see* Vitamin C
Aseptic meningitis, HIV-related, 188
Aseptic necrosis of bone *see* Avascular (aseptic) necrosis of bone
Ash-leaf macules, 1253
Asian sudden unexpected nocturnal death syndrome, *789*
Aspartate, 210, 1098
Aspartate aminotransferase (AST), 324, 332
Aspartate transferase, hypothyroidism, 986
Aspergilloma, 876
Aspergillosis, *150*, 151
invasive, 876
Aspergillus, *90*
A. fumigatus, *70*, *82*, 151, *191*, 848
HIV-associated, 200–1
lung disease, 875–6, *875*
pneumonia, *858*
A. niger, 151
in organ rejection, 75
Aspiration pneumonia, 861–2
Aspirin
acute coronary syndromes, 753, *754*
angina pectoris, 750
anti-thrombotic action, 687, 1050
arterial thrombosis, 443, *443*
atrial fibrillation, 725
fixed dose prescribing, *925*
gastric bleeding, 267
mesangiocapillary glomerulonephritis, 590
metabolism, *949*
myocardial infarction, 757, 759
pericarditis, 795
polycythaemia vera, 420
renovascular disease, 608
resistance, 443
side-effects
haemolysis in G6PD deficiency, *414*
respiratory, *877*
stroke, 1132
see also Salicylates
Asteatotic eczema, 1239
Asterixis, 353, 630
Asthma, 815, **846–57**
acute severe, 856–7
aetiology and pathogenesis, 847–8, *847*
airway hyperresponsiveness, 848
atopy and allergy, 847
environmental factors, 847–8
genetics, 847
and air pollution, 830
attack, 856
catastrophic sudden severe (brittle), 857, *857*
classification, 846–7
clinical features, 852
extrinsic, 847
food allergy, 237
hygiene hypothesis, 72, 847, 1236
immediate, 850, *850*
inheritance, *44*
intrinsic, 847
investigations, 852–3
late phase response, 72, 850, *850*
management, 853–7, *853*, *854*
drug treatment, 853–6, *854*
inhaled therapy, *853*
pathogenesis, 850–2, *851*
inflammation, 850–1
remodelling, 851–2
PEFR, *825*
physical signs, *821*
precipitating factors, 848–50
allergens, 850, *850*
atmospheric pollution, 849
cold air and exercise, 848–9
diet, 849
drugs, 849–50, *849*
emotion, 849
occupational sensitizers, 848, *848*
prevalence, 846
prognosis, 857
prohibited drugs, *923*
Asthmatic bronchopulmonary eosinophilia, 875, *875*
Astigmatism, 1086
Astrocytes, 47
Astrocytoma, 1160
Astrovirus, *110*
Astroviruses, *110*
Asymmetric dimethyl arginine (ADMA), 631

Asymmetric septal hypertrophy, 704
Asymmetrical responses, 1125
Asystole, 711
Ataxia, 1111, 1114, *1128*
  congenital, 1169
  Friedreich's *see* Friedreich's ataxia
  gait, 1111
  sensory, 1097
  truncal, 1111, 1114
Ataxia telangiectasia, 70, 453, *453*, *1114*, 1170
  leukaemia risk, 467
Ataxic respiration, 1124
Atazanavir, *193*, *195*, 196
Atelectasis, 912, 913
Atenolol, 730, *731*
  acute coronary syndromes, *754*
  hypertension, 804
  side-effects, hepatotoxicity, *366*
Atheroembolic renal disease, 609, *609*
Atherosclerosis
  coronary *see* Coronary atherosclerosis
  and diabetes mellitus, 1050
  rheumatoid arthritis, 528
Atherosclerotic plaques
  advanced/raised fibrolipid, 744
  complicated, 744
  stabilization/remodelling, *754*, 755
  thrombosis mechanisms, *744*, *745*
  transitional, 744
  see also Atheroma
Atherosclerotic renovascular disease, 607, 608
Athetoid cerebral palsy, 1149
Athetosis, 1112
Athlete's foot, 1232
Atlizumab, Crohn's disease, *290*
Atopic eczema, 1236–8
  aetiology, 1236
  associated features, 1237
  clinical features, 1236–7, *1236*
  complications, 1237
  exacerbating factors, 1236
  investigations, 1237
  prognosis, 1237
  treatment, 1237–8, *1237*
Atopy, 72
  and asthma, 847
  environmental factors, 72
Atopy/atopic diseases, *see* specific diseases
Atorvastatin
  acute coronary syndromes, *754*
  hyperlipidaemia, *1066*
Atovaquone, babesiosis, 160
ATP, 19, 1098
ATP-binding cassette transporter 1 gene, 1062
ATP-binding cassette transporter deficiency, 367
ATP7B gene, 357
Atractylis gummifera, 952
Atria (heart)
  contraction, 684
  depolarization, 684
  ectopic beats, 726
  myocyte action potentials, 682, *682*
  see also Left atrium (atrial); Right atrium (atrial)
Atrial arrhythmias *see* Atrial tachyarrhythmias
Atrial fibrillation, 690, *721*, 724–5
  acute heart failure, 742
  anticoagulation, 725
  atrial septal defect, 775
  causes, *723*, 724
  DC-cardioversion, 724
  ECG, *723*, 724
  lone, 724
  management, 724–5, *730*
  mitral regurgitation, 763
  mitral stenosis, 761
  post-STEMI, 759
  symptoms and signs, 724
  tricuspid regurgitation, 768
  Wolff-Parkinson-White syndrome, *722*
Atrial flutter, *721*, 725
  catheter ablation, 732
  DC-cardioversion, 725
  ECG, *723*, 725
  management, *730*
Atrial myxoma, 787, *787*
Atrial natriuretic peptide, 351, 615, *651*, 652, 1023
  heart failure, 735–6
Atrial septal defect, 721, *774*, 775–6
  auscultation, *776*
  clinical features, 775
  investigations, 775–6
  ostium primum, 775
  ostium secundum, 775, *776*
  pathophysiology, *776*
  right bundle branch block, *720*
Atrial tachyarrhythmias, 723–6, *723*
  atrial fibrillation *see* Atrial fibrillation
  atrial flutter *see* Atrial flutter
  causes, *723*
  management, 729–33
  premature beats, *723*
Atrial tachycardia, *721*, 725–6
  catheter ablation, 732
  fibrillation *see* Atrial fibrillation
  flutter *see* Atrial flutter
  management, *730*
  multifocal, *721*
  with second-degree AV block, *723*
Atrioventricular (AV) block, 717–19
  first-degree, 717, *717*
    heart sounds, 692
  prohibited drugs, *923*
  second-degree, 718, *718*
  third-degree, 718–19, *718*, *719*
    aetiology, *719*
    broad complex escape rhythm, 719
    narrow complex escape rhythm, 719
Atrioventricular (AV) junctional tachycardias, 721–3, *722*
Atrioventricular (AV) node, 682
  blocking agents, 725
  blood supply, 685
  disease
    re-entry tachycardia, 721, *721*, *722*, *730*
    reciprocating complexes, *721*
    see also Atrioventricular (AV) block
Atrioventricular (AV) reciprocating tachycardia, 721–2, *721*, *722*
  management, *730*
Atropa belladonna, 952
Atrophy, *1227*
Atropine, *937*
  calcium channel blocker overdose, 941
  organophosphorus poisoning, 947
Attention, 1188
Attitude (clinician's), communication difficulties, 10–11
Atypical mole syndrome, 1258
Audiogram, *1079*
Audiometry, pure tone, 1078, *1079*
Auditory lesions, 1109
Aura, 1139
Auroscope, 1078
Auscultation
  abdomen, 243
  cardiac, 692
    aortic stenosis, 766
    atrial septal defect, *776*
    coarctation of aorta, *778*
    Fallot's tetralogy, *779*
    mitral regurgitation, *761*
    mitral stenosis, 761–2, *761*
    ventricular septal defect, *775*
    see also Heart sounds
  chest, 820
Austin Flint murmur, 767
Autoantibodies, 75, 414, 541
  autoimmune hepatitis, 344
  serum, 325
  systemic lupus erythematosus, 543, *544*
  systemic sclerosis, 546
  see also specific antibodies
Autoimmune chronic pancreatitis, 380
Autoimmune diseases, 73–5
  CD4 T cell activation, 74
  classification, 75
  and endocrine dysfunction, 965, *965*
  mechanism of immune damage, *74*, 75
  and primary biliary cirrhosis, 355
  regulatory T cells, 74
  restoration of tolerance in, 76
  rheumatic *see* Autoimmune rheumatic disease
  thymic education, 74
  see also Autoantibodies; Autoimmunity; Autoimmunity (type II hypersensitivity); specific diseases
Autoimmune gastritis, 262, 263, 399
Autoimmune haemolytic anaemias, 75, **415–17**
  causes, *415*
  chronic cold haemagglutinin disease, 417
  cold, *415*, 416–17
  direct antiglobulin test, 415, *415*
  immune destruction of red cells, 415–16, *416*
  paroxysmal cold haemoglobinuria, 417
  warm, *415*, 416
Autoimmune hepatitis, 343–4, 554–5
  antinuclear antibodies, *504*
  autoantibodies, 344
  clinical features, 344
  course and prognosis, 344
  HLA system in, *60*
  investigations, 344
  liver biopsy, 344
  pathogenesis, 343–4
  renal tubular acidosis, *676*
  transplantation, 348
  treatment, 344
Autoimmune hyperthyroidism *see* Graves' disease
Autoimmune polyendocrinopathy type 1, 1026
Autoimmune polyglandular syndrome type 1, 74
Autoimmune rheumatic disease, **541–8**
  antiphospholipid syndrome, 544–6
  overlap syndromes, 548
  polymyositis/dermatomyositis *see* Polymyositis/dermatomyositis
  Sjögren's syndrome *see* Sjögren's syndrome
  systemic lupus erythematosus *see* Systemic lupus erythematosus
  systemic sclerosis *see* Systemic sclerosis
  undifferentiated, 548
  see also Rheumatoid arthritis (RA)
Autoimmune thrombocytopenia, 416
Autoimmunity (type II hypersensitivity), 541
  ANCA antibody formation, 596
  diabetes mellitus, 1033
  see also Autoimmune diseases
Autologous stem cell transplantation, 462–3
Autologous transfusion, 428
Autonomic nervous system (ANS), chronic renal failure, 630
Autonomic neuropathy, 1176
  diabetes-related, 1056
  HIV-related, 188
Autonomy
  patients' rights, 1–2
  respect for, 3–6
Autoregulation, 681, 686, 1127–8
Autosomal dominant disorders, 42, *42*
  hypophosphataemic rickets, 670
  pseudo-hypoaldosteronism, *659*, 667
Autosomal recessive disorders, *42*, 43
  pseudo-hypoaldosteronism, *659*, 667
Autosomal-dominant polycystic kidney disease (ADPKD), 642–4
  clinical features, 642–3
  complications and associations, 643
  diagnosis, 643
  screening, 643
  therapy, 643–4
Avascular (aseptic) necrosis of bone, 199, 410, 501, *512*, *520*, 541, 554, 556, 560
Avascular necrosis (osteonecrosis) of femoral head, 513
Aveeno cream, *1270*
Average, 932–3
Avian influenza, 79, *81*, 91
Avidin, 223
Axonal degeneration, 1171
Azacytidine, myelodysplasia, 422
Azathioprine (AZA), 76
  ANCA-positive vasculitis, 597
  aphthous ulcers, 247
  autoimmune haemolytic anaemia, 416
  autoimmune hepatitis, 344
  bullous pemphigoid, 1256
  chronic discoid lupus erythematosus, 1252
  Crohn's disease, *289*, 290
  dermatomyositis, 1251
  erythema multiforme, 1250
  focal segmental glomerulosclerosis, 587
  genetic polymorphism in metabolism, *926*
  glomerular disease
  idiopathic pulmonary fibrosis, 873
  lupus nephritis, 591
  membranous glomerulonephropathy, 588
  microscopic colitis, 294
  multiple sclerosis, 1153
  pemphigus vulgaris, 1256
  pericarditis, 795
  polyarteritis nodosa, 551
  polymyositis/dermatomyositis, 548
  renal transplantation, 640
  rheumatoid arthritis, 532
  side-effects, hepatotoxicity, *366*
  systemic lupus erythematosus, 544
  thrombocytic purpura, 436
Azelaic acid, rosacea, 1247
Azimilide, supraventricular tachycardia, 723

Azithromycin, *94*, 97
  babesiosis, 160
  chancroid, 180
  Chlamydia, 176
  donovanosis, 180
  gastroenteritis, *132*
  syphilis, 179
  tuberculosis, 866
  urethritis, 177
Azoles, 100, *101*
Azoospermia, *995*, 1000
Aztreonam, 95
  pancreatitis, 379
  urinary tract infections, 602

B cell(s), 53
  affinity maturation, 64
  antigen receptors, 63
  class switching, 64
  deficiency, *69*, 70–1
  origins of, 388
  sensitization, 275
B cell lymphoma
  chromosome translocations, *480*
  diffuse, 482
  gastric, 262, 266
  lymphoplasmacytic, *480*, 482
B-lymphocytes *see* B cell(s)
Babesia, *82*
  B. divergens, 160
  B. microti, 160
Babesiosis, *82*, 160
Babinski reflex, 1112
Bacillary angiomatosis, *148*
Bacille Calmette-Guérin (BCG), 867
  bladder tumours, 645
  cancer treatment, 465
Bacillus
  B. anthracis, *80*, *123*, 142
    see also Anthrax
  B. cereus, *123*
    diarrhoea, *307*
    food poisoning, *133*
    gastroenteritis, *128*
Bacillus Calmette-Guérin, immunization, *103*
Bacillus cereus, gastroenteritis, 131
Back pain
  acute abdomen, 314
  elderly patients, 552
  lumbar *see* Low back pain
    see also Disc disease, prolapse/protrusion
Bacteraemia, 135–6, *135*, *136*
  definition, *899*
  infective endocarditis, 770
Bacteria, 80
  antibiotic resistance, *see* Antibiotics, resistance
  classification, 122–3, *123*
  Gram-negative, protein secretion, *86*
  in urine, 579
Bacterial infection(s), **122–40**, *123*
  conjunctivitis, 1088
  developing and tropical countries, **140–9**
  diagnosis and management, 123
  HIV-associated, 203–4
  respiratory tract, 125–8
  skin and soft tissue, 123–5, 1228–30
  transfusion-transmitted, 428
  see also specific infections
Bacterial meningitis, 134
Bacterial pericarditis, 795
Bacterial vaginitis, *172*
Bacterial vaginosis, 183
Bactericidal permeability-inducing protein, 275
Bacteriophages, 35
Bacteriuria, 579
  diagnosis, 602, *602*
  pregnancy, 603
  renal calculi, 612
  see also Urinary tract infections (UTI)
Bacteroides, *70*, 122, 280
  B. fragilis, *123*
    septicaemia, *135*
  bacterial vaginosis, 183
  pyogenic abscess, 360
  urethritis, 176
Bak proteins, 28
Balance impairment *see* Vertigo
Balanitis, 182
  circinate, 536
Balanoposthitis, 182
Balantidiasis, 162
Balantidium coli, *161*, 162
Balkan nephropathy, 606
Ball-and-socket joints, 499
Balloon enteroscopy, 246
Balloon flotation catheter, 902, *903*
  complications, *903*
  see also Swan-Ganz catheters
Balloon tamponade, 350
Bamboo spine, 534, *534*
Bannayan-Riley-Ruvalcaba syndrome, *301*, 302
Barbiturates
  poisoning, *935*
  side-effects, rash, *1265*
Bardet-Biedl syndrome, *1035*
Bariatric surgery, 232–3
Barium contrast studies, 245
  absorbable water soluble contrast agents, 245
  barium enema *see* Barium enema
  barium swallow *see* Barium swallow
  small bowel enema *see* Small bowel enema
  small bowel follow-through *see* Small bowel follow-through
  see also Barium meal
Barium enema, 245
  colorectal carcinoma, 304
  constipation, 295
  diverticular disease, 298, *299*
Barium follow-through, Crohn's disease, 289
Barium meal, 250
  double-contrast, 245
  gastric cancer, 265
Barium swallow, 245, 250
  achalasia, 254, *254*
  carcinoma of oesophagus, *257*
  systemic sclerosis, 546
Barlow's syndrome, 764–5
Barometric pressure, high altitude, *955*
Barotrauma, 914
  middle ear, 957
Barrett's oesophagus, 252–3, *252*
  screening and surveillance, 253
  treatment, 253
Bartonella
  B. bacilliformis, *82*, *148*
  B. henselae, *80*, 125, *148*
  B. quintana, *148*, *191*
  infective endocarditis, 770
Bartonellosis, 148, *148*
Bartter's syndrome, *658*, 665, *665*
Barttin, *658*, 665
Basal cell carcinoma, 1259, *1259*
Basal cell papilloma, 1257, *1257*
Basal metabolic rate (BMR), 208, *208*, *209*
  elderly people, 228
  equations for, **209**
  obesity, 230
Basic fibroblast growth factor (bFGF), 50, 451–2
Basic life support, 709–11, *710*
  airway, 709
  breathing, 709–10
  circulation, 710–11
Basidiobolus, 152
Basilar invagination, *1114*
Basilar migraine, 1109, 1137
Basiliximab, 76, *76*
Basophils, *56*, 431
  in anaphylaxis, 73
  IgE binding, 431
  mediators, *58*
Batteries (electric), ingestion of, 941
Battery, unlawful, 4
Bax proteins, 23, 28
Bayliss myogenic response, 686
Bazett's square root formula, 695, *695*
BCG vaccination *see* Bacille Calmette-Guérin (BCG)
Bcl-2 oncogene, 465
Bcl-2 proteins, 23, 28, 451
Bcl-2/*IgH* gene, *453*
Bcl-6, **33**
BCR gene, 467
BCR-ABL gene, 453, *453*, 467–8
Beau's lines, 1266
Beclometasone
  asthma, *854*, 855
  COPD, 839
  rhinitis, 833
Becquerels, 958
Bed bugs, 1235
Bee venom allergy, 72, 73
Before-and-after studies, 931
Behaviour, 1186
  altered, physical causes, *1190*
  illness-related, 1191
Behaviour therapy, 1208
Behavioural disturbance, 1191, *1191*
Behçet's syndrome, 552, 1159
Bejel, 142
Beliefs
  abnormal, 1187–8
  illness, 1187
Bell's palsy, 1098, 1107–8, 1115
  clinical trial, *930*
  management and course, 1107–8
Bell's phenomenon, 1107
Bels, 959
Bence Jones protein, 484, 485
Bendroflumethiazide, *657*
  heart failure, *739*
  hypercalciuria, 613
  hypertension, 803
Bendroflumethiazole
  fixed dose prescribing, *925*
  nephrotic syndrome, 585
Bends, 957
Benign lymphoid polyposis, *301*
Benign paroxysmal positional vertigo, 1079–80
  diagnosis, 1079
Benign prostatic hypertrophy, 645–6
  clinical features, 646
  management, 646
  prognosis, 646
  urinary tract obstruction, 617
Benzbromarone, gout, 538
Benzene exposure, and leukaemia, 467
Benznidazole, 158
Benzodiazepines, 1209
  overdose, *935*, 941
    antidotes, *937*
  withdrawal syndromes, *1209*
Benzylpenicillin, 127
  infective endocarditis, *773*
  meningitis, *1155*
  neurosyphilis, 1157
  pneumonia, 862
  side-effects, *928*
Beraprost, pulmonary hypertension, 783
Beriberi, 221, *789*, 1174
  diagnosis, 221
  dry, 221
  infantile, 221
  treatment, 221
  wet, 221
  see also Thiamin (vitamin B_1)
Berlin (Prussian) blue, *937*
Bernard-Soulier syndrome, 26, 437
Bernoulli equation, 699, 701
Bernstein test, 250
Berylliosis, 880
Beryllium poisoning, 868
Beta(β)-adrenoceptors, heart, 684, *684*
Beta$_2$(β_2)-agonists
  asthma, 854–5
  slow-acting, 854, 855
Beta(β)-blockers
  angina pectoris, 750
  arrhythmogenic right ventricular cardiomyopathy, 791
  atrial fibrillation, 725
  genetic polymorphism in metabolism, *926*
  heart failure, 738–9, *739*
  hypertension, 804
  hypertrophic cardiomyopathy, 790
  Marfan's syndrome, 781
  mitral valve prolapse, 764
  myocarditis, 788
  overdose, 941
  side-effects
    asthma, 849–50
    elderly patients, *924*
    rash, *1265*
    retroperitoneal fibrosis, 616
  specific conditions
  supraventricular tachycardia, 723
  variceal bleeding, 351
Betahistine, 1080
Betaine, 655
Betamethasone valerate, *1237*
Bevacizumab, *464*
  age-related macular degeneration, 1094
  colorectal carcinoma, 490
  renal cell carcinoma, 645
Bexarotene, *464*
Bezafibrate, hyperlipidaemia, *1066*
Bezold-Jarisch reflex, 717
Bicarbonate, 672
  administration
  in dialysate, *634*
  interstitial fluid, *650*
  intracellular fluid, *650*
  plasma, *650*
  renal calculi, 613
  renal reabsorption, 672–3, *672*
Bifidobacterium, 280
  B. infantis, 312
Biglycan, 500
Biguanides, 1038, *1038*
  overdose, 940
Bile
  flow, 322
  formation, 321–2
  secretion, 321–2
  see also Bile acids/salts; Bilirubin
Bile acids/salts, 1062
  detergent function, 322
  and diarrhoea, 281
  drugs binding, 282
  enterohepatic circulation, 274
  loss of, 277
  malabsorption, 307–8, *308*
  metabolism, 321–2
  pool composition, 367
  primary, 322
  receptors, 322
  recirculation, 274, 323, *323*, 1062
  secondary, 322
  synthesis, *322*
Bile canaliculi, 320, 322
Bile ducts, 320
  common, 320, 375
    stones *see* Common bile duct stones
    stricture, *374*
  congenital dilatation, 360
  malignant obstruction, 374–5, *374*

stenting, 372
urgent endoscopic clearance, 372
Bile pigment stones, 367–8
black, 368
brown, 368
Bile salt export pump, 322, 330
Bilevel positive airway pressure (BiPAP), 916
COPD, 840, 841
Bilharzia *see* Schistosomiasis
Biliary cirrhosis
primary *see* Primary biliary cirrhosis (PBC)
secondary, 356
Biliary colic, 368
Biliary pain, 242
Biliary tract
structure, 320
tumours, 373–5
secondary, 374
see also Gall bladder
Bilio-pancreatic diversion, 233
Bilirubin
dipstick tests, 325
metabolism, 322–3
raised levels
common bile duct stones, 370
gallstones, 368
pernicious anaemia, 400
see also Jaundice
serum, 324
pyogenic abscess, 361
Bilirubinuria, 325
Biliverdin, 323, 418
Binet staging system, *474*
Binswanger's disease, 1130
Biochemistry
adrenal gland, 1008
chronic renal failure, 627
coeliac disease, 278
diabetes insipidus, 1017
liver, 324, 332
renin-angiotensin-aldosterone system, 1023
Biofeedback, 1208
faecal incontinence, 297
Biofilms, 85
Biolimus, 750
Biological agents, 960
Biomedical outcome, communication effects, 9–10
Biopsy
bone, 560, *560*
brain, 1121
brain tumours, 1161
cardiac, 737
endomyocardial, 788, 793–4, 797
iliac crest, 567
liver, 328, *328*, 341, 865
muscle, 1122, 1178–9
pericardial, 796
peripheral nerve, 1122
pleural, 827, *827*, *828*
renal, 582, *582*
chronic renal failure, 628
glomerular disease, *585*
scalene node, 828
skin, *1228*
small intestine, 277
tonsillar, 1121
transbronchial, 865, 870
video-assisted thoracoscopic, 829
Bioterrorism/biowarfare, 959–60, *959*
emergency planning, 960
potential pathogens, 959–60
anthrax, 960
botulism, 960
plague, 960
smallpox, 959–60, *960*
Biotin, 223
recommended IV intake, *234*
Bipolar affective disorder, 1198
mania, 1205–6
see also Mania
treatment, 1201
see also Depression
Birbeck granules, 874
Bird fancier's lung, *876*
Birth defects in epilepsy, 1142–3
Birth marks, 1267–8, *1268*
Bisacodyl, *297*
Bismuth
H. pylori eradication, 261
microscopic colitis, 294
Bisoprolol, 730
heart failure, *739*
hypertension, 804
Bisphosphates
myeloma, 484
see also individual drugs
2,3-Bisphosphoglycerate (2,3-BPG), 391
Bisphosphonates
breast cancer, 489
osteoporosis, 563, 564, *564*
Paget's disease, 566
pain relief, *495*
prostate cancer, 492
side-effects, elderly patients, *924*
see also individual drugs
Bites
insects, 951, 1235
snakes, 951, *951*
spiders, 951
Bitot's spots, 218
Bivalirudin
acute coronary syndromes, *754*
arterial thrombosis, *443*
venous thrombosis, 445
Biventricular cardiac failure *see* Heart failure
Björnstad syndrome, *1078*
Black hairy tongue, 247
Black pigment gallstones, 368
Blackflies, *82*
Blackheads, 1246
Blackouts, 1097
Blackwater fever, 155
Bladder
anatomy, 1119
catheterization, 1165
control, 1119
diabetic neuropathy, 1056
fullness, 1119
functions, 1119
irritable, 602
nerve supply, *1119*
outflow obstruction *see* Bladder outflow obstruction
stones, *see* Urinary tract stones
tumours, 645
Bladder cancer, epidemiology, *449*
Bladder outflow obstruction
symptoms, *614*
see also Urinary tract obstruction
Blalock shunt, 779
Blastocystis hominis, 162
Blastomyces, *B. dermatitidis*, *82*, 152
Blastomycosis, *150*, 152
Bleach
allergy, *848*
ingestion of, 944
Bleeding
gastrointestinal *see* Gastrointestinal bleeding
urinary tract, *see* Haematuria
see also Blood loss; Haemorrhage
Bleeding disorders
investigation, 434–5
platelet disorders *see* Platelets, disorders
see also Coagulation; specific disorders
Bleeding time, 434
haemophilia A, *439*
von Willebrand disease, *439*
Bleomycin, 460
germ cell tumours, 492
Hodgkin's lymphoma, *461*
non-Hodgkin's lymphoma, *461*
side-effects
haemolytic uraemic syndrome, 598
respiratory, *877*
skin changes, 1252
systemic sclerosis, 545
Blepharitis, 1087, *1087*, 1090
Blepharospasm, 1089, 1150
Blighia sapida, 952
Blind therapy, 92
Blinding (clinical trials), 929
Blindness
cortical, 1102
see also Visual impairment/loss
Blister cells, *414*
Blister fluid, *1228*
Blisters
overdose/poisoning, 936
see also Bullae; Vesicles
Blood, 428–30
cells *see* Blood cells
circulation *see* Circulation
clotting *see* Coagulation
components, *see individual components*
haemostasis, 431–5
see also Coagulation; Thrombosis
peripheral, 389
pressure *see* Blood pressure
restoration of volume, *see* Blood transfusions
in rheumatoid arthritis, 528
volume *see* Blood volume
whole, 429
Blood cells, 387
formation *see* Haemopoiesis
see also Leucocytes; Red blood cells (erythrocytes)
Blood clotting *see* Coagulation
Blood components, 428–30
Blood count, 434
acute leukaemia, 468
chronic lymphocytic leukaemia, 473
chronic myeloid leukaemia, 473
Hodgkin's lymphoma, 477
infective endocarditis, 772
iron deficiency anaemia, 396
liver disease, 325
musculoskeletal problems, 503
myeloma, 484
non-Hodgkin's lymphoma, 480
normal, 389
polyarteritis nodosa, 550
rheumatoid arthritis, 528
sarcoidosis, 870
sickle cell anaemia, 410
small intestinal disease, 276
systemic lupus erythematosus, 543
systemic sclerosis, 546
Blood culture, *89*
Blood diseases *see* Haematological disease
Blood donors, 425
Blood film, 393, *394*, 434
acute leukaemia, 468, *469*
chronic lymphocytic leukaemia, 473
chronic myeloid leukaemia, 473
dimorphic, 393
G6PD deficiency, *414*
hereditary spherocytosis, 405, *405*
iron deficiency anaemia, 396
megaloblastic anaemia, *398*
myeloma, 485
pyruvate kinase deficiency, 414
sickle cell anaemia, 410
Blood flow
and regulation of metabolism, 214
renal *see* Renal blood flow
see also Circulation
Blood gas analysis *see* Arterial blood gases
Blood glucose *see* Glucose, blood
Blood group antigens
Blood groups, 424
ABO system *see* ABO blood group
pretransfusion testing, 424
Rh system *see* Rh blood group
Blood islands, 387
Blood loss *see* Anaemia; Bleeding; Haemorrhage
Blood ordering, 425
elective surgery, 425
emergencies, 425
Blood oxygenation, 892–3
alveolar oxygen tension, *892*
mixed venous oxygen tension/saturation, 893, *893*
oxyhaemoglobin dissociation curve, 892, *892*
pulmonary gas exchange, 893
Blood pressure, 691, *691*, 798
abnormal, *798*
see also Hypertension; Hypotension
ambulatory monitoring, 801–2, *801*
control, 1022–3
see also Antihypertensive drugs
measurement, *691*, 900, *901*
on standing, *see* Postural hypotension
Blood products, 429
Blood salvage, 428
Blood substitutes, 906
Blood transfusions, **423–30**
antibody screening, 424–5
autologous, 428
bacterial contamination, 428
barcode identification, 426, *427*
compatibility testing, 424–5
complications, 425–8, *426*, 905–6
coagulopathy, 906
hyperkalaemia, 906
hypocalcaemia, 906
immunological, 426–7
increased oxygen affinity, 906
metabolic acidosis/alkalosis, 906
microembolism, 906
non-immunological, 428
prevention, 426
temperature changes, 905
crossmatching, 425
donor selection, 425
errors, 424, 425, *426*
avoidance of, 426, *427*
incompatibility, 427
exchange transfusion, 418
massive, 441
procedure, 424–5
shock, 905–6
specific conditions, sickle cell anaemia, 411
unnecessary, avoidance of, 428
viral contamination, 428
see also Blood grouping; Blood groups
Blood vessels, 685–7
autoregulation, 681, 686
Bayliss myogenic response, 686
vasodilator washout effect, 686
vasomotor control, 686–7
see also entries beginning vaso-/vascular; specific vessels
Blood volume
gastrointestinal bleeding, 267–8
Bloom's syndrome, 453, *453*
Blue bloaters, 835, 836
Blue diaper syndrome, 1071
Blue naevi, 1257
Blushing, 1247
Bocavirus, *104*, 109
BODE index, 841, *842*
Body core temperature, 954
Body mass index (BMI), 214, 1034
obesity, 228, 230, *231*

Body packers, 936
Body stuffers, 936
Bodyweight *see* Weight
Boerhaave's syndrome, 256
Bohr effect, 391
Boils, *86*, 1229
Bolivian haemorrhagic fever, *114*
Bombesin, 229
Bone, 557
  aseptic necrosis *see* Avascular (aseptic) necrosis of bone
  basic multicellular unit, 558, *558*
  biopsy, 560, *560*
  cortical (compact), 557
  densitometry, *see* Bone scans (scintigraphy)
  density, 562
  diseases *see* Bone diseases/disorders
  epiphyseal, 499
  formation, biochemical markers, 560
  growth and remodelling, 558
  hypercalcaemia, 1020
  isotope scanning *see* Bone scans
  juxta-articular, 499–500
  lamellar, 557
  malignancy, 555, *555*
  mass, 562, *563*
  matrix components, 557
  metaphysis, 499
  metastases, 452
  myeloma involvement, 484
  pain *see* Bone pain
  resorption, biochemical markers, 560
  structure, 557, *557*
  trabecular (cancellous), 557
  woven, 557
Bone age, 978
Bone cells, 557–8
  osteoblasts, 557
  osteoclasts, 58, 557–8
  osteocytes, 557
Bone densitometry, coeliac disease, 278
Bone diseases/disorders, 557–68
  adynamic, 628
  HIV-related, 198
  infections, 541
    see also Osteomyelitis
  investigation, 559–60
  osteonecrosis, 565
  osteoporosis *see* Osteoporosis
  Paget's disease *see* Paget's disease
  renal osteodystrophy, 628, *628*
  sickle cell anaemia, 410
Bone marrow, 387
  anaemia, 393, *393*
  aplasia, 402
    sickle cell anaemia, 410
  aspirate
    acute leukaemia, 468–9, *469*
    chronic lymphocytic leukaemia, 473
    chronic myeloid leukaemia, 473
    Hodgkin's lymphoma, 477
    myeloma, 485, *485*
    non-Hodgkin's lymphoma, 481
  failure, *see* Aplastic anaemia
  fibrosis, 628
  infiltration, 484
  iron, *396*
  iron deficiency anaemia, 396
  megaloblasts, *397*
  pernicious anaemia, 399–400, *400*
  red, 387
  sampling, *393*
  suppression, 461
  toxins, 628
  yellow, 387
Bone marrow transplantation
  aplastic anaemia, 403
  myelodysplasia, 422
  paroxysmal nocturnal haemoglobinuria, 419
  sickle cell anaemia, 411
  thalassaemia, 408–9
Bone mineral density (BMD), *561*, 563
  see also Osteoporosis
Bone pain
  hypercalcaemia, 1020
  myeloma, 484
Bone scans (scintigraphy), 560
  Paget's disease, 565
  skull and spine, 1120
  specific conditions, musculoskeletal problems, 505
Book marking, 51
Bordetella
  B. bronchiseptica, 127
  B. parapertussis, 127
  B. pertussis, *123*, 127
    pneumonia, 862
    see also Pertussis
Bornholm disease *see* Myositis
Borrelia, *123*
  B. burgdorferi, *80*, *82*, 139
    see also Lyme disease
  B. duttoni, *82*, 147
  B. recurrentis, *82*, 146
Bortezomib, *464*
  mantle cell lymphoma, 482
  myeloma, 486
Bosentan
  pulmonary hypertension, 783
  systemic sclerosis, 547
Botulism, 134
  as biological agent, *959*, 960
  immunization, *102*
Bouchard's nodes, 521
Bovine spongiform encephalopathy (BSE), 122
Bowel *see* Intestinal; Intestine
Bowel sounds, 314
Bowen's disease, 1258
Bowman's capsule, 572, 582
Bowman's space, 572
  macrophages and epithelial cells in, 595, *595*
Brachytherapy, 465
  bronchial carcinoma, 883
Bradyarrhythmias, post-STEMI, 759
Bradycardias, 689, 711, 717
  definition, 715
  heart sounds, 692
  neurally mediated systems, 717
  pacing, temporary, 713, *713*
  sinus *see* Sinus bradycardia
  sinus node, 716
  see also Atrioventricular (AV) block; Syncope
Bradykinesia, 1112, 1113
Bradykinin, 285, 686
  asthma, 852
Brain
  arterial supply, *1127*
  biopsy, 1121
  diffuse disorders
  glucose uptake, 1029
  imaging
    computed tomography, 1119–20
    digital cerebral angiography, 1120
    Doppler ultrasound, 1120
    functional MRI, 1120
    MRI, 1120
    PET, 1120
    radionuclide studies, 1120
    SPECT, 1120
    see also specific techniques
  tumours *see* Brain tumours
Brain abscess, *86*, 1141, 1159, *1159*
  clinical features, 1159
  and epilepsy, 1141
  management, 1159
Brain damage/injury, 1162–3
  clinical course, 1163
  competency, 6
  immediate management, 1163
  mechanisms, 1163
  rehabilitation, 1163
  see also Head injury; Skull fracture
Brain death, 920–1, 1121
  diagnostic tests, 920–1
  exclusions, 920
  preconditions, 920
Brain diseases/disorders
  Alzheimer's disease *see* Alzheimer's disease
  clinical features, 1098–100, *1099*
  degenerative
  HIV-related, 1158
  Huntington's disease *see* Huntington's disease (HD)
  Parkinson's disease *see* Parkinson's disease
Brain natriuretic peptide, 1023
  heart failure, 735–6
  synthetic, 736
Brain tuberculoma, 1159
Brain tumours, 1160–2, *1160*
  clinical features, 1161
  direct effects, 1161
  epidemiology, *449*
  and epilepsy, 1141
  investigations, 1161
  management, 1161–2
  see also specific cancers
Brainstem, 1109
  demyelination, 1152
  infarction, 1129–30, *1129*, *1130*
  lesions, 1110–11
Brainstem death, *1123*
Bran, 211
Branched-chain ketoacid dehydrogenase deficiency, *1070*
BRCA1, **33**, 454, 487
BRCA2, 454
Breast
  cancer *see* Breast cancer
Breast cancer, 453, **487–9**
  advanced, 489
  aetiology and pathology, 487
  biological therapy, 488
  bisphosphonates, 489
  chemotherapy, *461*, 488, 489
  endocrine therapy, 488, *488*, 489
  epidemiology, *449*
  familial, *453*
  geographical distribution, 451
  her2/cerbB2 targeted therapy, 488
  histology, *487*
  investigations, 487
  menopausal status, 488
  neoadjuvant and primary systemic treatment, 489
  pericarditis, 795
  prognostic factors, 487–8
    early breast cancer, 487–8
    poor, *487*
  radiotherapy, 488
  screening, 454
  surgery, 488
  symptoms and signs, 487
Breast cancer resistant protein, 393
Breast-feeding
  anticonvulsants, 1143, *1143*
  prescribing in, 924
Breath, 1124
Breath tests
  ^{13}C urea, 245, 260, *260*
  ^{14}C-glycocholic acid, 277
  hydrogen, 277
  nitric oxide exhalation, 826, 853
Breath-holding attacks, 1145
Breathing, 709–10, 814–15
  exercises, 840
  expiration, 814
    forced, 814, 816, *816*
  inspiration, 814
  periodic, 688
  tidal, 814
Breathlessness, 815, 819–20
  cancer-related, 496
  relief of, 840
  see also Dyspnoea
Brill-Zinsser disease, 147
Briquet's syndrome, 1195
British Medical Association (BMA), 2
Brittle asthma, 857, *857*
Brittle bone disease, 568
Broad-based gait, 1097
Broca's aphasia, 1099, *1099*
Brock's syndrome, 864
Brodie's abscess, 541
Bromocriptine
  acromegaly, 980
  hyperprolactinaemia, 982
  pituitary tumours, *973*
  side-effects, retroperitoneal fibrosis, 616
Bronchial carcinoid, 884
Bronchial carcinoma, **880–3**
  cell types, 880
  computed tomography, 823, *823*
  cough, 819
  direct spread, 881
  investigations, 882
  metastatic complications, 881
  mortality, *880*
  non-metastatic extrapulmonary manifestations, 881–2, *881*
  non-small-cell, 880–1
  palliative care, 883
  pericarditis, 795
  small-cell, 881
  symptoms, *881*
  treatment, 882–3
    radiation therapy, 883
    surgery, 883
  see also Lung cancer/tumours
Bronchial provocation tests, 848
Bronchial veins, 814
Bronchiectasis, 819, 842–4
  aetiology, 842–3, *843*
  clinical features, 843
  complications, 844
  computed tomography, 823, *823*
  investigations, 843, *843*
  prognosis, 844
  sputum, 819
  treatment, 843–4
    anti-inflammatory agents, 844
    antibiotics, 843–4
    bronchodilators, 844
    postural drainage, 843
    surgery, 844
Bronchioles, 811–12
Bronchitis
  acute, 835
  chronic
    pathological changes, *836*
    see also Chronic obstructive pulmonary disease (COPD)
Bronchoalveolar cell carcinoma, 881
Bronchoalveolar lavage
  extrinsic allergic alveolitis, 877
  idiopathic pulmonary fibrosis, 873
Bronchodilators
  asthma, 855
  bronchiectasis, 844
  COPD, 839
Bronchomotor tone, 815
Bronchopneumonia, 861
  see also Pneumonia
Bronchopulmonary infections
  see also Respiratory tract infections
Bronchoscopy
  fibreoptic *see* Fibreoptic bronchoscopy

Bronchus, 811–12
  collapse, *821*
  inflammation, 846
  rupture, 884
Bronchus-associated lymphoid tissue (BALT), 53
Bronze skin, 356
Brown pigment gallstones, 368
Brown-Séquard syndrome, 1112, *1117*, 1118
Brucella, 122, *123*
  B. abortus, *80*
  B. melitensis, *80*, 540
Brucellosis, *80*, 138, 540
  as biological agent, 960
  clinical features, 138
  diagnosis, 138
  management and prevention, 138
  pyrexia, *91*
  returning travellers, *90*
Brugada syndrome, 727, *789*
  ECG, *727*
Brugia
  B. malayi, *82*, *163*, *164*
  B. timori, *163*, *164*
Bruising
  easy bruising syndrome, 435
  factious, 435
  thrombocytopenia, 436
  vitamin K deficiency, 440
Bubonic plague, 146
Budd-Chiari syndrome, 327, 359–60, 419
  causes, 359
    steroids, 365
  clinical features, 360
  differential diagnosis, 360
  investigations, 360
  prognosis, 360
  treatment, 360
Budesonide
  asthma, *854*, 855
  Crohn's disease, 289
  microscopic colitis, 294
Buerger's disease, 806, 809
Buffer systems, 672, *672*
  acid excretion, 674, *674*
Buffy coat analysis, 155
Building-related illnesses, 960–1
  legionnaires' disease, *82*, 128, 961
  non-specific, 960
  specific, 961
Bulbar palsy, 1110–11, 1166
Bulimia nervosa, 248, 1220
Bulla, *1227*
Bullous disorders, 1255–7, *1255*
  immunobullous disease, 1255–7
  mechanobullous disease, 1257
Bullous impetigo, *86*, 1228–9
Bullous pemphigoid, 1256, *1256*
Bumetanide, *657*
  ascites, 353
  bioavailability, 657
  heart failure, *739*
Bundle branch block, 719–20, *719*
  bifascicular, *719*, 720
  causes, 720, *720*
  clinical features, 720
  complete block, 719
  conduction delay, 719
  hemiblock, 720
  left, 720, *720*
  right, 719–20, *719*
  with supraventricular tachycardia, *726*
Bundle branches, 682
  block *see* Bundle branch block
Bunyaviruses, *110*, *114*, 116
Bupropion, 830
Burkholderia, *123*
  B. cepacia, *82*
    cystic fibrosis, 845
  B. pseudomallei, *82*, 148
Burkitt's lymphoma, 49, *105*, 482–3
  chromosome translocations, 479–80, *480*
  clinical features, *483*
  treatment, 483
Burning mouth syndrome, 246
Burns, electrical, 959
Bursa, 499
Bursitis
  Achilles tendon, *515*, 516
  anserine, 513
  calcaneal, 515
  infrapatellar, 513
  rheumatoid arthritis, 527
  shoulder, 507–8
Buruli ulcer, 142
Buserelin, prostate cancer, 646
Busulfan, 460
  myelofibrosis, 421
  polycythaemia vera, 420
  side-effects
    bone marrow aplasia, 402
    respiratory, *877*
  thrombocythaemia, 421
butyrate, and inflammatory bowel disease, 287
Butyrophenones, 1217
Byler's disease *see* Progressive familial intrahepatic cholestasis syndromes
Bypass surgery, 751–2, *751*, 758
Byssinosis, 880
Bystander activation, 74

^{13}C urea breath test, 245, 260, *260*
^{14}C-glycocholic acid breath test, 277
c-fos, **33**
  Paget's disease, 565
C-jun, **33**
c-myc gene, *453*
C-peptide, 1029, *1030*
C-reactive protein, *56*, 68, 389
  and coronary artery disease, 747
  in diabetes, 1034
  in infection, *88*
  measurement, *69*
  musculoskeletal problems, 503
  myeloma, 484
  pancreatitis, 378
C-type peptide, heart failure, 735–6
C1 esterase deficiency, 71
C1q nephropathy, 590, 591
C3
  focal segmental glomerulosclerosis, 587
  glomerular disease, *585*
  lupus nephritis, 591
C3a, 57
C3b, *416*
C3d, *416*
C4
  glomerular disease, *585*
  lupus nephritis, 591
C4a, 57
C5a, 57
CA15-3, as tumour marker, *456*
CA19-9, 383
  as tumour marker, *456*
CA125, 454
  ovarian cancer, 491
  as tumour marker, *456*, *493*
CAAT box, 30
Cabergoline, *976*
  acromegaly, 980
  hyperprolactinaemia, 982
  Parkinson's disease, 1147
  pituitary tumours, *973*
  side-effects
    constrictive pericarditis, 797
    mitral regurgitation, 763
    retroperitoneal fibrosis, 616
Cachexia
  cancer-related, 455, 495
  cardiac, 763
  HIV, 189
  see also Muscle atrophy (wasting)
Cachexia factor, 214
CADASIL, 1127
Cadherins, *25*, 26, 452, 583, 655
  rheumatoid arthritis, 524
Cadmium, *226*
Cadmium poisoning, uraemia, 619
Caeruloplasmin, 321, 357
  serum, 325
    Wilson's disease, 1122
Café-au-lait spots, 1170, 1253, 1254, 1265
Caffeine, 1198
Calabar swellings, 165
Calcaneal bursitis, 515
Calcific tendonosis, 507
Calcification, 693
  aorta, 693
  coronary artery, 693
  eggshell, 878
  myocardial, 693
  pericardial, 693
  valvular, 693
  see also Calculi
Calcimimetics, parathormone suppression, 633
Calciphylaxis, 630, *630*
Calcitonin, 559
  chronic regional pain syndrome type 1, 517
  osteoporosis, 563, 564
Calcitonin gene-related peptide, *272*, 814
  asthma, 852
Calcitonin-related peptide, 1136
Calcitriol
  hyperphosphataemia, 633
  osteoporosis, 564
  rickets/osteomalacia, 567
Calcium, 226
  absorption, 558–9
  cardiac contraction, 683, 684
  deficiency, 226
  dialysate
    haemodialysis, *634*
    peritoneal dialysis, *637*
  dietary intake, 558
  disorders, 1019–*22*
  see also Hypercalcaemia; Hypocalcaemia
  distribution, 558–9, *558*
  interstitial fluid, *650*
  intracellular fluid, *650*
  muscle contraction, 683
  osteoporosis prevention, 563, 564
  plasma/serum, *650*
    bone disorders, *559*
    heart failure, 735
    low, 276
    neurology, *1122*
    pancreatitis, *378*
    renal calculi, 613
    rickets/osteomalacia, 567
  restriction
  RNI, 226
    elderly people, 228
  sources of, 226
  total plasma, 559
  urinary, 559–60
    renal calculi, 613
Calcium bilirubinate, 368
Calcium channel blockers
  angina pectoris, 750
  atrial fibrillation, 725
  Conn's syndrome, 1024
  hypertension, 804–5
  overdose, 941
    antidotes, *937*
    clinical features, 941
    metabolic acidosis, *939*
    treatment, 941–2
  side-effects, postural hypotension, *659*
  see also individual drugs
Calcium channels, 683
  long-lasting (L), 683
  transient (T), 683
Calcium folinate, 460
Calcium homeostasis, 558–9, *558*
Calcium ions, 24
Calcium oxalate stones, *609*, 613
Calcium phosphate stones, *609*
Calcium pyrophosphate, 356, 521
  pseudogout, 539
Calcium-dependent adherins *see* Cadherins
Calculi
  bladder, aetiology, 611
  common bile duct *see* Common bile duct stones
  renal *see* Renal calculi
  salivary gland, 248
  vesical *see* Renal calculi
Caliciviruses, *110*, 112, *113*
California encephalitis, 1156
Caloric tests, 1109, 1124
Calprotectin, 289
CAMP, 24
CAMP response element, 31
CAMP response element binding proteins *see* CREB
Campbell de Morgan spots, 1258
Campylobacter, *70*
  antibiotics, *132*
  C. jejuni, *80*, *123*
    diarrhoea, *307*
    food poisoning, *133*
    gastroenteritis, *128*, 129
Canal of Schlemm, 1086
Cancer
  aetiology, 51
    see also Carcinogens; Smoking
  biology, 451–3
    angiogenesis, 451–2
    apoptosis and growth, 451
    invasion and metastasis, 452, *452*
    proliferation, 451
    tumour immunology, 451
  causative factors, *450*
  dietary factors, 207
  epidemiology, 449–51, *449*
  epigenetics, 51
  familial syndromes, *453*
  gene therapy, 47
  genetic basis, **49–51**, 452–3, *453*
  histology, 455–6
  imaging, 456–7
  neuropathy, 1175
  oncogenesis, 49
  oncological emergencies, 458–9
  performance status, 456, *456*
  physical examination, 455
  screening, 454
    lead time bias, 454, *454*
    see also Genetic screening
  secondary *see* Metastases
  staging, 456, *456*
  symptoms, 454–5
  treatment *see* Cancer treatment
  see also Malignancy/malignant disease; Metastases; specific cancers; Tumours
Cancer treatment, **457–66**
  adjuvant therapy, 457
  aims, 457
  antimetabolites, 460
  antitubulin agents, 460
  biological therapy, **463–5**
    haemopoietic growth factors, 464
    immunotherapy, 465
    interferons, 463–4
    interleukins, 464
    intracellular signal inhibitors, 464, *464*
    monoclonal antibodies, 464, *464*
  curative, 457
  DNA damaging, 459–60, *459*
  DNA repair inhibitors, 460

Cancer treatment (*Continued*)
drug resistance, 462
endocrine therapy, 463, *463*
high-dose therapy, 462
measuring response to, 458, *458*
multiple courses, *459*
palliative, 458
palliative care, 494–7
radiation therapy, 465–6, *466*
combination chemo-radio-therapy, 466
side-effects, 466
remission, 458
side-effects, 460–2, *460*
bone marrow suppression and immunosuppression, 461
cardiotoxicity, 461
extravasation of intravenous drugs, 460–1
hair loss, 461
mucositis, 461
nausea and vomiting, 461
nephrotoxicity, 462
neurotoxicity, 462
secondary malignancies, 462
sterility and premature menopause, 462
solid tumours, **486–94**
see also individual tumours
see also Haemopoietic stem cell transplantation
Candesartan
heart failure, *739*
hypertension, 804, 1023
Candida, *70*, *82*, *90*, *191*
C. albicans, 149, 182, 1233–4, *1233*
septicaemia, *136*
treatment, 1234
vaginal discharge, *172*
see also Candidiasis
CAPD peritonitis, *637*
dermatophytosis, 153
HIV-associated, 200
hypersensitivity, 1194
oral, 247
Candidiasis, 149–50, 182–3
chronic mucocutaneous, 150
clinical features, 150, 182–3
diagnosis, 150, 183
oesophageal, 255
oral, 247
treatment, 150, 183
Cannabis, 1215
multiple sclerosis, 1153
poisoning, *935*, 942
Cannulae
internal jugular vein, *901*
radial artery, 900, *900*
Capecitabine, 460
breast cancer, 489
colorectal carcinoma, *461*
gastric cancer, 266, 490
oesophageal cancer, 489
Capillary leak syndrome, 114, 1249
treatment, 1250
Capillary wall, 649
Caplan's syndrome, 527, 871, 878
Capnocytophaga canimorsus, skin and soft tissue infections, *124*
Capnography, 911
Capreomycin
resistance, 866
tuberculosis, 866
Captopril
heart failure, *739*
hyperaldosteronism, 1024
hypertension, 804
side effects, membranous glomerulopathy, 588
Carbamate insecticides, *935*, 942
antidotes, *937*
Carbamazepine
cytochrome P450 induction, *925*
diabetes insipidus, *976*, 1018
diabetic neuropathy, 1055
epilepsy, *1143*
dose, *1143*
therapeutic range, *1143*
mania, 1206
overdose, *935*, 940
prostadynia, 603
side-effects, *1143*
acute tubulointerstitial nephritis, *604*
rash, *1265*
teratogenesis, *924*
therapeutic drug monitoring, *927*
trigeminal neuralgia, 1107
Carbapenems, 95–6
Carbaryl, 183
Carbenoxolone
side-effects
endocrine disease, *964*
hypertension, 799
Carbimazole, 989, *989*
side-effects
agranulocytosis, 989
bone marrow aplasia, 402
hepatotoxicity, *366*
teratogenesis, *924*
Carbocisteine, COPD, 839
Carbohydrate
absorption, 271–2, *273*
dietary fibre, 211
dietary requirements, 211–12
diabetes mellitus, 1038
intake, 211, 212
metabolism, 321
inborn errors, 1068–9, *1069*
Carbon dioxide
dissociation curve, 893, *893*
partial pressure, 814, 893
COPD, 836
Carbon monoxide, 959
poisoning, *935*, 942, *1114*
antidotes, *937*
metabolic acidosis, *939*
Carbon monoxide transfer test, 852
Carbon tetrachloride
acute tubular necrosis, 617
hepatotoxicity, *366*
Carbonic acid, 672
Carbonic anhydrase, 672
Carbonic anhydrase inhibitors, clinical uses, 658
Carboplatin, 460
breast cancer, 488
germ cell tumours, 492
lung cancer, *461*, 486, 487
ovarian cancer, 491
γ-Carboxyglutamate, 219
Carboxyhaemoglobin, 942
Carbuncles, *86*, 1229
renal, 603
Carcinoembryonic antigen, *493*
colorectal carcinoma, 304
Carcinogens, 51
Carcinoid syndrome, 222, 284, 1145
Carcinoid tumours, 284–5
bronchial, 884
Carcinoma *see* Cancer; and individual carcinomas
Carcinoma embryonic antigen, *456*
Carcinoma embryonic antigen-related cell adhesion molecule 6, 286
Cardiac arrest
unexpected, 711, *711*
ventricular tachycardia, 726
see also Myocardial infarction
Cardiac arrhythmias, **715–33**
bradycardia *see* Bradycardias
long-term management, *see* Antiarrhythmic drugs
mechanisms, 716–17, *716*
accelerated automaticity, 716, *716*
re-entry (circus movements), 716–17
triggered activity, 716–17, *716*
overdose/poisoning, 936
palpitations, 688–9, 722
post-STEMI, 759
sinus node, 716
Stokes-Adams attacks, 689
syncope, *see* Syncope
tachycardia *see* Tachycardias
see also Atrial tachyarrhythmias; Ventricular tachyarrhythmias
Cardiac biopsy, 737
Cardiac catheterization, 708
aortic regurgitation, 767
aortic stenosis, 766
aortic valve disease
constrictive pericarditis, 797
Fallot's tetralogy, 779
heart failure, 737
mitral regurgitation, 764
mitral stenosis, 762
mitral valve disease
primary restrictive non-hypertrophic cardiomyopathy, 793
pulmonary stenosis, 769
right-heart bedside, 714
Cardiac complications
HIV, 189
myocardial infarction, 758–9
post-transplantation, *741*
Cardiac conduction, post-STEMI disturbances, 759
Cardiac cycle, *683*, 684, *684*
Cardiac disease *see* Cardiovascular disease; Heart disease
Cardiac enzymes, myocarditis, 788
Cardiac failure *see* Heart failure
Cardiac glycosides, heart failure, 739, *739*
Cardiac ischaemia, *251*
Cardiac masses
echocardiography, 704
magnetic resonance imaging, 707
Cardiac output (CO: Q1), 891, *891*
heart failure, 734
high (increased)
measurement, 903–4, *904*
echocardiography, 904
lithium dilution, 904
oesophageal Doppler, 904, *904*
pulsed-wave Doppler, 699
reduced/reduction
total, 890
Cardiac pacemakers, 713–14, *713*
biventricular, 714
heart failure, 741
choice of, 714
dual chamber, 714
rate responsive, 714
single chamber, 714
Cardiac pacing
permanent, 713–14
temporary, 713, *713*
transcardiac, 713
transvenous, 713
Cardiac resuscitation *see* Resuscitation
Cardiac status, grading system, 688, *688*
Cardiac syndrome X, 749
Cardiac tamponade, 796
cancer-related, 459
haemodynamic changes, *899*
Cardiac transplantation, 741, *741*
complications, *741*
contraindications, *741*
see also Heart-lung transplantation
Cardiac vectors, 696, *696*
Cardiogenic shock, *742*, *894*
clinical features, 898
haemodynamic changes, *899*
Cardiomegaly, 655
sickle cell disease, 410
Cardiomyopathy, **788–94**, *789*
acquired, 794
arrhythmogenic right ventricular, *789*, 791
conduction system disease, 792
dilated, 792, *792*, *793*
echocardiography, 704
echocardiography, 704
hypertrophic, 788–91, *788*
clinical features, 788–9
echocardiography, 704
investigations, 790, *790*
signs, 789–90
symptoms, 789
treatment, 790–1
ion channelopathies, 792
iron overload, 410
left ventricular non-compaction, *789*, 791–2
magnetic resonance imaging, 707
peri-partum, 794
primary restrictive non-hypertrophic, 792–4
clinical features, 793
investigations, 793–4
treatment, 794
restrictive, echocardiography, 704
right bundle branch block, *720*
stress, 794, *794*
tachycardia, 794
Cardiorespiratory assessment, *824*
Cardiothoracic ratio, 692
Cardiotoxicity, chemotherapy-induced, 461
Cardiovascular disease, **681–810**
chronic renal failure, 630–1
coronary artery calcification, 631
pericarditis, 631
risk failure, 630–1
HIV-related, 189
homocysteine and B vitamins, 224
investigations, *see specific techniques*
ischaemic *see* Ischaemic heart disease
prevention, 747
renal transplant-associated, 642
risk prediction, 747, *748*
see also Heart disease
Cardiovascular support
see also Cardiac resuscitation
Cardiovascular system
diabetic neuropathy, 1056
effects of mechanical ventilation, 914
examination, 690–2
arterial pulse, 690–1, *690*
blood pressure, 691, *691*
jugular venous pressure, 691–2, *691*
precordium, 692
systemic lupus erythematosus, 543
see also Circulation; Heart
Cardioversion *see* DC-cardioversion
Cardioverter-defibrillator *see* Implantable cardioverter-defibrillator (ICD)
Carditis, 137
Carina, 811, 818
Carmustine, 460
Caroli's disease, 360, 368, 371
beta-Carotene, 218
Carotene dioxygenase, 218
Carotid artery dissection, *1110*
pathophysiology, 1126
Carotid pulse, 690
Carotid sinus
massage, 697, *698*
sensitivity, 717
syncope, 689
Carotid sinus syndrome, 717

Carpal tunnel syndrome, 501, 508–9, 1116, 1172
  rheumatoid arthritis, 528
Carpometacarpal osteoarthritis, 521
Carrier proteins, *22*
Carrión's disease, *82*, 148, *148*
Carteolol, glaucoma, 1091
Cartilage, 519
  hyaline, 500
Carvajal syndrome, 791
Carvedilol, heart failure, *739*
Casal's necklace/collar, *222*
Case-control studies, 931
Caseation, 864
Caspases, 28, *28*, 66, 451
Caspofungin, 100, *100*
Castleman's disease, 108, 479, 484
Casts, urinary
    granular, 627
    red-cell, *579*, 627
    white-cell, 579
Cat scratch disease, *80*, 125, *148*
Catalase, *224*
Cataplexy, 1145
Cataracts, 1090–1, *1091*
  aetiology, 1090, *1091*
  clinical features, 1091
  diabetes mellitus, 1050
  extraction, 1052
  investigations, 1091
  treatment, 1091
Catecholamines
  cardiac contraction, 684, *684*
  excess/increased
  and metabolism, 214
  resting plasma, 1025
  synthesis and metabolism, *1025*
  urinary, 1025
  see also Adrenaline (epinephrine); Dopamine; Norepinephrine (noradrenaline)
β-Catenin, 382
  haemopoiesis, 388
Catenins, 583, 655
Cathartics, 938
Cathelicidins, 274–5, 1226
Cathepsin(s), 56, *57*
Cathepsin K, 558
Catheter ablation, 723, 731–2
  atrial flutter, 725
Catheter-related sepsis, *136*, 235–6
  bladder catheterization, 603
Catheterization
  bladder, 1165
  cardiac *see* Cardiac catheterization
Catheters
  balloon flotation, 902, *903*
  Swan-Ganz, 902
Cauda equina syndrome, 512, 1177
Cavernous haemangioma, 1267–8, *1268*
Cavernous sinus, 1106
  thrombosis, *1102*
Cawthorne-Cooksey exercises, 1109
CCL11, 57
CCL21, *67*
CCL24, 57
CCR2, *67*
CCR5, 185
CCR6, 1226
CCR7, *67*
CD1c, *59*
CD2AP, 583
CD3, systemic sclerosis, 545
CD4
  antibodies, 523
  systemic sclerosis, 545
  see also T cell(s), helper
CD5, systemic sclerosis, 545
CD8, *see* T cell(s), cytotoxic
CD11a, 67
CD14, 58
CD16, 58, 66
CD18, 67
CD20 positive B cell ablation, 524
CD25, 66, 76
CD34, 47
CD38, 47
CD40, 70
CD44, 500
CD55, 418, 419
CD56, 66
CD59, 418, 419
CD71, IgA nephropathy, 593
CD80, *59*
  focal segmental glomerulosclerosis, 586
CD86, *59*
CD89, IgA nephropathy, 593
CD95, 28
CD117, *493*
CD123, *59*
CD154, 70
Cefaclor, *96*
  bronchiectasis, 843
  COPD, 839
  empyema, 863
Cefadroxil, *96*
Cefalexin, *96*
  urinary tract infections, 603
Cefamandole, *96*
Cefepime, *96*
Cefixime, *96*
  COPD, 839
  gonorrhoea, 176
Cefotaxime, *94*, *96*
  meningitis, *1155*
  spontaneous bacterial peritonitis, 353
Cefoxitin, *96*
Cefpirome, *96*
Cefpodoxime, *96*
Cefprozil, *96*
Cefradine, *96*
Ceftazidime, *96*
  bronchiectasis, 843
  epiglottitis, 834
  spontaneous bacterial peritonitis, 353
Ceftriaxone, *93*, *96*
  chancroid, 180
  donovanosis, 180
  gonorrhoea, 176
  side-effects, hepatotoxicity, *366*
Cefuroxime, *96*
  empyema, 863
  pancreatitis, 379
  pneumonia, 862, 863
  urinary tract infections, 602
Celecoxib, musculoskeletal problems, 517, *518*
Cell, *19*, **19–29**
  chromosomes, *19*
  cytoplasm, *19*, 22–3
  cytoskeleton, *19*
  endoplasmic reticulum, *19*, 22–3, *23*
  Golgi apparatus, *19*, 23
  intercellular connections, 26
  intracellular transport, 21, *21*
  lysosomes, *19*, 21, 23
  mitochondria, *19*, 23–4
  nuclear envelope, *19*
  nucleolus, *19*
  nucleus, *19*, **27–9**
  peroxisomes, 23
  ribosomes, *19*
  secondary messengers, 24
  see also individual cell types
Cell adhesion molecules (CAMs), 20, *25*, 26–7, 452
  cadherins, *25*, 26
  immunoglobulin superfamily, *25*, 26–7
  integrins, *25*, 26
  selectins, *25*, 27
  see also individual molecules
Cell cycle, 27–9, *27*
  apoptosis, 27–9, *28*
  cyclin, 27
  cyclin-dependent kinases, 27
Cell death, programmed *see* Apoptosis
Cell lysis, 85
Cell membrane, **19–22**, *20*, 649
  dynamics, 19–21
  exocytosis, 21
  phagocytosis, 21
  pinocytosis, 21
  proteins, *19*, 20, *20*
  receptors, *22*
  transport, 21–2
Cell migration, 66–8, *67*
  dendritic cells, *67*
  neutrophils, *68*
Cell *see also individual cell types*
Cell volume, 655
Cell-mediated immunity, 140, *140*, 188, *867*, 1231
  defective/deficiency, 118, 355, 1170
Cellular naevi, 1257
Cellular theories of ageing, 226–7
Cellulitis, *86*, 1229
  crystal, 538
Cellulose, 211
Central nervous system (CNS)
  chronic renal failure, 630
  disease *see* Neurological disease
  malignancy
    metastases, 457
    see also Brain tumours
  radiation effects, 958
  see also Brain; Spinal cord
Central pontine myelinolysis, 663
Central post-stroke pain, 1118
Central retinal vein thrombosis/occlusion, *1102*
Central sleep apnoea syndrome, 688, 842, 1145
Central tolerance *see* Tolerance
Central venous pressure (CVP), 901–2, *901*, *902*
  problems in measuring, 902
  Seldinger technique, *902*
Centri-acinar emphysema, 836
Cephalic tetanus, 134
Cephalosporins, 95, *95*, *96*
  side-effects
    acute tubulointerstitial nephritis, *604*
    respiratory, *877*
  urinary tract infections, 602
  see also individual drugs
Cephamycins, 95, *95*
Cerebellar ataxia, gait, *1096*, 1097
Cerebellar degeneration, *455*
Cerebellar haemorrhage, 1134
Cerebellar tonsils, 1114
Cerebellopontine angle, 1106, 1107, 1109
  lesions, 1106
Cerebellum, 1109, 1111, *1111*, 1113–14, *1114*
  lesions, 1114, *1114*
Cerebral amyloid angiopathy, 1134
Cerebral amyloidosis, 1073
Cerebral arteries, *1128*
  middle, infarct, *1129*
  thrombosis, 442
Cerebral blood flow, *1127*
  autoregulation, 1127–8
Cerebral cortex, 1099–100
  lesions, 1109
Cerebral dominance, 1099
Cerebral embolus, infective endocarditis, *770*
Cerebral haemorrhage
  in haemophilia A, 438
  see also Intracerebral haematoma/haemorrhage
Cerebral infarction, 1129–30, *1129*
  see also Stroke
Cerebral lymphoma, 483
Cerebral oedema, 1048, 1096
  fulminant hepatic failure, 343
  high-altitude, 956, *956*, *1114*
Cerebral palsy, *1114*, 1169
  athetoid, 1149
Cerebral-evoked potentials, 1121
Cerebroglycan, 500
Cerebrospinal fluid (CSF)
  changes/abnormalities, acute leukaemia, 469
  culture, *89*
  examination, 1121, *1121*, 1125
  normal, *1121*
  pleocytosis, 1156, *1156*
Cerebrovascular disease, **1126–36**
  see also Stroke; Transient ischaemic attacks (TIAs)
Certolizumab, Crohn's disease, *289*, 290
Cervical cancer
  epidemiology, *449*
  geographical distribution, 451
  screening, 454
Cervical discs
  prolapse, 506
  protrusion
    central, 1177, *1177*
    lateral, 1176, *1176*
Cervical lymph node enlargement, differential diagnosis, *476*
Cervical spondylosis, 506
Cervix
  carcinoma *see* Cervical cancer
  strawberry, 182
Cestodes (tapeworms), *163*, 170–1
  Taenia saginata, *163*, 170
  Taenia solium, *163*, 170–1
Cetirizine, rhinitis, 833
Cetuximab, *464*
  colorectal carcinoma, 490
CFU-Eo, 388
CFU-GEMM, 388
CFU-*GM*, 388
CFU-*Meg*, 388
Chagas' disease, 80, 158, 788, *1179*
  achalasia, 254
  clinical features, 158
  diagnosis, 158
  insect vectors, *82*
  management and control, 158
  megacolon, 297
  right bundle branch block, *720*
  Romaña's sign, 158
Chalazion, 1087, *1087*
Chancre, *172*
Chancroid, 179–80
  clinical features, 180
  diagnosis and treatment, 180
  genital ulceration, *172*
Chandipura virus, 1156
Channel proteins, *22*
Channelopathies, 1178, 1182
Charcoal, activated
    multiple-dose, 938
    single-dose, 937–8
Charcot-Bouchard aneurysm, 1133–4
Charcot-Marie-Tooth disease (HMSN), 26, 39, 1175
Charcot's joints, 540, 555–6, 1055, 1117
Chédiak-Higashi syndrome, 71, *430*
  host factors, *86*
Cheese washer's lung, *876*
Cheiroarthropathy, 1254
Chelation therapy, 945
Chemo-immunotherapy, 481
Chemo-radiotherapy, 466
  colorectal carcinoma, 490
  lung cancer, 486
Chemoattractants *see* Chemokines
Chemokines, 55, 56, 65
  pleiotrophy, 55
Chemoreceptor trigger zones, 242
Chemotaxins, 430
Chemotaxis, 55

Chemotherapy
antimicrobial, **91–102**
cancer *see* Cancer treatment
side-effects, endocrine disease, *964*
tuberculosis, 100
see also specific drugs
Chenodeoxycholic acid, 308, 322, 370
Cherry angioma, 1258
Chest
examination, 820–1, *820*, *821*
surface anatomy, *813*
X-ray *see* Chest X-ray
Chest compressions, 710
Chest pain, 516, 688, *688*, 820
acute coronary syndrome, 752
angina, *see* Angina pectoris
aortic dissection, 808
cor pulmonale, 783
functional, 310–11
mitral valve prolapse, 764
myocardial infarction, 755
pulmonary embolism, 784
tuberculosis, 865
Chest wall
disorders, 884–6
trauma, 884
Chest X-ray, 244, 692–4, *692*, *813*, 822–3, *822*
achalasia, 253
Addison's disease, 1011
aortic regurgitation, 767
aortic stenosis, 766
ARDS, *919*
asthma, 853
atrial septal defect, 775
bronchial carcinoma, 882
bronchiectasis, 843
calcification, 693
constrictive pericarditis, 797, *797*
COPD, 838
cor pulmonale, 783
Cushing's syndrome, 1013
dilated cardiomyopathy, 792
extrinsic allergic alveolitis, 877
Fallot's tetralogy, 779
glomerular disease, *585*
goitre, 993
heart failure, 737
heart size, 692–3
Hodgkin's lymphoma, 477, *477*
idiopathic pulmonary fibrosis, 872
infective endocarditis, 772
lung fields, 693–4, *694*
Marfan's syndrome, 781
mitral regurgitation, 763
mitral stenosis, 762
myocarditis, 788
neurology, *1122*
non-Hodgkin's lymphoma, 481
pancreatitis, 378
pericardial effusion, *796*
pneumonia, 858–9, *858*
primary restrictive non-hypertrophic cardiomyopathy, 793
pulmonary embolism, 785
pulmonary hypertension, 783
pulmonary stenosis, 768
tricuspid stenosis, 768
tuberculosis
adult post-primary, 865, *865*
miliary, 865
ventricular septal defect, 775
Cheyne-Stokes respiration, 688, 1124
Chi-squared test, 933
Chickenpox *see* Herpes zoster virus; Varicella (chickenpox)
Chikungunya, 541
Chilblain lupus, 1252
Children
abuse, 1210
akinetic-rigid syndromes, 1148
arthritis, 553–4
athetoid cerebral palsy, 1149
competence to give consent, 5–6
dermatomyositis, 547
epilepsy, *see* Epilepsy
failure to thrive, 214
Henoch-Schönlein purpura, 554
hypermobility syndrome, 554
hyperthyroidism, 988
hypothyroidism, 985
idiopathic musculoskeletal pain, 554
immune thrombocytopenic purpura, 436
low back pain, 554
malnutrition, *216*
care setting, 217
prevention, 217
prognosis, 217
refeeding, 216, *217*
rehabilitation, 217
vitamin A in, 219
minimal-change nephropathy, 585–6, *586*
obstructive megaureter, 616
osteochondritis, 554
Perthes' disease, 554
prescribing in, 924
psoriatic arthritis, 553
rheumatic fever, 554
rickets *see* Rickets
transient synovitis of hip, 554
undernutrition, 215
Wilson's disease (hepatolenticular degeneration), 1148
see also Infants; Neonates
Chillblains, 955
Chin lift, 710
Chinese herb nephropathy, 606
Chinese restaurant syndrome, 237
Chitinase, 848
Chitins, 848
Chitotriosidase, 1071
Chlamydia, *123*
C. pneumoniae, 128
pneumonia, *858*, 860
C. psittaci, *80*, 128
pneumonia, *858*, 860
C. trachomatis, 143, 176
clinical features, 176
conjunctivitis, 1088
diagnosis, 176
non-specific urethritis, 535
nucleic acid amplification test, 175, 176
screening for, **174**
treatment, 176
urethral discharge, *172*
vaginal discharge, *172*
infective endocarditis, 770
visual loss, 1094
Chlamydial conjunctivitis, 1088
Chloasma, 1265
Chlorambucil, 459
acute lymphocytic leukaemia, 475
focal segmental glomerulosclerosis, 587
lymphoplasmacytic lymphoma, 481
membranous glomerulonephropathy, 588
Chloramphenicol, *94*, 97–8, *97*
epiglottitis, 834
indications, 97
interactions, 97
mechanism of action, 97
meningitis, *1155*
side-effects
bone marrow aplasia, 402
haemolysis in G6PD deficiency, *414*
neuropathy, *1174*
structure, *97*
toxicity, 98
Chlorhexidine gluconate, 247
Chloride
dialysate
haemodialysis, *634*
peritoneal dialysis, *637*
interstitial fluid, *650*
intracellular fluid, *650*
losses, 679
from gut, *662*
plasma, *650*
2-Chloroadenosine acetate, hairy cell leukaemia, 475
Chloroquine, *156*, *157*
overdose, 940
resistance, 156
side-effects
haemolysis in G6PD deficiency, *414*
inclusion bodies, 597
neuropathy, *1174*
rash, *1265*
systemic lupus erythematosus, 544
Chlorphenamine, rhinitis, 833
Chlorproguanil, *156*
Chlorpromazine
amfetamine overdose, 940
schizophrenia, 1217
side-effects
bone marrow aplasia, 402
endocrine disease, *964*
hepatotoxicity, *366*
long QT syndrome, *727*
rash, *1265*
Chlorpropamide
diabetes insipidus, *976*, 1018
diabetes mellitus, *1038*
side-effects, endocrine disease, *964*
Chlortalidone, *657*
Chloruretic agents, 679
Chokes, 957
Choking, *1128*, 1145
Cholangiocarcinoma, 364, 374
Cholangiography
endoscopic retrograde (ERC), common bile duct stones, 371, *371*
magnetic resonance (MRC), 370, *371*
percutaneous transhepatic (PTC), 327
Cholangitis, 327
primary sclerosing *see* Primary sclerosing cholangitis (PSC)
Cholecalciferol, 323, 559
Cholecystectomy, 369
laparoscopic, 372
post-cholecystectomy syndrome, 370
Cholecystitis
acute, 368
cholecystectomy, 369
chronic, 372
non-calculous, 372
see also Gallstones
Cholecystokinin, 229, *272*, 322, 375
Choledochal cyst, 373
Cholera, *83*, 143–4
antibiotics, *132*
clinical features, 143–4
diagnosis, 144
El Tor biotype, 143
immunization, *103*, 144
management, 144
intravenous rehydration, 144
oral rehydration solutions (ORS), *132*, 144
Cholera toxin, *306*
Cholestasis
benign recurrent intrahepatic, 330
blood picture, 325
γ-glutamyl transpeptidase in, 324
intrahepatic, of pregnancy, 362
progressive familial intrahepatic, 330
Cholestatic jaundice (acquired), 330–2, *330*
causes, *330*
clinical features, 331, *331*
differential diagnosis, 330–1
investigations, 331–2, *332*
vitamin K deficiency in, 220
see also Cholestasis; Jaundice
Cholestatic viral hepatitis, 335
Cholesteatoma, 1078
Cholesterol, 211, 321
and coronary artery disease risk, *1063*
diet low in, 1065
see also Lipid-lowering therapy
dietary intake, 1065
familial hypercholesterolaemia *see* Familial hypercholesterolaemia
high-density lipoprotein *see* High-density lipoprotein cholesterol
intake, 212
primary biliary cirrhosis, 355
reduction, *see* Lipid-lowering therapy
synthesis, *322*
total, and LDL, 1062
Cholesterol absorption inhibitors, *1066*
Cholesterol embolization, 609, *609*
Cholesterol gallstones, 367
risk factors, *367*
Cholesterol-binding resins, *1066*
Cholesterolosis, 372
Cholesteryl ester transfer protein inhibitors, *1066*
Cholic acid, 322, 367
Cholinergic nerves *see also* Acetylcholine
Chondrocalcinosis, 356, 521, *522*, 1020
elderly patients, 553
Chondrocytes, 500
Chondroitin sulphate, 500, 519
supplements, 522
Chondroma, 884, 974
Chondromalacia patellae, 513
Chondrophytes, 518
Chorea, 1111, 1112, 1149
causes, *1149*
Chorionic villus sampling, 45
Christmas disease, 439
Chromium, *226*
allergy, *848*
recommended IV intake, 233
Chromoblastomycosis, *150*, 153
Chromogranin, *272*, *493*
Chromogranin A, 1025
Chromosomal disorders, 38–40, *40*
abnormal numbers, 38–9
abnormal structure, 39
analysis of, 39–40
mitochondrial, 39
trisomy 21, *30*, 38
see also specific disorders/abnormalities
Chromosomal translocations, 49, *453*
Chromosomes, *19*, *30*, 33, **36–7**
analysis, *37*
cell division, *see* Cell cycle
constitution, *30*
disorders *see* Chromosomal disorders
human, 36
mitochondrial, 37, *37*
see also Mitochondrial genome
nomenclature, *30*
sex determination *see* Sex chromosomes
telomeres, 37
X chromosome, 37
Chronic actinic dermatitis, 1249
Chronic allograft nephropathy, 641
Chronic bullous dermatosis of childhood, 1256–7

Chronic discoid lupus erythematosus, 1252, *1252*
Chronic fatigue syndrome (CFS), 516, 1192–3
  aetiology, 1192–3, *1192*
  management, 1193
  prognosis, 1193
Chronic granulomatous disease, 71
Chronic hepatitis *see* Hepatitis
Chronic inflammatory demyelinating polyradiculoneuropathy (CIDP), 1173
Chronic lymphocytic leukaemia (CLL) *see* under Leukaemia
Chronic myeloid leukaemia (CML) *see* under Leukaemia
Chronic non-suppurative destructive cholangitis *see* Primary biliary cirrhosis (PBC)
Chronic obstructive pulmonary disease (COPD), **835–42**
  acute exacerbation, 841, *841*
  air travel, 840
  airways in, 815, 816
  definition, 835
  diagnosis, 837, *838*
  epidemiology and aetiology, 835
  haemoptysis, 819
  investigations, 838
  management, 837–8, *837*
    drug therapy, 838–9
    pulmonary rehabilitation, 840
    smoking cessation, 838, *839*
  nocturnal hypoxia, 839–41
  pathogenesis, 837
  pathophysiology, 835–6, *835*, *836*
  physical signs, *821*
  prognosis, 841–2, *842*
  signs, 837
  surgery, 840–1
  symptoms, 837
Chronic pain syndromes, 516–18, 1193–4
  aetiology, 1193
  chronic regional pain syndrome type 1, 517
  management, 1193–4, *1194*
  temporomandibular pain dysfunction syndrome, 517
  upper limb, work-related, 516–17
  widespread, 510
  widespread pain, 516
  see also Chronic fatigue syndrome (CFS)
Chronic renal failure (CRF), **625–42**
  causes, *625*
  complications, 628–31
    anaemia, 628
    calciphylaxis, 630
    cardiovascular disease, 630–1
    endocrine, 629–30
    gastrointestinal, 629
    malignancy, 631
    management, 631
    metabolic, 629
    muscle dysfunction, 630
    nephrogenic systemic fibrosis, 629
    nervous system, 630
    renal osteodystrophy, 628–9
    skin disease, 629
  early referral, 634
  end-stage, 631
  examination, 626–7, *627*
  history, 626
  investigations, 627–8
    haematology, 627
    immunology, 627
    radiology, 628
    renal biopsy, 628
    serum biochemistry, 627
    urinalysis, 627
    urine biochemistry, 627
    urine microscopy, 627
  management, 632–4
    correction of complications, 632–3
    renoprotection, 632, *632*
  progression, 631–2
  renal replacement therapy, 634–9
  staging, *626*
  symptoms, 626
Chrysops, 165
Churg-Strauss syndrome, 504, *549*, 596, 870–1
  ANCA antibodies in, 596
  clinical features, *550*
  diagnosis, *550*
  eosinophilia, 503
Chuvasch polycythaemia, 420
Chvostek's sign, 1021, 1022
Chylomicrons, 1061, *1061*, 1063
  disorders of, 1063
Chylothorax, 886
Chymotrypsin, 273
  faecal, 376
Chymotrypsinogen, 375
Ciclosporin (cyclosporin), 76
  aplastic anaemia, 403
  asthma, *854*, 855
  atopic eczema, 1238
  autoimmune hepatitis, 344
  Behçet's disease, 552
  chronic discoid lupus erythematosus, 1252
  Crohn's disease, *289*
  dermatomyositis, 1251
  focal segmental glomerulosclerosis, 587
  genetic polymorphism in metabolism, *926*
  IgA nephropathy, 594
  membranous glomerulonephropathy, 588
  minimal-change nephropathy, 586
  paroxysmal nocturnal haemoglobinuria, 419
  pemphigus vulgaris, 1256
  polymyositis/dermatomyositis, 548
  psoriatic arthritis, 535
  pyoderma gangrenosum, 1251
  renal transplantation, 640
  rheumatoid arthritis, 532
  side-effects
    haemolytic uraemic syndrome, 598
    hepatotoxicity, *366*
    hypertrichosis, 1003
  therapeutic drug monitoring, *927*
  thrombocytic purpura, 436
Cicuta, 952
Cicutoxin, 952
Cidofovir, 101, *101*
  cytomegalovirus, 107
Cigarette smoking *see* Smoking
Ciguatera fish poisoning, 950
Ciguatoxin, *133*
Cilia
  immotile cilia syndrome, 818
  nasal, 814
  trachea, 811–12, *812*
Ciliary dyskinesia, 818
Cimetidine, 252
  cytochrome P450 inhibition, *925*
  side-effects
    acute tubulointerstitial nephritis, *604*
    hepatotoxicity, *366*
Cinacalcet, parathormone suppression, 633, 1020
Cinnarizine, 1080
Ciprofibrate, hyperlipidaemia, *1066*
Ciprofloxacin, *93*, *94*, 98
  anthrax, 142
  bronchiectasis, 843
  chancroid, 180
  Crohn's disease, *289*
  gastroenteritis, *132*
  gonococcal arthritis, 540
  gonorrhoea, 176
  H. pylori eradication, 261
  pneumonia, 863
  prostatitis, 603
  side-effects
    acute tubulointerstitial nephritis, *604*
    long QT syndrome, *727*
  spotted fevers, 148
  tuberculosis, 866
  urinary tract infections, 602
Circadian rhythm
  bronchomotor tone, 815
  cortisol levels, 1012–13
  hormone release, 967
  reverse, in infertility treatment, 1005
Circinate balanitis, 536
Circle of Willis, 1127, *1127*
Circulation, 710–11
  fetal, 687–8, *687*
Circumflex artery, 685
Circus movements, 716–17, *716*
Cirrhosis, **345–7**
  aetiology, 345
  alcoholic *see* Alcoholic cirrhosis
  causes, *345*
  compensated, 338
  complications, 348–54, *348*
    ascites, 351–3
      see also Ascites
    hepatopulmonary syndrome, 354
    porto-pulmonary hypertension, 354
    portosystemic encephalopathy, 353–4
    renal failure (hepatorenal syndrome), 354
    see also Portal hypertension; Variceal haemorrhage
  course, 347
  decompensated, 338
  extracellular volume regulation, 653
  hepatomegaly, *331*
  imaging, 346–7
  Indian childhood, 364
  investigations, 346
  liver biopsy, 347
  liver transplantation, 347–8
  macronodular, 346
  management, 347
  micronodular, 345
  oedema, 655
  pathogenesis, 345, *345*
  pathology, 345–6, *346*
  primary biliary *see* Primary biliary cirrhosis (PBC)
  prognosis, 347, *347*
  scoring systems, *347*
  secondary biliary, 356
  severity, 346
  type, 346
Cisplatin, 460
  colorectal carcinoma, 490
  gastric cancer, 266, *461*, 490
  germ cell tumours, 492
  lung cancer, *461*, 486
  oesophageal cancer, 489
  side-effects
    haemolytic uraemic syndrome, 598
    nephrotoxicity, 462
    neuropathy, *1174*
Citalopram, overdose, 949
Citric acid cycle, 210
Civil law, 2
  negligence, 3
CJD *see* Creutzfeldt-Jakob disease
Cladribine, hairy cell leukaemia, 475
Clara cells, 811
Clarithromycin, 97
  cytochrome P450 inhibition, *925*
  H. pylori eradication, 261
  pneumonia, 862
  tuberculosis, 866
Clasp-knife effect, 1112
Class switching, 64
Clathrin, 21
Claudication
  coarctation of aorta, 777
  intermittent, 805
  spinal canal, 806
Claudin(s), 26
Claudin 19, 669
Clavulanic acid, 95
Cleft palate, inheritance, *44*
Climate change, 953
Clindamycin, 99
  babesiosis, 160
  bacterial vaginosis, 183
  pitted keratolysis, 1230
  septic arthritis, 539
  toxoplasmosis, 201
Clinical choices
  ethical consideration, 1
  patient's best interest, 6
  see also Ethics
Clinical error, 3
  negligence vs., 3
Clinical negligence, 3
  error vs., 3
  evidence, 3
  irrelevancy of experience, 3
  withholding information about risk, 4
Clinical practice
  clinical importance of trust, 2
  communication *see* Communication
  consumerism, 9
  doctor-patient relationship, 2
    failure, 10, 14–15
    good, 9–10
    importance of communication, 9–10
  duties of care, 1–2
    autonomy, respect for, 1–2, 3–6
    equality, 7–8
  good medical practice (GMP), 2
  legal issues *see* Legal issues
Clinical resources, 7–8
Clinical training, communication skills, 14, 18–19
Clinical trials
  controlled observational studies, 931
  dangers, 932
  new drug evaluation, 932, *932*
  randomized controlled trials, 924, 929–31
  statistical analyses, *see* Statistical analyses
  uncontrolled observational studies, 931
Clinician satisfaction, 9–10
Clobazam, epilepsy, *1143*
Clobetasol propionate, *1237*
  lichen planus, 1245
  sarcoidosis, 1253
Clobetasone butyrate, *1237*
Clodronate
  breast cancer, 489
  hypercalcaemia, 458
Clofazime, leprosy, *141*
Clomifene, infertility, 1005
Clomipramine, narcolepsy/cataplexy, 145
Clonazepam
  epilepsy, *1143*
  side-effects, hepatotoxicity, *366*
  trigeminal neuralgia, 1107
Clonidine, hypertension, 805
Clonidine suppression test, 1025
Cloning, 35
  in silico, 36
Clonorchis sinensis, *163*, 170
  cholangiocarcinoma, 364

Clopidogrel, 687
  acute coronary syndromes, 753, *754*
  arterial thrombosis, 443, *443*
  myocardial infarction, 759
  side-effects, thrombotic thrombocytopenic purpura, 437
  stroke, 1132
Clostridial infections, 131
Clostridium, 122
  C. botulinum, 85, *123*, 134
    diarrhoea, *307*
    food poisoning, *133*
    see also Botulism
  C. difficile, 88, 131
    antibiotics, *132*
    diarrhoea, *307*
    pseudomembranous colitis, *132*, 307
  C. perfringens, *82*, 85, *123*, *1179*
    diarrhoea, *307*
    food poisoning, *133*
    gastroenteritis, *128*
  C. tetani, *82*, 85, *123*, 134
  see also Tetanus
Clotrimazole, 100, *100*
  candidiasis, 183
  ringworm, 1233
Clotting *see* Coagulation
Clotting factors *see* Coagulation factors
Clozapine
  genetic polymorphism in metabolism, *926*, *927*
  schizophrenia, 1217–18
Clubbing (fingers) *see* Finger clubbing
Clue cells, 183
Cluster of differentiation (CD) classification, 53, *55*
  see also individual CD antigens
Cluster headache, 1138
CMV *see* Cytomegalovirus (CMV)
CNT 01275, Crohn's disease, *290*
Co-amoxiclav, 125
  COPD, 839
  pneumonia, 862, 863
  side-effects, hepatotoxicity, *366*
  sinusitis, 1081
  urinary tract infections, 602
Co-analgesics, 1118
Co-receptor blockers, *193*, 194
Co-transporters, 271
Co-trimoxazole
  donovanosis, 180
  gastroenteritis, *132*
  myelotoxicity, 189
  side-effects
    hepatotoxicity, *366*
    rash, *1265*, *1266*, 1269
  urinary tract infections, 603
Coagulation, 432–3, *433*
  disorders *see* Coagulation disorders
  factors *see* Coagulation factors
  in infection, *88*
  inhibitors, 441
    see also Anticoagulants/anticoagulation
  limitation of, 433–4
  mechanism, *433*
  pathway, 433
  tests *see* Coagulation screens/tests
  see also Haemostasis
Coagulation cascade, activation in shock, 897–8
Coagulation disorders (coagulopathies), 434
  acquired, 440–1
    coagulation inhibitors, 441
    DIC *see* Disseminated intravascular coagulation (DIC)
    liver disease, 440
    massive transfusion, 441
    vitamin K deficiency *see* Vitamin K, deficiency
  inherited, 438–40
    haemophilia A, 438–9
    haemophilia B, 439
    von Willebrand disease, 440
  screens *see* Coagulation screens/tests
  and venous thrombosis, 442–3
  see also Haemophilia(s); Thrombosis
Coagulation factors
  deficiencies, *see* Haemophilia(s); and individual factors
  synthesis of, 321
Coagulation screens/tests, 434–5
  activated partial thromboplastin time (APTT), 433, 435
  correction tests, 435
  prothrombin time *see* Prothrombin time (PT)
  thrombin time (TT), 435
Coagulopathy
  of cancer, 455
  transfusion-related, 906
Coarctation of aorta, *774*, 777–8
  clinical features, 777–8, *778*
  investigations, 778
  treatment, 778
Coartemether, *156*
Cobalamin *see* Vitamin B_{12} (cobalamin)
Cobalt, *226*
Cocaine, 1215
  poisoning, *935*, 942
    metabolic acidosis, *939*
  side-effects, hepatotoxicity, *366*
Cocaine-and-amfetamine-regulated transcript *see* CART
Coccidioides, *C. immitis*, *82*, 151, *191*
Coccidioidomycosis, *150*, 151–2
Cochlea, 1077
Cochlear nerve, 1108
  lesions, 1108
Cockroach allergy, 848
Cockroft-Gault equation, *575*
Codeine
  genetic polymorphism in metabolism, *926*
  musculoskeletal problems, *518*
Codeine phosphate, 307
Codon restriction theory of ageing, 226
Coeliac axis compression syndrome, 283
Coeliac axis nerve block, 381
Coeliac disease, **277–9**
  aetiology, 277–8
  clinical features, 278
  complications, 279
  diagnosis, 278–9
    absorption tests, 279
    biochemistry, 278
    bone densitometry, 278
    haematology, 278
    histology, 278, *279*
    HLA typing, 278
    radiology, 278
    serology, 278
    wireless capsule endoscopy, 279
  environmental factors, 278
  HLA system in, *60*
  immunology, 277–8
  inheritance, 278
  liver involvement, 365
  and small intestine tumours, 284
  splenic atrophy, 423
  treatment and management, 279
  see also Dermatitis herpetiformis
Cognitive behaviour therapy (CBT)
  anxiety, 1209
  depression, 1204
  obsessive-compulsive disorder, 1211
Cognitive deficits/impairment, Parkinson's disease, 1146
Cognitive state, 1188
  assessment, 1188–9
    see also Mental state examination (MSE)
Colchicine, 24
  aphthous ulcers, 247
  Behçet's disease, 552
  erythema nodosum, 1250
  familial Mediterranean fever, 556
  gout, 538
  pericarditis, 795
  primary restrictive non-hypertrophic cardiomyopathy, 794
Cold, **954–5**
  hypothermia, 954–5
Cold air, and asthma, 848–9
Cold haemagglutinin disease, 417
Cold injury, 955
  chillblains, 955
  frostbite, 955
  non-freezing, 955
Cold sores, 105, *105*, 331, 1230–1
Cold water immersion, 954
Colds *see* Common cold (acute coryza)
Colesevalam, hyperlipidaemia, *1066*
Colestipol, hyperlipidaemia, *1066*
Colestyramine, 282
  hyperlipidaemia, *1066*
  pruritus, 355
Colic
  biliary/gallstone, 368
  ureteric, 611–12
  see also Abdominal pain
Colipase, 375
Colitis
  ischaemic, 297–8
  microscopic inflammatory, 294
  total, 292
  ulcerative *see* Ulcerative colitis
Collagen, 500, 1227
  anchoring fibril, 568
  basement membrane, 568
  fibril-associated, 568
  fibrillar, 568
  filament-producing, 568
  network-forming, 568
  structure, 500
  synthesis
  types of, 500
Collagen defects, 568
Collagen disorders, 1263–4, *1264*
Collagen propeptides, type 1, 560
Collagenase, 56, *57*, 323, 430
  cirrhosis, 345
Collagenous colitis, microscopic, 294
Collapse, 1097
Collapsing pulse, 690
Collectins, *56*
Colles' fracture, 561
Colloid osmotic pressure, 321
Colloid solutions, 429
Colloids, 660, 906
  distribution, 650, *650*
Colon, **294–5**
  cancer *see* Colorectal carcinoma
  endoscopy, *see* Colonoscopy; Endoscopy
  physiology, 294–5, *295*
  sigmoid, examination, 243–4
  tumours, 300–2
    polyps and polyposis syndromes, 300–2, *301*
    see also Colorectal carcinoma
  see also specific disorders
Colonic pseudo-obstruction, 317, *317*
Colonoscopy, 246, *246*
  colorectal carcinoma, 304, 305, *305*
  Crohn's disease, 289
  diverticular disease, 298
  GI bleeding, 270
  ulcerative colitis, 291
  virtual, 244, 298, 305
Colony-forming units (CFU), 388
Colony-stimulating factors, 388–9, *388*
Colophony fumes, allergy, *848*
Colorectal carcinoma, 302–5, 490
  chemotherapy, *461*
  clinical features, 304
  epidemiology, *449*
  family history, 303, *303*
  follow-up, 305
  genetics, 302–3, *303*
  geographical distribution, 451
  investigation, 304
  lifetime risk, *304*
  pathology, 303–4, *304*
  prevalence, 302
  prevention and screening, 305
  prognostic factors, 490
  risk factors, *303*
  screening, 454
  staging, *305*
  survival, *305*
  treatment, 304–5, 490
Coma, 1123, 1129
  causes, 1123–4, *1123*
  diabetic *see* Diabetic coma
  Glasgow Coma Scale, *1123*
  investigations, 1125
  management, 1125
  mechanisms, 1123, *1123*
Combination therapy
Combivir, 195
Comedones, 1246
Commensal organisms, 600
Common bile duct stones, 327, 370–2
  differential diagnosis, 371
  investigations, 370
  management, 371–2
  physical examination, 370
Common cold (acute coryza), 819, 831
Common peroneal nerve palsy, 1172
Common variable immunodeficiency, 71
Common warts, 1231, *1231*
Communication, **8–17**
  complaints/lawsuits, 14–15
    resolution, 15
  difficulties for clinicians, 10–11
    attitude, 10
    distancing strategies, 10
    empathy, 10
    lack of knowledge, 10
    lack of skill, 10
    patients/topics, 10
    personal, 10
    time, 10
  difficulties for patients, 10, 11
    anxiety, 11
    conflicting information, 10
    disinclination to disclose concerns, 10
    forgetfulness, 10
    impaired communication, 10, 16
    inferiority, 10
    misconceptions, 10
  doctor-patient relationship
    good communication, 9–10
    in healthcare, 9–10
    poor communication, 10–11, 16
  failed communication, 10
    factors leading to, 10
    monetary costs, 10
  good communication, 9, 10
    qualities, 10
  healthcare, 8–11
    adherence to treatment, 9
    clinician satisfaction, 10
    health outcomes, 9
    social outcomes, 9
  improvement, 16–17
  medical interview *see* Medical interview
  patient-centred, 8–9

research, 18
training, 16–17
Compartment syndrome
heel, 516
lower limb, 807
Compensatory anti-inflammatory response syndrome, 895, *900*
definition, *899*
Competence
children, 5–6
definition, 5
English vs Scottish law, 6
other, 6
patient, 5–6
patient's best interest, 7
psychiatric illness, 6
refusal of treatment, 6, 7
Complement, 55, *56*, 418, 818
activation, 55, *56*
C3, 55, 418
measurement, *69*
C4, measurement, *69*
C5, 55, 418
chronic renal failure, 627
deficiency, *69*, 71–2
functions, 55
host defence, 600
iC3b, 56
lytic pathway defects, *70*
musculoskeletal problems, 504
regulatory proteins, 71–2
response to shock, 895
systemic lupus erythematosus, 543
Complementary DNA *see* DNA (deoxyribonucleic acid)
Component therapy, 429
Compound naevi, 1257
Compulsions, 1187
Computed tomography angiography, 1131
Computed tomography (CT)
abdominal, 244, *245*
brain, 1119–20
colonography, 244
GI bleeding, 270
heart, 705–6
coronary angiography, 706, *706*
coronary artery calcification, 705–6, *706*
Hodgkin's lymphoma, 477, *478*
kidney, 580
lungs, 823, *823*
pneumocolon, 244
small intestine, 277
specific conditions
ARDS, *917*
bronchial carcinoma, 882
bronchiectasis, 843, *843*
cancer, 457
chronic renal failure, 628
cirrhosis, 346, *346*
COPD, 838
diverticular disease, 298–9, *299*
epilepsy, 1141
extrinsic allergic alveolitis, 877
hepatocellular carcinoma, *364*
idiopathic pulmonary fibrosis, 872, *873*
liver disease, 326–7, *326*
musculoskeletal problems, 505
osteoporosis, 562
pancreatitis, *378*, *381*
pulmonary embolism, 785, *785*
retroperitoneal fibrosis, 617, *617*
sarcoidosis, *869*
stroke, *1130*, 1131
tuberculosis, 865
urinary tract infections, 602
spiral *see* Spiral computed tomography
unprepared, 244
GI bleeding, 270
Conditioning, classical
Conduction, cardiac *see* Cardiac conduction
Conduction system disease, 792
Condylomata lata, 178
Confidence intervals, 932
Confidentiality, 4–5
acceptable breaches, 5
respecting confidentiality in practice, 5
Conflicting information, communication difficulties for patients, 10, 11
Confusion, 1123
Confusional states, *1099*
Congenital adrenal hyperplasia, *1007*, 1014–16
clinical features, 1014
investigations, 1014
pathophysiology, 1014
treatment, 1014–15
Congenital antibody deficiency *see* Antibodies
Congenital erythropoietic porphyria, *1074*, 1075
Congenital heart disease, **773–80**
aetiology, 773
atrial septal defect, 775–6, *776*
classification, *774*
coarctation of aorta, 777–8, *778*
cyanotic *see* Fallot's tetralogy
echocardiography, 704
genetic causes, 773
genetic counselling, 774
inheritance, *44*
magnetic resonance imaging, 706–7
patent ductus arteriosus, 777, *777*
presentation, 774, *774*
symptoms and signs, 773–4
transposition of great arteries, 779–80
valvular defects *see* Heart valves; Valvular heart disease
ventricular septal defect, 774–5, *775*
Congenital hyperbilirubinaemias (non-haemolytic), 329–30
conjugated
benign recurrent intrahepatic cholestasis, 330
Dubin-Johnson syndrome, 330
progressive familial intra-hepatic cholestasis syndromes, 330
Rotor's syndrome, 330
unconjugated, 329–30
Crigler-Najjar syndrome, 329–30
Gilbert's syndrome, 329
Congenital immunodeficiency *see* Immunodeficiency
Congenital intrahepatic biliary dilatation, 360
Congenital nephrotic syndrome, 586
Congenital syphilis, 178
clinical features, *178*
Congenital venous webs, 359, 360
Congestive cardiac failure *see* Heart failure
Congo-Crimean haemorrhagic fever, *82*, *110*, *114*, 116
Conium maculatum, 952
Conivaptan, *657*
Conjugate lateral gaze, 1104
abnormalities of, 1105, 1124
Conjunctiva, 1085
inflammation *see* Conjunctivitis
splinter haemorrhages, infective endocarditis, *770*
Conjunctivitis, *111*, 1087–9, *1092*
allergic, 1089
bacterial, 1088
chlamydial, 1088
differential diagnosis, *1088*
viral, 1088–9, *1088*
adenoviral, 1088
herpes simplex, 1088–9
molluscum contagiosum, 1089
Connective tissue
degradation, 500–1
diseases *see* Autoimmune rheumatic disease *see also* Collagen
Connective tissue-derived growth factor, 852
Connexins, 26
Conn's syndrome, 653, 1023–4
causes, 1024
clinical features, 1024
investigations, 1024
pathophysiology, 1023
treatment, 1024
Conocybe, 952
Consciousness, 1122
clouding, 1123
loss, *see* Coma; Unconsciousness
poisoning, 934
Consent, 3–5
battery, 4
competence, 5–6
disclosure, 4
express vs complicit, 4
express vs implied, 4
parental, 6
risk, 4
Constipation, 242, 295–7, *295*
assessment, 295
causes, *295*
defecatory disorders, 296
normal-transit, 295–6
slow-transit, 296
treatment, 296–7
Constrictive pericarditis *see* Pericarditis
Consumerism, modern healthcare, 9
Consumption coagulopathy *see* Disseminated intravascular coagulation (DIC)
Contact eczema, 1239–40, *1239*
allergic, 1239–40
irritant, 1240
topical preparations, 1270
treatment, 1240
Contact lens keratitis, 1090, *1090*
Continuous ambulatory peritoneal dialysis (CAPD), 637
gastrointestinal complications, 629
peritonitis, *637*
Continuous erythropoiesis receptor activator, 633
Continuous positive airway pressure (CPAP), 915
COPD, 841
Continuous renal replacement therapy, acute tubular necrosis, 623
Contraceptives
oral *see* Oral contraceptive(s)
Contrast agents
side-effects, *928*
Contrast nephropathy, 624–5
Controlled mechanical ventilation (CMV), 912, 914–15
see also Mechanical ventilation
Controlled observational studies, 931
Conversion disorders *see* Dissociative (conversion) disorders
Convulsions
overdose/poisoning, 936
see also Seizures
Cooley's anaemia, 408
Coombs' test *see* Direct antiglobulin (Coombs') test
Cooperativity, 391
Copper, 225
absorption, 357
deficiency, 225
excretion, 357
poisoning, antidotes, *937*
recommended IV intake, 233
RNI, 225
serum, Wilson's disease, 357, 1122
sources of, 225
toxicity, 225
see also Wilson's disease
urinary, Wilson's disease, 357
Coprinus atramentarius, 952
Copy Number Variation Project, 42
Cor pulmonale *see* Pulmonary heart disease (cor pulmonale)
Cord-blood-derived embryonic like stem cells, 47
Corkscrew oesophagus, 254
Cornea, 1085, 1089–90
abrasions, 1089, *1089*
calcification, 1020
foreign body, 1089, *1090*
inflammation *see* Keratitis
keratitis, 1090, *1090*
oedema, 1085
trauma, 1090, *1090*
ulcers/ulceration, dendritic ulcers, *1089*
Coronary angiography, 706, *706*, *708*
computed tomography, 706, *706*
angina pectoris, 749
dilated cardiomyopathy, 792
Coronary arteries, 681, 684–5, *685*
calcification, 631, 693, 705–6, *706*
disease *see* Coronary artery disease (CAD)
Coronary artery bypass grafting, 751–2, *751*
myocardial infarction, 758
Coronary artery disease (CAD), 744
left bundle branch block, *720*
magnetic resonance imaging, 707–8, *708*
National Service Framework, 747–8
prevention, 745
risk factors, 744–7, *745*
age, 745
alcohol, 747
C-reactive protein, 747
cholesterol, *1063*
coagulation factors, 747
diabetes mellitus, 746
diet and obesity, 746
family history, 745
gender, 745
genetic factors, 747
hyperlipidaemia, 746
hypertension, 746
lipoprotein (a), 747
NSAIDs, 747
psychosocial well-being, 747
sedentary lifestyle, 747
smoking, 745–6
see also Atheroma; Coronary atherosclerosis; Ischaemic heart disease
Coronary artery thrombosis, 442
Coronary atherosclerosis, 744
development, 744
pathogenesis, 744
see also Thrombosis
Coronary circulation, 684–5, *685*
arteries *see* Coronary arteries
Coronary heart disease (CHD) *see* Coronary artery disease (CAD)
Coronary revascularization, 755
Coronaviruses, *80*, *110*, 119
Corproporphyrins, 1073
Corpus striatum, 1113, *1113*
Correlation, 932–3

Correlation coefficient, 932
Corrigan's sign, 690
Cortical venous thrombosis, 1136
Corticospinal (pyramidal) system, 1111–12, *1111*
  lesions of, 1111–12, *1111*
Corticosteroids, 965
  inhaled, 855
  intra-articular
  side-effects, 855, *1015*
    hypertension, 799
    myopathy, 1179
    rash, *1265*
    uraemia, 619
  specific conditions
    ALI/ARDS, 919
    allergy, 72
    ANCA-positive vasculitis, 597
    asthma, *854*, 852–3, 855
    autoimmune haemolytic anaemia, 416
    Behçet's disease, 552
    Churg-Strauss syndrome, 871
    COPD, 839
    cryoglobulinaemic renal disease, 591
    giant cell arteritis, 550
    gout, 538
    IgA nephropathy, 594
    juvenile idiopathic arthritis, 554
    laryngitis, 833
    membranous glomerulonephropathy, 588
    mesangiocapillary glomerulonephritis, 590
    minimal-change nephropathy, 586
    myocarditis, 788
    osteoarthritis, 522
    pericarditis, 795
    polyarteritis nodosa, 551
    psoriatic arthritis, 535
    relapsing polychondritis, 557
    renal transplantation, 640
    retroperitoneal fibrosis, 617
    rheumatoid arthritis, 529
    rhinitis, 833
    sarcoidosis, 556
    Sjögren's syndrome, 548
    systemic lupus erythematosus, 544
    thrombocytic purpura, 436
  supervision of therapy, 1015–16
  surgical guidelines, *1016*
  see also Cortisol; Glucocorticoids; Hydrocortisone; individual types
Corticotrophin-releasing hormone (CRH), 970, *971*, 1008
  Cushing's syndrome, 1013
Cortinarius
  C. orellanus, 952
  C. speciossimus, 952
Cortisol, *971*, 1030
  potency, *1008*
  release
    circadian rhythm, 967, *968*
    in shock, 895
  urinary free, Cushing's syndrome, 1012
Corynebacterium
  C. diphtheriae, *123*, 126
    see also Diphtheria
  C. minutissimum, 1228
Coryza *see* Common cold (acute coryza)
Cosmids, 35
Costochondral joints, 499
Costochondritis, 533, 820
Cotton wool spots, 1051
Cough, 818, 819
  asthma, 852
  bovine, 819
  cancer-related, 496
  chronic, 845–6
  productive, 819
  specific conditions
Cough centre, 819
Cough headache, 1137
Cough reflex, 818, 845
  poisoning, 934
Coumarins, 445
Councilman bodies, 333
Counselling, genetic *see* Genetic counselling
Counter-regulatory anti-inflammatory response (CARS), 899
Countercurrent mechanism, 573
Couple therapy, 1204
Courvoisier's sign, 383
Cowden syndrome, *301*, 302
Cowpox, *104*, 109
Coxibs, 518
  gout, 538
  rheumatoid arthritis, 529
Coxiella burnetii, 139
  infective endocarditis, 770
  pneumonia, *858*
Coxsackievirus, *110*, 111–12, *1179*
  pericarditis, 794
Crabs *see Pediculosis*
Crackles, 821
Cradle cap, 1239
Cramps
  heat-induced, 954
  writer's, 1150
Cranberry juice, urinary tract infections, 603
Cranial arteritis *see* Giant-cell arteritis
Cranial nerves, *1100*, **1100–10**
  abducens (VI), 1106
  accessory (XI), 1110
  facial (VII), 1107–8
  glossopharyngeal (IX), 1110
  hypoglossal (XII), 1110
  oculomotor (III), 1105
  olfactory nerve (I), 1100
  optic nerve (II), 1100–4
  trigeminal (V), 1106–7
  trochlear (IV), 1105
  vagus (X), 1110
  vestibulo-cochlear (VIII), 1108–9
  see also individual nerves
Cranial polyneuropathy, 1173
Craniopharyngioma, *972*, 974
Craniotabes, 567
Creams, 1270
Creatine kinase (CK)
  hypothyroidism, 986
  polymyositis/dermatomyositis, 547
Creatine phosphokinase (CPK)
  muscular dystrophies, 1178
  serum levels, *1122*
Creatinine
  plasma/serum, *574*, *585*
    causes of alteration, *620*
    chronic renal failure, 627
    cirrhosis, 346
    systemic lupus erythematosus, 543
    systemic sclerosis, 546
Creatinine clearance, 574–5, *574*
  estimation of, *575*
  glomerular disease, *585*
Creatinine-kinase-MB, *752*
CREB, 24, 31, **33**
Crepitus, 503
CREST syndrome, 546
Cretinism, 225, 987
Creutzfeldt-Jakob disease (CJD), 121–2, 1158
  diagnosis, 48
  EEG, 1121
  iatrogenic, 1158
  incidence, 1158
  mortality, *122*
  new variant *see* Variant Creutzfeldt-Jakob disease (vCJD)
  sporadic, 1158
Cricoarytenoid joint, rheumatoid arthritis, 527, 871
Cricothyroidotomy, 914
Crigler-Najjar syndrome, 329–30
Crisantaspase, acute lymphoblastic leukaemia, 471
Critical care *see* Intensive care
Critical illness, 890
  myopathy, 916
  polyneuropathy, 916, 1176
Crocidolite (blue asbestos), 879
Crohn's disease, 245, 282, **288–91**
  anal/perianal complications, *288*
  clinical features, 288
  examination, 288
  extragastrointestinal manifestations, *288*
  histology, *287*
  investigations, 288–9
    blood tests, 288–9
    disease activity, 289
    radiology and imaging, 289
    stool cultures, 289
  medical management, 289, *289*
  microscopic changes, 287
  remission induction, 289, *289*
  remission maintenance, 290
  and small intestine tumours, 284
  surgical management, 290
Cross-over studies, 929
Crossmatching, 425
  electronic, 425
  immediate spin, 425
Croup, 833
Cruciate ligament tear, 514
Crusting, *1227*
Cryoglobulinaemia
  essential, *549*
  renal tubular acidosis, *676*
  rheumatoid factor, *503*
Cryoglobulinaemic renal disease, 591, *591*
Cryoglobulinaemic vasculitis
  clinical features, *550*
  diagnosis, *550*
Cryoglobulins, 591
  chronic renal failure, 627
  glomerular disease, *585*
Cryoprecipitate, 429
  thrombotic thrombocytopenic purpura, 437
  von Willebrand disease, 440
Crypt cells, *394*
Cryptococcosis, *150*, 151
Cryptococcus
  C. neoformans, *70*, *82*, *90*, 151, *191*
    HIV-associated, 200
  HIV-associated, 200
Cryptogenic organizing pneumonia, 873
Cryptorchidism, 1000
Cryptosporidiosis, 162, 282
  HIV-associated, 201
Cryptosporidium
  C. parvum, *90*, *161*, *190*, 201
  diarrhoea, 309
Crypts of Lieberkühn, 270
Crystal arthritis, 504, 522, **536–9**
  aetiology, 536
  gout and hyperuricaemia *see* Gout; Hyperuricaemia
  pseudogout, 501, 504, 539
Crystal cellulitis, 538
Crystal deposition diseases, *see* Gout
Crystal meth, 939
Crystal osteoarthritis, 521, *522*
Crystal synovitis, analgesia, 518
Crystalloids, 906
Culex
  C. fatigans, 164
  C. pipiens, 116
  C. tritaeniorhynchus, 115
Cullen's sign, 377
Cultural barriers to communication, 15
  beliefs, 15
  language, 15
  non-verbal communication, 15
Culture, 1185–6
Cumulative mutation theory of ageing, *227*
Cuneate nuclei, 1115
Curies, 958
Curling ulcers, 262
Cushing's syndrome, *228*, 972, *972*, 1012–14
  causes, *1012*
  clinical features, 1012, *1013*
    diabetes, 1031
  depression, 1200
  diagnosis, 1013–14
  drug-induced, *964*
  myopathy, 1179
  pathophysiology and causes, 1012
  rheumatological problems, 556
  skin manifestations, 1254, 1265
    striae, 1263
  tests for, 968, *969*
  treatment, 1013–14
Cutaneous larva currens, 167
Cutaneous larva migrans, *80*, 168
  rash, *88*
Cutaneous leishmaniasis *see* Leishmaniasis, cutaneous
Cutaneous leucocytoclastic vasculitis, *549*, 551–2
Cutaneous plague, 146
CXCL8, 55, 56
CXCL12, 55
CXCR4, 185
Cyanide poisoning, 943
  antidotes, *937*
  clinical features, 943
  metabolic acidosis, *939*
  treatment, 943
Cyanosis, 690, 820
  congenital heart disease, 773
Cyclic AMP, 214
Cyclic AMP phosphodiesterase inhibitors, autosomal-dominant polycystic kidney disease, 643
Cyclic citrullinated peptide antibodies, 503
Cyclic GMP, 686, 896
Cyclin, 27
Cyclin-dependent kinases, 27
Cyclizine
  nausea, 461, 496
  sickle cell anaemia, *411*
Cyclo-oxygenase, 262–3, 517–18, 577
Cyclopenthiazide, hypertension, 803
Cyclopentolate, uveitis, 1093
Cyclophosphamide, 76, 459
  acute lymphocytic leukaemia, 475
  ANCA-positive vasculitis, 597
  breast cancer, *461*, 488, 489
  Burkitt's lymphoma, 483
  cryoglobulinaemic renal disease, 591
  diffuse large B cell lymphoma, 482
  focal segmental glomerulosclerosis, 587
  Hodgkin's lymphoma, *461*
  idiopathic pulmonary fibrosis, 873
  IgA nephropathy, 594
  lupus nephritis, 591
  membranous glomerulonephropathy, 588
  minimal-change nephropathy, 586
  multiple sclerosis, 1153
  non-Hodgkin's lymphoma, *461*
  pemphigus vulgaris, 1256
  polymyositis/dermatomyositis, 548
  rheumatoid arthritis, 532

side-effects, hepatotoxicity, *366*
systemic lupus erythematosus, 544
systemic sclerosis, 547
Wegener's granulomatosis, 870
Cyclosarin, 946
Cycloserine, tuberculosis, 866
Cyclospora cayetanensis, *161*, 162
Cyclosporin *see* Ciclosporin (cyclosporin)
Cylindroma, 884
CYP7A1, 321
CYP8B1, 321
Cyproterone acetate
acne vulgaris, 1246
hirsutism, 1005
Cyst(s)
choledochal, 373
dermoid, *888*
epidermoid, 1257
hydatid, *331*, *362*
mucinous, 384
ovarian, 243, 316
pilar, 1257
popliteal (Baker's), 514
serous, 384
Cystathionine synthase deficiency, *1070*
Cystatin C, 575
Cysteinyl leukotriene receptor antagonists, 72
Cystic duct, 320
Cystic fibrosis, 42, 380, 381–2, *844*, **844–5**
ciliary dyskinesia, 818
clinical features, 844–5
gastrointestinal, 365, 845
nutritional, 844
respiratory, 845
and diabetes, 1031
diagnosis, 845
gene therapy, 46, *46*, 845
prognosis, 845
screening, 845
treatment, 382, 845
Cystic fibrosis chloride ion channel, 22
Cystic fibrosis transmembrane conductance regulator, 322, 844
Cystic fibrosis transmembrane regulator gene, 46, *46*, 381
Cysticercosis, 170–1, *1179*
treatment, 171
Cystine stones, *609*, 613
Cystinosis, 1071
Cystinuria, 575, 1070–1
renal calculi, 611
screening test, 613
Cystitis, 601
interstitial, 602
Cystoscopy, urinary tract obstruction, 615
Cysts
choledochal, 373
dermoid, *888*
epidermoid, 1257
hepatic, 642
hydatid, *331*, *362*
liver, 360
lung, 880
ovary, 243, 316
pilar, 1257
pleuropericardial, 888
popliteal (Baker's), 514
Cytarabine *see* Cytosine arabinoside
Cytochrome C, 28, *28*
Cytochrome P450 enzymes
inducers, *925*
inhibitors, *925*
see also Genetic polymorphisms
Cytoid bodies, 543, 1051
Cytokeratin 7, *493*
Cytokeratin 20, *493*
Cytokines, 54, 65
acute inflammation, 68
endocrine effects, 54
in fever, 87
immunosuppression, 76
osteoarthritis, 520
paracrine function, 54
pleiotropy, 54
regulation of metabolism, 214
response to shock, 895
synergy, 54
see also individual cytokines
Cytology, 827
Cytolysins, 66
Cytolytic/cytocidal infection, 85
Cytomegalovirus (CMV), *70*, *90*, *104*, *105*, 107, *191*
clinical features, 107, *107*
diagnosis, 107
gastrointestinal conditions, 202
hepatitis, 342
HIV-associated, 202, *202*
in organ rejection, 75
owl-eye inclusions, *107*, 202
pneumonia, *858*, 859
pyrexia, *91*
rash, *88*
retinitis, 188, 202, *202*
transfusion-transmitted, 428
treatment, 107
Cytopenia, idiosyncratic, 431
Cytoplasm, *19*, 22–3
Cytoplasmic linker protein, 24
Cytosine, 29, *30*
Cytosine arabinoside, 460
acute lymphoblastic leukaemia, 471
acute myeloid leukaemia, 470
Burkitt's lymphoma, 483
Cytotoxic antibiotics, 460
Cytotoxic drugs
high-dose therapy, 462
resistance, 462
side-effects, 460–2, *460*, *928*
bone marrow suppression and immunosuppression, 461
cardiotoxicity, 461
extravasation, 460–1
hair loss, 461
mucositis, 461
nausea and vomiting, 461
nephrotoxicity, 462
neurotoxicity, 462
sterility and premature menopause, 462
teratogenicity, *924*
see also Cancer treatment
Cytotoxic hypersensitivity *see* Autoimmunity (type II hypersensitivity)
Cytotoxic T cells *see* T cell(s), cytotoxic
Cytotoxic-associated protein, 258–9
Cytotoxin, *128*, 129

D-dimer, pulmonary embolism, 785
Da Costa's syndrome, 688
Dabigatran, 446
arterial thrombosis, *443*
atrial fibrillation, 725
Dacarbazine, 460
Hodgkin's lymphoma, *461*
Daclizumab, 76, *76*
Dacryocystitis, 1087, *1087*
Dactylitis, 409, 509, 534, 1253
Dagon disease, *1069*
Dairy products, allergy to, 73
Dalfopristin, 99, 124
Danaparoid, 445
Danazol
side-effects, hepatotoxicity, *366*
thrombocytic purpura, 436
Dandy-Walker syndrome, 1162
Danon's disease, 788
Dantrolene, amfetamine overdose, 940
Dapsone
aphthous ulcers, 247
bullous pemphigoid, 1256
dermatitis herpetiformis, 1256
erythema nodosum, 1250
leprosy, *141*
linear IgA disease, 1257
malaria, *156*
pyoderma gangrenosum, 1251
side-effects
haemolysis in G6PD deficiency, *414*
haemolytic anaemia, 415
neuropathy, *1174*
subacute lupus erythematosus, 1252
thrombocytic purpura, 436
Daptomycin, 99–100, 124
Darbepoietin-alfa, 633
Darier's sign, 1265
Darunavir, *193*, 196
Dasatinib, *464*
Datura stramonium, 952
Daunorubicin
acute lymphoblastic leukaemia, 471
acute myeloid leukaemia, 470
DC-cardioversion, 712–13, *712*
atrial fibrillation, 724
atrial flutter, 725
indications, 713
De Musset's sign, 767
De Quervain's tenosynovitis, 508
de Quervain's thyroiditis, 988
De Toni-Fanconi-Debré syndrome, 1069
see also Fanconi syndrome
Deafness, 1077, 1078–9, *1078*, 1108
acoustic neuroma, 1079
causes, *1108*
communication problems, 16
conductive, *1078*
noise trauma, 1079
noise-induced, 959
otitis media, 1078–9
otosclerosis, 1079
perforated tympanic membrane, 1078
presbycusis, 1079, *1079*
Rinne test, 1078
sensorineural, *1078*
Weber test, 1078
Death
sudden *see* Sudden death
see also Mortality
Death domains, *28*
Death effector domains, *28*
Debrisoquin, metabolic ratios, *926*
Decay accelerating factor, 418
rheumatoid arthritis, 524
Decibels, 959
Decompression sickness, 957
bends, 957
chokes, 957
Decongestants, 833
Decorin, 500
Decubitus angina, 748
Decubitus ulcers *see* Pressure sores (decubitus ulcers)
Deep brain stimulation, 1203
Deep venous (vein) thrombosis (DVT), 809–10
air travel, 955
clinical features, 809–10
hip replacement surgery, 446
investigations, 810
prevention, 810
prognosis, 810
risk, *445*
treatment, 810
DEET, 157
Defecation
disorders of, 296
obstructive disorders *see* Constipation
physiology of, 295
see also Diarrhoea
Defence mechanisms (immunological) *see* Host defence
Defence mechanisms (mental), 1189
Defensins, 56, *57*, 274, 1226
Defibrillation, 711
Defibrillators, 711
automated external, 711
biphasic, 711
implantable *see* Implantable cardioverter-defibrillator (ICD)
Defibrotide, veno-occlusive disease, 360
Dehydration
diabetic ketoacidosis, *1046*, 1047
older people, 1048
7-Dehydrocholesterol, 559
Dehydroepiandrosterone sulphate, PCOS, 1004
Déjérine-Sottas disease, 1175
Delayed after depolarizations, 716
Delayed hypersensitivity (type IV), 68
see also Hypersensitivity diseases
Deletions, 39, 41
Deliberate self-harm (DSH), 1206, *1206*
see also Suicide
Delirium, 1123, 1218–19, *1218*
causes, *1219*
prognosis, 1219
Delirium tremens (DTs), 359, 1213
treatment, *1214*
Delusions, 1187–8, *1187*
of passivity, 1188
secondary, 1188
Demeclocycline, 1019
Dementia, *1167*, **1167–9**
Alzheimer's disease *see* Alzheimer's disease
causes, *1167*
dialysis-related, 630
differential diagnosis, 1167
frontotemporal, 1168
incidence, *1096*
investigations, 1168, *1169*
with Lewy bodies, 1168
management, 1168
memory loss, 1100
multi-infarct, 1130
pellagra, 222
social consequences, 1168
subcortical encephalopathy, 1168
vascular, 1168
Demodex, 1247
Demyelination, 1171
see also Neuropathies
Dendritic cells, 55, *56*, 58–9, *58*, *59*
activation, 58, *60*, 819
asthma, 850
maturation, *59*
migration through lymph nodes, *67*
molecules on, *59*
myeloid, *56*, 58, *59*
plasmacytoid, *56*, 58, *59*
T cell activation, 59, 65
Dengue fever, *82*, *110*, 114–15
classic, 115
clinical features, 115
diagnosis, *115*
haemorrhagic, *114*, 115
prevention, 115
rash, *88*
returning travellers, *90*
Denial, 1189
Dental caries, 248
prevention by fluoridation, 226
Dent's disease, 671
Deoxyadenosulcobalamin, 399
Deoxycholic acid, 308, 322, 367
Deoxyhaemoglobin, 391
Deoxyuridine suppression test, 398
Depersonalization, 1188
Depolarization

Depression, 1186, 1198–204, *1199*
  aetiology, 1200–1
    biochemical, 1200
    childhood traumas, 1201
    genetic, 1200
    hormonal, 1200
    neural changes and neuronal growth, 1200
    personality, 1201
    sleep, 1200–1
    social, 1201
  bipolar disorder *see* Bipolar affective disorder
  clinical features, 1198–9
  differential diagnosis, 1200, *1200*
  epidemiology, 1199–200
  investigations, 1200
  life-limiting illness, 496
  physical causes, *1190*
  prognosis, 1204
  treatment, 1201–4
    antidepressants, 1201–2
    electroconvulsive therapy (ECT), 1203
    exercise, 1201
    psychological, 1204
    see also individual conditions
Derealization, 1188
Dermatitis
  chronic actinic, 1249
  pellagra, 222, *222*
  perioral, 1247
  radiation-associated, 958
  see also Eczema
Dermatitis herpetiformis, 279–80, 1256
  splenic atrophy, 423
Dermatobia hominis, 172
Dermatofibroma, 1257
Dermatological disorders *see* Skin diseases
Dermatomes, spinal root(s), *1116*
Dermatomyositis *see* Polymyositis/dermatomyositis
Dermatophagoides pteronyssinus, 831
Dermatophyte infections, 1232–3
Dermatophytes, *191*
Dermatophytosis, *150*, 153
Dermatoscopy, *1228*
Dermis, 1227
Dermographism, 1243
Dermoid cyst, *888*
Desensitization, 72–3, *73*
Desferrioxamine, *937*
  haemochromatosis, 357
  iron overload, 408
Designer drugs, 76
Desloratadine, rhinitis, 833
Desmin, 320, *493*
Desmogleins, 26, 1225, 1228, 1255
Desmoplakin, 791, 1225
Desmopressin, 664
  diabetes insipidus, 1018
  haemophilia A, 439
  replacement therapy, *976*
  von Willebrand disease, 440
Desmosine, 500
Desmosomes, *25*, 26
Desogestrel, 1006
Desquamative gingivitis, 247
Desquamative interstitial pneumonia, 873
Developed countries, protein-energy malnutrition (PEM), 212–15
Developing/tropical countries
  bacterial infections, **140–9**
    gastrointestinal infections, 143–5
      cholera *see* Cholera
      enteric fever, 144–5
    skin
      soft tissue and eye disease, 140–3
      anthrax, *80*, *83*, 142
      bejel, 142
      endemic treponematoses, 142–3
      leprosy, 140–1, *140*, *141*
      mycobacterial (Buruli) ulcer, 142
      pinta, 142–3
      trachoma, 143
      yaws, 142
    systemic infections, 145–9
      actinomycosis, 149
      bartonellosis and ehrlichiosis, 148
      melioidosis, 148–9
      Nocardia infections, 149
      non-tuberculous mycobacterial infections, 146, *146*
      plague, 146
      relapsing fevers, 146–7
      tuberculosis, 145, *145*
      typhus, 147–8, *147*
Dexamethasone, 461
  acute lymphoblastic leukaemia, 471
  Addison's disease, *1012*
  laryngitis, 833
  mountain sickness, 956
  pain relief, *495*
  potency, *1008*
  uveitis, 1093
Dexamethasone suppression test, 1009–10, *1010*
  Cushing's syndrome, 1012
Dexamethasone test, *969*
Dexamfetamine, 939
  narcolepsy/cataplexy, 145
Dexfenfluramine, side-effects, respiratory, *877*
Dextrans, 906
Dextropropoxyphene, musculoskeletal problems, 517
Diabetes insipidus, *664*, 1017–18
  biochemistry, 1017
  clinical features, 1017, *1017*
  cranial, *1017*
  nephrogenic *see* Nephrogenic diabetes insipidus
  treatment, 1018
Diabetes mellitus, **1029–75**
  adolescence, 1046
  antinuclear antibodies, *504*
  asymptomatic, 1036
  classification, 1031–2, *1032*
  clinical features
    acute abdomen, *315*
    Charcot's joints, 555
    cutaneous manifestations, 1254
    haemochromatosis, 356
    malabsorption, 282
    postural hypotension, *659*
  complications, 1036, 1049–58
    eye disease *see* Diabetic eye disease
    infections, 1057
    macrovascular, 1050, *1050*
    microvascular, 1050
    neuropathy *see* Diabetic neuropathy
    pathophysiology, 1049
    renal *see* Diabetic kidney
    skin and joints, 1057
  and coronary artery disease, 746
  diagnosis, 1036–7, *1036*
  see also Glucose tolerance test (GTT)
  drug-induced, *964*
  emergencies, ketoacidosis *see* Diabetic ketoacidosis
  gestational, 1058–9
  latent autoimmune, in adults, 1032
  metabolic control, 1044–5
    glycosylated haemoglobin and fructosamine, 1044
    home blood glucose testing, 1044
    importance of, 1044–5
    reversal of complications, 1045
    targets, 1044, *1044*
    urine tests, 1044
  metabolic emergencies, 1046–9, *1046*
  monogenic, 1035–6, *1035*
  pancreatic endocrine failure, 381
  permanent neonatal, 1035
  physical examination, 1036
  practical aspects, 1046
  pregnancy, 1058–9
  presentation
    acute, 1036
    subacute, 1036
  psychosocial implications, 1045–6
  regular checks, 1045, *1045*
  secondary, causes, *1032*
  surgery, 1058
  treatment/management
    diet, 1037–8, *1037*
    injection therapies, 1039–40
    insulin *see* Insulin
    oral therapy, 1037–45, *1038*
    see also individual therapeutic drugs
    patient education and community care, 1037
  type 1, *965*
    autoantigens, *74*
    autoimmunity, 1033
    biochemical, *1032*
    causes, 1032
    clinical, *1032*
    environmental factors, 1033
    epidemiology, *1031*, 1032
    genetic susceptibility, 1032–3
    heredity, *1031*, 1032–3
    HLA system in, *60*
    incidence, *1032*
    pathogenesis, *1032*
    pre-condition, 1033
    prevention, 1033
  type 2, 1033–5
    biochemical, *1032*
    causes, 1034–5
    clinical, *1032*
    environmental factors, 1034
    epidemiology, *1032*, 1033–4
    genetic causes, *1035*
    heredity, *1032*, 1034
    immune involvement, 1034
    inflammatory changes, 1034
    insulin secretion abnormalities, 1034–5
    pathogenesis, *1032*
    and prevention, 1035
    treatment, 1038–9, *1039*, 1041, *1041*
  unstable, 1059
  see also entries beginning diabetic
Diabetic amyotrophy, 1056
Diabetic cheiroarthropathy, 1057
Diabetic coma, 1043
  see also Diabetic ketoacidosis
Diabetic dermopathy, 1254
Diabetic eye disease, 1050–1, *1051*
  cataracts, 1050
  examination, 1052
  management, 1052–3, *1052*
  see also Diabetic retinopathy
Diabetic foot, 1056–7, *1056*, *1057*
  management, 1057, *1057*
Diabetic ketoacidosis, 677, *677*, 1036, 1046–7, *1046*
  clinical features, 1047
  diagnosis, 1047
  management, 1047–8, *1048*
  pathogenesis, 1046–7, *1046*
  in pregnancy, 1058
  recurrent, 1059
Diabetic kidney, 1053
Diabetic nephropathy, 589, 1050, 1053–5
  diagnosis, 1054
  epidemiology, 1053
  infective lesions, 1054
  ischaemic lesions, 1054
  management, 1054–5
  natural history, *1054*
  pathophysiology, 1053–4
  see also Diabetic kidney
Diabetic neuropathy, 1050, 1055–6, *1055*, 1174
  acute painful, 1055
  autonomic, 1056
  bladder involvement, 1056
  cardiovascular system, 1056
  gastrointestinal tract, 1056
  incidence, *1096*
  mononeuritis, 1056
  sequelae, 1055
  symmetrical sensory polyneuropathy, 1055
Diabetic retinopathy, 1036, 1050, 1051–2, *1052*
  background, 1051
  circinate, 1052
  maculopathy, 1051–2
  mixed, 1052
  pathological changes, *1053*
  pre-proliferative, 1051
  prevalence, *1052*
  proliferative, 1051, 1052–3
  treatment, 1052
Diabetic stiff skin, 1254
DIABLO protein, 28
Dialysate, 634, *634*
Dialysis
  acute tubular necrosis, 623
  complications, 638–9
  side-effects, pruritus, 629
  see also Haemodialysis; Peritoneal dialysis; Renal replacement therapy
Dialysis amyloidosis, 638–9
Dialysis dementia, 630
Dialysis disequilibrium, 630, 636
Dialysis machine, 634
Diamond-Blackfan syndrome, 403
Diamorphine
  acute coronary syndromes, *754*
  myocardial infarction, 757
  sickle cell anaemia, *411*
Diapedesis, 67
Diaphragm, 812–13
  bilateral weakness/paralysis, 886–7
  eventration, 887
  fatigue, 886
  hernia, 887
  unilateral paralysis, 886
Diarrhoea, 242, **305–9**
  acute, 307
  antibiotic-associated, *132*, 307
  bacterial causes, *129*
  bile acid malabsorption, 307–8
  bile-salt induced, 281
  causes, *307*
  chronic, 307, *308*
  factitious, 308–9
    dilutional diarrhoea, 309
    purgative abuse, 308–9
  functional, 313
  HIV infection, 309, *309*
  mechanisms of, 305–7, *306*
    abnormal motility, 307
    inflammatory, 307
    osmotic, 305–6
    secretory, 306
  small intestinal disease, 275
  specific conditions
    pellagra, 222
    travellers' *see* Travellers' diarrhoea
    see also Gastroenteritis
Diarrhoea-associated haemolytic uraemic syndrome, 598
Diarrhoeal disease, mortality, *80*
Diastolic blood pressure, 736

Diazepam
  amfetamine overdose, 940
  convulsions, 936
  genetic polymorphism in metabolism, *926*
  osteoporosis, 563
Diazoxide
  insulinoma, 1060
  side-effects, hypertrichosis, 1003
Dickkopf, 558
Dickkopf-1, 484
Diclofenac, *495*
  gout, 538
  musculoskeletal problems, 517, *518*
  renal colic, 612
  sickle cell anaemia, *411*
  side-effects, *928*
    hepatotoxicity, *366*
Dicobalt edetate, *937*, 943
Didanosine, *192*, *195*
DIDMOAD (Wolfram) syndrome, 1017
Diet
  diseases/disorders, 207
    acute tubular necrosis, 623
    asthma, 849
    cancer association, 207, 449
    coronary artery disease, 746
    diabetes mellitus, 1037–8, *1037*
    heart failure, 737
  estimated average requirement (EAR), 207
  gluten-free, 279
  health promotion, 212
  lipid-lowering, 1065
  low-calorie, 232
  low-carbohydrate, 232
  low-energy-density, 232
  low-fat, 232
  lower reference nutrient intake (LRNI), 207
  prescribing, 1038
  processed foods, 207
  protein-energy malnutrition *see* Protein-energy malnutrition (PEM)
  recommended daily amounts (RDAs), 207
  reference nutrient intake (RNI), 207
  reference values, 207
  requirements, **208–12**
    carbohydrate, 211–12
    energy, 208–9
    fat, 210–11
    protein, 209–10
    vitamins *see* Vitamin(s)
  salt-restriction, 656
  specific conditions
    gout, 538
    obesity, 231–2
  very low-fat, 232
  see also Food; Nutrition
Dietary fibre, 211, 212
  increasing intake, 1065
Dietary guidelines, 212
Dietary induced thermogenesis, *208*
Dietary requirements *see* Diet, requirements
Diethylcarbamazine (DEC)
  filariasis, *163*, 165
  loiaisis, *163*
Diethyleneglycol poisoning
  antidotes, *937*
  metabolic acidosis, *939*
Diffuse idiopathic skeletal hyperostosis (DISH), 512
Diffuse large B cell lymphoma (DLBCL), 482
Diffuse parenchymal lung disease, *868*, **868–78**
Diffusion defects, 893
Diflucortolone valerate, *1237*
DiGeorge syndrome, 39, 70, 1022
Digital cerebral angiography, 1120
Digital subtraction angiography (DSA), 708
Digital vascular imaging, 327
Digitalis purpurea, 952
Digitoxin, poisoning, antidotes, *937*
Digoxin, *731*
  atrial fibrillation, 725
  heart failure, *739*
  myocarditis, 788
  overdose, 943
    antidotes, *937*
    cardioversion, 713
    clinical features, 943
    treatment, 943
  side-effects, elderly patients, *924*
  therapeutic drug monitoring, *927*
Dihydroartemisinin, *156*
Dihydrocodeine
  breathlessness, 840
  chronic cough, 846
  musculoskeletal problems, 517, *518*
Dihydrofolate, 400
Dihydrotestosterone, 1008
1,25-Dihydroxycholecalciferol (1 25(OH)2D3), 577
  see also Vitamin D
24,25-Dihydroxycholecalciferol *see also* Vitamin D
Dilated cardiomyopathy *see* Cardiomyopathy
Diloxanide furoate, 161
Diltiazem, 731, *731*
  angina pectoris, 750
  atrial fibrillation, 725
  side-effects, hepatotoxicity, *366*
  supraventricular tachycardia, 723
Dimercaprol, *937*
  side-effects, haemolysis in G6PD deficiency, *414*
Dimercaptopropanesulphonate (DMPS), *937*
Dimeric transmembrane receptors, 966
Dinoflagellate plankton toxin, *133*
Dinorphin, 1118
Dipalmitoyl lecithin, 811, 817
Dipeptidyl peptidase 4, 1039
2,3-Diphosphoglycerate (2,3-DPG) *see* 2,3-Bisphosphoglycerate
Diphtheria, *83*, 126–7
  antitoxin, *127*
  clinical features, 126
  cutaneous, 126
  diagnosis, 126
  immunization, *102*, *103*
  laryngeal, 126
  nasal, 126
  pharyngeal, 126
  prevention, 127
  treatment, 127
Diphtheritic neuropathy, 1173
Diphylloobothrium latum, *163*, 171
Diplopia, 1105, *1128*
Diprobase cream, *1270*
Diprobase ointment, *1270*
Dipyridamole
  arterial thrombosis, 443, *443*
  IgA nephropathy, 594
  mesangiocapillary glomerulonephritis, 590
  stroke, 1132
Direct allorecognition, 75
Direct antiglobulin (Coombs') test, 405, *415*
  autoimmune haemolytic anaemia, warm, 416
  cold haemagglutinin disease, 417
Disability, 502
Disc disease
  cervical *see* Cervical discs
  prolapse/protrusion
    cervical, 506
    lumbar, 510–11, *511*
Disclosure
  concerns, 11
  legal consent, 4
Discoid eczema, 1238
Discrimination in clinical practice, 7
Disease
  early origins in older adults, *227*
  and environment, 953
Disease-modifying anti-rheumatic drugs (DMARDs)
  biological, 531
  rheumatoid arthritis, 529–31, *530*
  see also individual drugs
Disopyramide, 730, *731*
  hypertrophic cardiomyopathy, 790
  side-effects, long QT syndrome, *727*
Displacement, 1189
Disposable soma theory of ageing, 227
Dissecting cellulitis, 1267
Disseminated intravascular coagulation (DIC), 427, 440–1, *441*, *455*, 898
  causes, 440–1
  clinical features, 441
  investigations, 441
  treatment, 441
Disseminated sclerosis *see* Multiple sclerosis
Dissociation curves
  carbon dioxide, 893, *893*
  oxyhaemoglobin, 892, *892*
Dissociative (conversion) disorders, 1195–7, *1196*
  aetiology, 1196–7, *1196*
  clinical features, 1196
  differential diagnosis, 1196
  management, 1197
  prognosis, 1197
Distal interphalangeal joints (DIP), 509
  nodal osteoarthritis, 509, 520
Distancing strategies (clinical communication), 10
Disulfiram
  alcohol dependence, 1214
  side-effects
    hepatotoxicity, *366*
    neuropathy, *1174*
Dithranol, psoriasis, 1242
Diuretics, **656–9**, *657*
  carbonic anhydrase inhibitors, *657*
  clinical uses, 657–8
  COPD, 839
  effects on renal function, 659
  heart failure, 738, *739*
  hypertension, 803–4
  loop *see* Loop diuretics
  myocarditis, 788
  potassium sparing, *657*
    clinical uses, 657–8, *659*
  resistance to, 658
  side-effects
    acute tubulointerstitial nephritis, *604*
    elderly patients, *924*
    endocrine disease, *964*
    renal calculi, 611
  specific conditions
  thiazide *see* Thiazide diuretics
  tricuspid stenosis, 768
  vasopressin/ADH receptor blockers (aquaretics), *657*
Divalent metal transporter 1, 394
Diverticular disease, 298–9, *299*
  acute abdomen, *314*
  clinical features and management, 298
  complications, 299
Diverticulitis, 298
  acute, 298–9
  treatment, 299
Diverticulosis, 298
Diverticulum
  Meckel's *see* Meckel's diverticulum
  oesophageal, 254–5
  Zenker's, 254
Diving, 956–7
  ascent, 957
  decompression sickness, 957
    bends, 957
    chokes, 957
  depth and pressure, *957*
  descent, 957
  lung rupture, 957
  nitrogen narcosis, 957
  oxygen narcosis, 957
  pneumothorax, 957
  surgical emphysema, 957
Diving reflex, 958
Dizziness, 1097
DMSA, *937*
DNA amplification, *453*
DNA damaging agents, 459–60
DNA (deoxyribonucleic acid)
  antibodies *see* Anti-DNA antibodies
  cloning, 35
  complementary (cDNA)
  double-stranded, *30*
    antibodies, *69*
  function, 29–33
    see also Gene(s)
  genomic, preparation, 33
  repair
  satellite, 33
  sequencing, 35–6
  structure, 29–33, *30*
    see also Chromosomes
  supercoiling, *30*, 33
  transcription, 29–30, *31*
  translation, 29–30, *31*
DNA methylation, 51
DNA polymerase, 35
DNA repair, 453
DNA repair inhibitors, 460
DNA viruses, *104*, **104–9**
  adenoviruses, 104
  erythroviruses, 109
  hepadna viruses, 109
  herpesviruses, 104–8, *105*
  papovaviruses, 108–9
  poxviruses, 109
  see also individual viruses
DNA-binding proteins, 31, *32*, **33**
  helix-loop-helix motifs, *32*, **33**
  helix-turn-helix motifs, *32*, **33**
  leucine zipper, *32*, **33**
  zinc finger, *32*, **33**
DNAase, 845
Dobutamine, 907, *908*
Docetaxel, 460
  breast cancer, *461*, 488, 489
  gastric cancer, 490
  prostate cancer, 492
Docosahexaenoic acid, 210, *210*, 211
Doctor-patient relationship, 2
  communication, 9–10
    empathy, 10, 12, 14
    failure, 10–11, 14–15
    good, 9–10
  discord, 10
  see also Communication
Docusate sodium
  enema, *297*
  sickle cell anaemia, *411*
Dofetilide, 730
  supraventricular tachycardia, 723
Doll's head reflex, 1105
Domperidone, 252, 461
Donath-Landsteiner test, 417
Donovan bodies, 180
Donovanosis, 180
  genital ulceration, *172*
Dopamine, 907, *908*, *971*, 1098, 1113
  acute tubular necrosis, 622
  calcium channel blocker overdose, 941

Dopamine agonists, *976*
acromegaly, 980
hyperprolactinaemia, 982
Parkinson's disease, 1147–8
pituitary tumours, *973*
Dopamine antagonists, side-effects, endocrine disease, *964*
Dopamine transporter imaging, brain, 1120
Dopamine-responsive dystonia, 1150
Dopexamine, 907, *908*
Doppler echocardiography, 699, 701
colour flow Doppler, 699, *701*
continuous-wave Doppler, 699, 701, *702*
pulsed-wave Doppler, 699
tissue Doppler, 701
Doppler ultrasound
brain, 1120
kidney, 580
oesophageal, cardiac output measurement, 904, *904*
renovascular disease, 608
Dorsal tenosynovitis, 508
Dosage regimens
fixed, 925, *925*
titrated, 925
Dose selection, 927
Dosulepin
chronic pain syndromes, 1194
chronic widespread pain, 516
Double Y syndrome, *40*
Double-contrast barium meal, 245
Down-beat nystagmus, 1109
Down's syndrome *see* Trisomy 21
Doxapram, 841
Doxazosin, hypertension, 805
Doxorubicin, 460
bladder tumours, 645
breast cancer, 489
Burkitt's lymphoma, 483
Hodgkin's lymphoma, *461*
ovarian cancer, 491
pancreatic neuroendocrine tumours, 384
side-effects
bone marrow aplasia, 402
nausea and vomiting, 461
Doxycycline, *156*
Chlamydia, 176
donovanosis, 180
gonococcal arthritis, 540
Lyme arthritis, 540
lymphogranuloma venereum, 177
malaria, *157*
spotted fevers, 148
syphilis, 179
urethritis, 177
DPC4 gene, 382
Dracunculiasis, 166
Dracunculus medinensis, *163*, 166
Dressing apraxia, *1099*, 1100
Dressler's syndrome, 794
Driving
diabetes mellitus, 1046
epilepsy, 1143, 1144
heart failure, 738
Drop attacks, 1144
Drowning/near-drowning, 957–8
dry drowning, 957
emergency treatment, 958
prognosis, 958
wet drowning, 957–8
Drug(s), **923–9**
absorption, in renal disease, 618
abuse *see* Drug misuse and dependence
affordability, 927, *928*
antiretroviral *see* Antiretroviral drugs
clinical trials *see* Clinical trials
co-morbidity, *923*
conditions/disorders due to *see also* Adverse drug reactions
dose, 925–7
dose selection, 927
elimination, 576
renal disease, 619
end-organ sensitivity, 618
endocrine disease, *964*
enzyme induction, 925–6, *925*
see also Cytochrome P450
hepatotoxicity, 365–7, *366*
information sources, 933
interactions, 197
see also individual drugs
measurement of effects, 927
metabolism *see* Drug metabolism
nephrotoxicity, 617–19
postrenal, 618
prerenal, 617
renal, 617–18
overdose, *see* Poisoning
pharmacodynamics *see* Pharmacodynamics
pharmacokinetics *see* Pharmacokinetics
plasma concentrations, 927
prescribing, 923–5, *923*
breast-feeding mothers, 924
elderly, 924
errors *see* Adverse drug reactions
neonates, infants and children, 924
pregnancy, 924
selection, 924–5
side-effects *see* Adverse drug reactions
therapeutic monitoring, 927, *927*
and uraemia, 619
volume of distribution, renal disease, 618
see also specific drugs/types
Drug fevers, *91*
Drug metabolism, 323, 365, 925, *925*
first pass, 925
in renal disease, 618
Drug misuse and dependence, 1214–16, *1214*
alcohol *see* Alcohol misuse/dependence
amfetamines, 1179, 1214–15
cannabis, 1215
causes, 1214
cocaine, 1215
ecstasy, 1215
hallucinogenic drugs, 1215
myopathy, 1179
opiates, 1179, 1215
psychoses, 1216
purgatives, 308–9
solvents, *1114*, 1214
tranquillizers, 1215
treatment, 1215–16
Drug reactions, HIV, 189
Drug resistance
antimicrobial chemotherapy, 92–4, *93*, *94*
antiretroviral drugs (ARVs), 196–7, *196*
aspirin, 443
cancer treatment, 462
cytotoxic drugs, 462
diuretics, 658
tuberculosis, 866
Drug-eluting stents, 750–1, *751*
Drug-induced antibodies, 414
Drug-induced cancer, 450
Drug-induced haemolytic anaemia, 417
Drug-induced hepatitis, 344
Drug-induced hypersensitivity syndrome, 1266
Drug-induced rashes, 1265–6, *1265*, *1266*
in HIV patients, 1269
Drug-induced respiratory reactions, 877–8, *877*
Drug-induced thrombocytopenia, 436–7
Drusen, 1094, *1094*, 1102
Dual energy X-ray absorptiometry (DXA)
indications for, *563*
musculoskeletal problems, 505
osteoporosis, 562, *562*
Dubin-Johnson syndrome, 330
Duchenne muscular dystrophy (DMD), 41, 1181
clinical features, 1181
gene therapy, 46–7
investigations, 1181
management, 1181
Duchenne-Becker muscular dystrophy, *789*
Ducts of Bellini, 571
Ductular cholestasis *see* Cholestasis
Ductus arteriosus, 687
Duffy antigen, 155
Duffy blood group, 424
Duloxetine, diabetic neuropathy, 1055
Dumping syndrome, 262
Duncan's syndrome, 70
Duodenal ulcers, 259, *261*
aetiology and epidemiology *see Helicobacter pylori*
epidemiology, 260
perforation, 261–2
see also Peptic ulcer disease
Duodenitis, 260
Duodenum, 258
crypt cells, *394*
mucosa, 258
Duplications, 39
Dupuytren's contracture, 329, 509
Dural venous sinus thrombosis, 1136
Duroziez's sign, 767
Dusts, 831–2, 848
irritant, 849
silica, 545
wood, *848*
Dwarfism
Laron type, 979
see also Achondroplasia (dwarfism)
DXA scanning *see* Dual energy X-ray absorptiometry (DXA)
Dyes, allergy, *848*
Dynein, 24, 812, *812*
Dysarthria, 1100, 1114, *1128*, 1146
Dysdiadochokinesis, 1114
Dysdifferentiation theory of ageing, 226
Dysentery, *83*
antibiotics, *132*
bacterial causes, *129*
returning travellers, *90*
Dyserythropoiesis, 397
Dysfibrinogenaemia, 435
Dyshormonogenesis, 985
Dyskeratosis congenita, *430*
Dyskinesias, 1113, 1149–51
ciliary, 818
paroxysmal, 1145
tardive, 1148
see also individual conditions
Dyslexia, 1099
Dysmetria, 1114
Dysmnesia, 1188
Dyspareunia, *995*
Dyspepsia, 241, *314*
functional, 311
management, 263
Dysphagia, 241, 249, 1084–5, *1085*
carcinoma of oesophagus, 256
causes of, *249*
Dysphasia *see* Aphasia
Dysphonia *see* Hoarseness
Dysplastic naevus syndrome, 1258
Dyspnoea, 688, 819
aortic regurgitation, 767
aortic stenosis, 765
atrial septal defect, 775
cor pulmonale, 783
Fallot's tetralogy, 778
mitral regurgitation, 763
mitral stenosis, 761
paroxysmal nocturnal, 688
pulmonary embolism, 784
tricuspid stenosis, 769
see also Breathlessness
Dysraphism, 1169
Dyssynergia, 1114
Dysthymia, 1199
Dystonias, 1112, 1148, 1150
classification, *1150*
dopamine-responsive, 1150
oromandibular, 1150
primary torsion (PTD), 1150
treatment, 1150–1
Dystrophia myotonica, 43, *43*, *789*, 1181–2
and diabetes, *1031*
genetic anticipation, 45
Dystrophin, 29
Dysuria, 602

E2F-DP1, 50
E45 cream, *1270*
Ear, **1077–80**
anatomy and physiology, 1077, *1077*
disorders, 1078–80
discharge (otorrhoea), 1078
hearing loss *see* Deafness
tinnitus, 1080
vertigo, 1079–80
examination, 1078
inner, 1077
middle, 1077
outer, 1077
wax, 1077, 1078
see also Hearing; specific disorders
Ear drum *see* Tympanic membrane
Early after depolarizations, 716
Easy bruising syndrome, 435
Eating disorders, 1219–20
anorexia nervosa, 1219–20
atypical, 1220–1
bulimia nervosa, 1220
diabetic patients, 1045
obesity *see* Obesity
see also Anorexia; Anorexia nervosa; Bulimia nervosa
Ebola virus, *110*, *114*, 121
Ebstein's anomaly, 721, 768, *774*
EBV *see* Epstein-Barr virus (EBV)
Ecchymoses, *see* Bruising
Ecchymosis, *1227*
Eccrine sweat glands, 1227
Echinocandins, 100, *100*
Echinococcus
E. granulosus, *163*, 171, 361, *1179*
E. multilocularis, *163*, 171, 362
Echocardiography, 699–704, *700*, *701*, 904
aortic regurgitation, 767
clinical uses, 703–4
contrast echo, 704
Doppler *see* Doppler echocardiography
heart failure, 737
intravascular (coronary) ultrasound, 704
M mode, 699
myocardial perfusion, 702–3
physics, 699
quantification, 701–2
specific conditions
angina pectoris, 749
aortic aneurysm/dissection, 703–4
aortic stenosis, 766, *766*
arrhythmogenic right ventricular cardiomyopathy, 791
atrial septal defect, 776
cardiac masses, 704
cardiomyopathies, 704
congenital heart disease, 704
constrictive pericarditis, 797

COPD, 837
cor pulmonale, 784
dilated cardiomyopathy, 792, *793*
heart failure, 704
hypertrophic cardiomyopathy, *790*
infective endocarditis, 704, 772, *772*
ischaemic disease, 704
Marfan's syndrome, 781
mitral regurgitation, 763–4
mitral stenosis, 703, 762
pericardial effusion, 704, *704*, 796
primary restrictive non-hypertrophic cardiomyopathy, 793
prosthetic heart valves, 704
pulmonary embolism, 786
pulmonary hypertension, 783
pulmonary stenosis, 769
tricuspid stenosis, 768
valve regurgitation, 703
valve stenosis, 703
ventricular septal defect, 775, *775*
transoesophageal, 702
infective endocarditis, 772
mitral stenosis, 762
transthoracic, 701
infective endocarditis, 772
myocardial infarction, 757
wall motion stress, 702
see also Ultrasound
Echovirus, *70*, *110*, 111–12
Eclampsia, 362–3
Econazole, 100, *100*
Economics, non-adherence to treatment, 10
Ecstasy, 1215
overdose, *935*
clinical features, 939–40
treatment, 940
Ecthyma, 1228
Ectopic ACTH syndrome, 1027
Ectopic hormone secretion, 1027
Ectopic pregnancy, ruptured, 315–16
Ectropion, 1087
Eculizumab, paroxysmal nocturnal haemoglobinuria, 419
Eczema, *72*, **1236–40**
asteatotic, 1239, 1253
atopic, 1236–8
chronic renal failure, 629
classification, *1236*
contact/irritant, 1239–40, *1239*
discoid (nummular), 1238
food allergy, 237
hand, 1238
nappy rash, 1268
photosensitive, 1249
seborrhoeic, 1238–9, *1239*
HIV-related, 1269
venous, 1239, *1239*
see also Dermatitis
Eczema craquelé, 1239
Eczema herpeticum, 1237
Edinburgh Postnatal Depression Scale, 1200
Edinger-Westphal nucleus, 1103
Education (clinician), 15, 16–17
Education (patient)
heart failure, 737
Edwards' syndrome *see* Trisomy 18
Efalizumab, 76, *76*
psoriasis, 1242
Efavirenz, *193*, 195, *195*
side-effects, rash, 1269
teratogenicity, 199
Effect size, 930
Effective arterial blood volume, 650
reduced, 655
Eflornithine, 158
hirsutism, 1005
Ehlers-Danlos syndrome, 42, *435*, 568, 1225, 1263
mitral regurgitation, 763
renal tubular acidosis, *676*
Ehrlichia chafeensis, *148*
Ehrlichiosis, 148, *148*
Eicosanoids, *see* Leukotriene(s); Prostaglandin(s)
Eicosapentanoic acid, 210, *210*, 211, 496
Eisenmenger's syndrome, 774, 783
heart-lung transplantation, 846
Ejection fraction, 702
Elaidic acid, *210*
Elastase, 56, *57*, 273, 819
faecal, 376, 380
Elastin, 500
Elastography
cirrhosis, 346
non-alcoholic fatty liver disease, 345
Elation, 1187
Elbow pain, 508
epicondylitis, 508
golfer's elbow, 508
rheumatoid arthritis, 526
tennis elbow, 508, *508*
Elderly
adverse drug reactions, *924*
antidepressants in, 1203
back pain, 552
chondrocalcinosis, 553
cold haemagglutinin disease, 417
epilepsy, 1142
giant cell arteritis, 552
glomerular filtration rate, 646
gout, 553
heart disease, 797–8
hyperthyroidism, 988
malnutrition, 228
neck pain, 552
nutritional requirements, 227–8
polymyalgia rheumatica, 552
presbycusis, 1079, *1079*
prescribing in, 924, *924*
pseudogout, 553
pupil in, 1103
renal disease, 646–7
rheumatic complaints, 552–3, *552*
rheumatoid arthritis, 552, **552**
septic arthritis, 553
stroke, 1133
see also Stroke
urinary incontinence, 647
urinary tract infections, 647
see also Ageing
Electric shock, 959
Electrocardiogram (ECG), 694–7
12-lead, 695, *695*
24-hour ambulatory *see* Holter monitoring
advanced cardiac life support, 711
cardiac vectors, 696, *696*
event recording, 697
exercise, 696–7, *697*
angina pectoris, 749
heart rate variability, 697
Holter monitoring, 697, *698*
heart failure, 737
normal, *696*
P wave, 684, 695, *695*
PR interval, 684, 695, *695*
prolongation, 717, *717*
Q wave, 695
QRS complex, 684, 695, *695*
QT interval, 695
correction, 695, *695*
QT prolongation, 695–6
long QT syndrome *see* Long QT syndrome
torsade de pointes *see* Torsades de pointes
R wave, 695
S wave, 695
signal-averaged, 697
specific conditions
acute coronary syndromes (ACS), 752
angina pectoris, 749
aortic regurgitation, 767
aortic stenosis, 766, *766*
arrhythmogenic right ventricular cardiomyopathy, *789*, 791, *791*
atrial arrhythmias, *723*
atrial septal defect, 776
Brugada syndrome, 727, *727*
constrictive pericarditis, 797
COPD, 837
cor pulmonale, 783–4, *784*
dilated cardiomyopathy, 792
Fallot's tetralogy, 779
heart failure, 737
hypertrophic cardiomyopathy, *790*
infective endocarditis, 772
Marfan's syndrome, 781
mitral regurgitation, 763, *764*
mitral stenosis, 762, *762*
myocardial infarction, 755–7, *756*, *757*
myocarditis, 788
pericarditis, 795, *796*
poisoning, 939
primary restrictive non-hypertrophic cardiomyopathy, 793
pulmonary embolism, 785, *786*
pulmonary hypertension, 783
pulmonary stenosis, 769
tricuspid stenosis, 768
ventricular tachycardia, 726, *727*
ST segment, 696
depression, 696–7
elevation, 697
T wave, 695
T wave alternans, 697
waveform, 695–6, *695*
Electroconvulsive therapy (ECT), 1203
Electrocution, 959
Electroencephalography (EEG), 1120–1
coma, 1125
Creutzfeldt-Jakob disease, 1121
epilepsy, 1120–1, *1120*, 1141
portosystemic encephalopathy, 354
Electrolyte balance, 207
acute tubular necrosis, 623
Electrolytes
absorption, 274
diabetic ketoacidosis, *1049*
disturbances
acute abdomen, 315
Addison's disease, 1011
cirrhosis, 346
infection, *88*
hyperosmolar hyperglycaemic state, *1049*
infective endocarditis, 772
interstitial fluid, *650*
intracellular fluid, *650*
loss of, 660
from gut, *662*
myeloma, 485
non-Hodgkin's lymphoma, 480
parenteral nutrition, 236
plasma, *650*
neurology, *1122*
sweat, 843, 844
urine, 627
see also Bicarbonate; Chloride; Potassium; Sodium; Urea
Electromyography, 1121, 1178
polymyositis/dermatomyositis, 548
Electrospray ionization, 48
Eletriptan, migraine, 1137
Elliptocytosis, hereditary, *405*, 406
Eltrombopag, thrombocytic purpura, 436
Embden-Meyerhof pathway, 412, *413*
Embolectomy, 786
Embolic disease, 807
Embolism
pulmonary *see* Pulmonary embolism
see also Thromboembolism
Embricitabine, *192*, *195*
Embryogenesis, 994
Embryology, 994
Emesis *see* Vomiting
Emmetropia, 1086, *1086*
Emollients, *1270*
Emotion
and asthma, 849
medical interviews, 14
response to bad news, 14
Empathy
failure in clinical communication, 10
medical interviews, 12, 14
techniques, 12
Emphysema, 835, 836
centri-acinar, 836
computed tomography, 823, *823*
and diving, 957
heart sounds, 692
irregular, 836
pan-acinar, 836
subcutaneous, 914
see also Chronic obstructive pulmonary disease (COPD)
Empty sella syndrome, 975
Empyema, 368, 863, 886
subdural, 1159
Enalapril
heart failure, *739*
hyperaldosteronism, 1024
hypertension, 804
side-effects, hepatotoxicity, *366*
Encephalitis, *111*, 1156–7
acute disseminated, 1157
acute viral, *83*, 1156–7
as biological agent, 960
and epilepsy, 1141
herpes simplex, *105*
HIV-related, 1159
Japanese, *81*, *110*, 115–16, 1156
measles-related, 118
progressive rubella, 1158
tick-borne, *82*, *110*, 116
Encephalitis lethargica, 1146
Encephalitozoon, *161*
Encephalofacial angiomatosis, 1170
Encephalomyopathy, *37*
Encephalopathies
bovine spongiform (BSE), 122
see also Creutzfeldt-Jakob disease (CJD)
chronic traumatic, 1163
hypertensive, 1130
hyponatraemic, 662
portosystemic *see* Portosystemic encephalopathy (PSE)
static myoclonic, 1150
subcortical, 1168
transmissible spongiform (TSE), **121–2**, 1073
see also Creutzfeldt-Jakob disease (CJD); Prion diseases
Wernicke-Korsakoff syndrome *see* Wernicke-Korsakoff syndrome
Endemic goitre *see* Goitre
Endemic treponematoses, 142–3
Endoanal ultrasound, 244
Endocannabinoid system, 230
Endocardial disease *see* Myocardial disease
Endocarditis, *86*
antibiotic prophylaxis, *93*
Fallot's tetralogy, 779
infective *see* Infective endocarditis
Loeffler's, *789*
rheumatoid arthritis, 528

Endocardium
  infection *see* Infective endocarditis
  inflammation *see* Endocarditis
Endocrine disease/dysfunction, **963–1027**, *964*
  aetiology, 963, 965
    autoimmune disease, 965, *965*
    endocrine tumours, 965
    enzyme defects, 965
    hormones as therapy, 965
    receptor abnormalities, 965
  history and examination, 963, *964*
  HIV-related, 189
  and hypertension, 799
  symptoms, 963
  see also specific diseases, endocrine systems, hormones
Endocrine function, 968–9
  basal blood levels, 968
  hormone concentrations, 969
  saliva, 968
  stimulation/suppression tests, 968, *968*, *969*
  stress-related hormones, 968
  urine collections, 968
  see also Hormones
Endocrine myopathy, 1179–80
Endocrine organs, *964*
  see also specific endocrine organs
Endocrine therapy, 463, *463*
  breast cancer, 488, *488*
Endocrine tumours, 965
Endocrinology *see* Hormones; specific endocrine organs
Endocyte pattern recognition receptors, 59
Endocytosis, *19*
  receptor-mediated, *1062*
Endogenous infection, 81
Endogenous opioid peptides, 1118
  release in shock, 895
Endomyocardial biopsy, 788, 793–4, 797
Endomysial antibody, 279
Endonuclease, 28
Endoplasmic reticulum, *19*, 22–3, *23*
Endorectal ultrasound, 244
Endorphins, 1118
Endoscopic retrograde cholangiography (ERC), common bile duct stones, 371, *371*
Endoscopic retrograde cholangiopancreatography (ERCP), 327, 377
  pancreatitis, 378
Endoscopic ultrasound (EUS), 244
  carcinoma of oesophagus, 257
  common bile duct stones, 371, *371*
  gastric cancer, 265
  liver disease, 326
  pancreas, 376–7, 380
Endoscopy, 246
  oesophageal, 251
  specific conditions
    cirrhosis, 346–7
    liver disease, 327
    variceal haemorrhage, 350, *350*, 351
  upper GI, 268–9, 270
  video capsule, 246
    GI bleeding, 270
  wireless capsule, coeliac disease, 279
Endostatin, 50
Endothelial-derived hyperpolarizing factor (ERHF), 686, *686*
Endothelin, 320, 577
  asthma, 852
  heart failure, 736
  pulmonary hypertension, 782
  vasomotor control, 686, *686*, 687
Endothelin-1, 545, 896
Endothelium, 431
Endotoxins, 85–6, 87, 122, 323, 895
Endotracheal intubation, 913, *913*
  complications, *913*
Enema, *297*
  barium *see* Barium enema
Energy
  balance, 208
    bodyweight, *see* Weight
  expenditure, 208–9, *208*, *209*
  from food, *208*
  intake, 208
  production from food, *see* Food
  requirements, 208–9
    adults, 234
    enteral nutrition, *235*
    parenteral nutrition, 236
  stores, 209, *209*
  see also Calories
Energy expenditure
  obesity, 230
  theory of ageing, 227
Enfurvitide, *193*
Enoxaparin
  acute coronary syndromes, *754*
  myocardial infarction, 757
Enoximone, 907
Entamoeba
  E. dispar, 161
  E. histolytica, 85, 161, *161*
    amoebic abscess, 361
    diarrhoea, *307*
Entecavir, *101*
  hepatitis B, 339
Enteral nutrition, 234–5
  Crohn's disease, *289*
  home, 237
  management, 235
  nasogastric tube, *235*
  standard diet, 235, *235*
Enteric fever, 144–5
  clinical features, 144
  diagnosis, 144
  management, 144
  prevention, 144–5
Enteric nervous system, 271
Enteritis, radiation, 282
Enterobacter, 280
  antibiotic resistance, *94*
Enterobius vermicularis *see* Threadworm (*Enterobius vermicularis*)
Enterochromaffin cells, 376
Enteroclysis *see* Small bowel enema
Enterococci, *123*
Enterococcus, 280
  E. faecalis
    pyogenic abscess, 360
    septicaemia, *135*
    urinary tract infection, *600*
  E. faecium, 99
Enterocolitis, *86*
  staphylococcal, *307*
Enterocytes, 270–1
Enterocytozoon bieneusi, *161*, 201
Enterohepatic recirculation, 274, 323, *323*, 1062
Enterokinase, 273
Enteropathic arthritis, 536
Enteropathy
  food-sensitive, 237
  gluten-sensitive *see* Coeliac disease
  protein-losing, 246, 282
    tests for, 277
Enteropathy-associated T cell lymphoma, 279
Enteroscopy, push, 270
Enterostatin, 229
Enterotoxin, *128*, 129
Enteroviruses, *90*, 111–12
  enterovirus 68-71, *110*
  rash, *88*
Enthesis, 500, 501
Enthesitis
  ankylosing spondylitis, 533
  elbow, 508
  heel, 515
  knee, 513
  reactive arthritis, 535
Enthesitis-related arthritis, 553
Enthesophytes, 518
Entrapment syndromes
  nerve root
    lumbar, *511*
    neck, 506, *506*
Entropion, 1087, *1087*
Environmental medicine, 953–61
  bioterrorism/biowarfare, 959–60
  building-related illnesses, 960–1
  cold, **954–5**
  disease and environment, 953
  diving, 956–7
  drowning/near-drowning, 957–8
  electric shock, 959
  heat, **953–4**
  high altitudes, 955–6, *955*
  ionizing radiation, 958–9
  noise, 959
  smoke, 959
  travel, 960
Environmental temperature, 953
Enzyme-linked surface receptors, *22*
Enzymes
  deficiencies, 965
  pancreatic, 375
  see also individual enzymes
Eosinophil(s), 56–7, *56*
  granules, 57
  late phase response, 72
  measurement, *69*
  role of, 57
Eosinophil cationic protein, 57, 850
Eosinophil neurotoxin, 57
Eosinophil peroxidase, 57
Eosinophilia, *88*, 431
  causes of, *431*
  chronic renal failure, 627
  Churg-Strauss syndrome, 503
  perennial rhinitis, 832
  pulmonary *see* Pulmonary eosinophilia
  pulmonary infiltration, 874–6, *875*
  tropical pulmonary, 164
  see also Eosinophil(s)
Eosinophilic gastroenteritis, 283
Eosinophilic myalgia syndrome, skin changes, 1252
Eosinophilic oesophagitis, 255
Eosinophils, 431
  asthma, 850
  origins of, 388
Eosinophiluria, 627
Eotaxin, 57, 850
Eotaxin-1, 57
Ephelides, 1264
Epicondylitis, 508
Epidemic haemorrhagic fever, *114*
Epidemiology, 830, *830*
Epidermal growth factor, 852
Epidermis, 1225–6
Epidermoid cyst, 1257
Epidermolysis bullosa, 568, 1257
  dystrophica, 1257
  junctional, 1257
  simplex, 1257
Epidermophyton, 153
Epididymo-orchitis, 118
Epidural haematoma, 1164
Epidural haemorrhage, 1164
Epigastric pain, 242
  see also Abdominal pain
Epigastric pain syndrome, 311
Epigenetics, 51
Epiglottitis, acute, 127, 834
Epilepsia partialis continua, 1142
Epilepsy, **1139–44**
  aetiology, 1140–1
    alcohol and drugs, 1141
    brain tumours and abscesses, 1141
    developmental anomalies, 1140–1
    encephalitis, 1141
    genetic predisposition, 1140–1
    metabolic abnormalities, 1141
    photosensitivity, 1141
    pyrexia, 1141
    sleep deprivation, 1141
    trauma, hypoxia and surgery, 1141
    vascular/degenerative brain disorders, 1141
  aura, 1139
  birth defects, 1142–3
  classification, 1139–40
  diagnosis, 1141
  driving, 1144
  drug withdrawal, 1143
  EEG, 1120–1, *1120*
  elderly patients, 1142
  and fertility, 1142
  generalized seizures, 1139, *1139*
  grand mal, 1140
  incidence, *1096*
  mechanisms and definitions, 1139, *1139*
  neurosurgery, 1143
  partial seizures, 1139
  petit mal, 1140
  progressive myoclonic, 1150
  prohibited drugs, *923*
  seizure threshold, 1139
  seizures types, 1140, *1140*
  social consequences, 1143
  treatment, 1141–2, *1142*, *1143*
    anticonvulsants, 1142, *1143*
    emergency measures, 1141
    refractory epilepsy, 1142
    status epilepticus, 1142, *1142*
Epiloia *see* Tuberous sclerosis (epiloia)
Epimysium, 501
Epinephrine *see* Adrenaline (epinephrine)
Epipodophyllotoxins, 460
Epirubicin
  breast cancer, *461*, 488, 489
  gastric cancer, 266, *461*, 490
  oesophageal cancer, 489
Episcleritis, 528
Epistaxis (nosebleeds), 1080–1, *1081*
  aetiology, *1081*
  aplastic anaemia, 402
  coarctation of aorta, 777
  thrombocytopenia, 436
Epithelial basement membrane, 851–2
Epithelial cells
Epithelial growth factor (EGF), 451
Epithelial membrane antigen, *493*
Epitopes, 58
Eplerenone, *657*
  clinical uses, 657
  heart failure, *739*
  myocardial infarction, 759
Epley manoeuvre, 1079, 1109
Epoetin-alfa, 633
Epoprostenol, arterial thrombosis, 443, *443*
Eprodisate, 589, 1073
Epsilon-aminocaproic acid, hyperfibrinolysis, 441
Epstein-Barr virus (EBV), *70*, *90*, *104*, 107–8, *107*
  associated immunodeficiency, 70
  and Burkitt's lymphoma, 482
  clinical features, 107–8
  diagnosis, 108
  HIV-related, 203
  and Hodgkin's lymphoma, 476
  and non-Hodgkin's lymphoma, 479
  pyrexia, *91*
  rash, *88*
  and systemic lupus erythematosus, 542
  treatment, 108

Eptifibatide
  acute coronary syndromes, *754*
  angina pectoris, 750
  arterial thrombosis, *443*
Epworth sleepiness scale, 842, *1084*
Equality, patients' rights, 1, 7–8
Equivalence trials, 929
  analysis, 930–1
erb-A, 49
Erectile dysfunction, 634, *995*, 1000, 1119
  diabetes mellitus, 1036, 1056
Ergotamine, side-effects, retroperitoneal fibrosis, 616
Erlotinib, *464*
  lung cancer, 486
Erosion, *1227*
Error catastrophe theory of ageing, 226
Erysipelas, 1228
Erythema ab igne, *566*
Erythema chronicum migrans *see* Lyme disease
Erythema gyratum repens, malignancy-related, *1255*
Erythema infectiosum, 109, 540, 1230
Erythema multiforme, *105*, 1250
  clinical spectrum, *1251*
  drug-induced, *1265*
Erythema neonatorum, 1268
Erythema nodosum, 1250, *1250*, 1253
  drug-induced, *1265*, *1266*
  leprosum, 141
  rheumatological problems, 555
  sarcoidosis, 869
Erythraemia, 642
Erythrasma, 1228, *1228*
Erythroblasts, 390
  iron, *396*
  nucleated, *405*
Erythrocyte sedimentation rate (ESR), *88*, 389, *389*
  Hodgkin's lymphoma, 477
  musculoskeletal problems, 503
  myeloma, 484
  non-Hodgkin's lymphoma, 480
  pyogenic abscess, 361
Erythrocytes *see* Red blood cells (erythrocytes)
Erythrocytosis *see* Polycythaemia
Erythroderma, 1249–50
  aetiology, 1249, *1249*
  clinical features, 1249
  complications, 1249
  drug-induced, *1265*
  HIV-related, 1269
  malignancy-related, *1255*
  treatment, 1249–50
Erythroid hyperplasia, 403
Erythromycin, *94*, 97, 126
  acne vulgaris, 1246
  atopic eczema, 1238
  cellulitis, 1228
  chancroid, 180
  chlamydial conjunctivitis, 1088
  cytochrome P450 inhibition, *925*
  erythrasma, 1228
  gastroenteritis, *132*
  lymphogranuloma venereum, 177
  pneumonia, 862
  septic arthritis, 539
  side-effects
    acute tubulointerstitial nephritis, *604*
    hepatotoxicity, *366*
    long QT syndrome, *727*
  syphilis, 179
Erythro(parvo)virus B19, arthritis in, 540
Erythropoiesis, 389–90, *390*
Erythropoietic protoporphyria, *1074*, 1075
Erythropoietin, 388, 389, 390, 576–7
  deficiency, 628, 633
  disadvantages of therapy, 633
  lymphoplasmacytic lymphoma, 481
  polycythaemia, 421
  recombinant, 397
Erythrovirus B19, *104*, 109
  bone marrow aplasia, 410
  rash, *88*
Erythroviruses, 109
Escherichia, 280
  E. coli, *70*, *80*, 84–5, *90*, 122
    CAPD peritonitis, *637*
    diffusely adhering, *128*
    enterohaemorrhagic, *128*, 130
    enteroinvasive, *128*, 130
    enteropathogenic, *128*
    enterotoxigenic, *128*, 130
    enterotoxin, *306*
    food poisoning, *133*
    gastroenteritis, 130
    pyogenic abscess, 360
    septicaemia, *135*, *136*
    urinary tract infection, 599, *600*
    ventilator-associated pneumonia, 914
  FimH, 599
  PapG, 599
Esmolol, *731*
  supraventricular tachycardia, 723
Esomeprazole, *252*
Essential fatty acids (EFAs), 210, *210*
  deficiency, 211
Essential hypertension, 607, 798–9
  environmental factors, 798–9
    alcohol intake, 798
    obesity, 798
    sodium intake, 798–9
    stress, 799
  fetal factors, 798
  genetic factors, 798
  management, 607
  pathophysiology, 607, 799–800
Estimated average requirement (EAR), 207
Etanercept
  ANCA-positive vasculitis, 597
  asthma, *854*, 856
  psoriasis, 1242
  psoriatic arthritis, 535
  rheumatoid arthritis, *530*, 531
Ethambutol, tuberculosis, 866
Ethanol *see* Alcohol
Ethics, 1–8
  clinical resources, 7–8
  clinical trials, 929
  ethical/legal boundaries, 6–8
    competent refusal, 6–7
    duty of equality, 7–8
    patient's best interest, 7
    prolonging life, 7
  importance to medical practitioners, 2
    clinical importance of trust, 2
    doctor-patient relationship, 2
    the law, 2
      European Convention on Human Rights, 2
      Human Rights Act (1998), 2
    professional concern, 1
    professional regulation, 2
      see also General Medical Council (GMC)
    rational self-interest, 2
    three duties of clinical care, 1–2
  medical mistakes, 3
    clinical error, 3
    clinical negligence, 3
    expert witnesses, 3
    importance of honesty, 4
    inexperience, 3
  respect for autonomy, 2, 3–6
    confidentiality, 4–5
    legally valid consent, 3–5
    vulnerable patients, 5–6
  why study medical ethics/law?, 1–2
Ethionamide, tuberculosis, 866
Ethnicity, 1185–6
Ethosuximide, epilepsy, *1143*
Ethyl alcohol (ethanol) *see* Alcohol
Ethylene glycol
  metabolism, *944*
  poisoning, 943–4
    antidotes, *937*
    clinical features, 943
    metabolic acidosis, *939*
    treatment, 943–4
Ethylene oxide, anaphylactic reactions, 635
Etidronate
  osteoporosis, *564*
  Paget's disease, 566
Etoposide, 460
  germ cell tumours, 492
  Hodgkin's lymphoma, *461*
  lung cancer, *461*, 487
  ovarian cancer, 491
  side-effects, leukaemia, 467
Etravirine, *193*, 195
Euchromatin, 33
Eukaryotes, 33, 80
Eumycetes, 153
European Convention on Human Rights, 2
Eustachian tube, 1077
Euthyroidism
  compensated, 967, 986–7
    thyroid function tests, *984*
Evans' syndrome, 416, 436
Everolimus, renal cell carcinoma, 645
Evidence-based medicine, 929–32
  controlled observational studies, 931
  randomized controlled trials, 924, 929–31
  uncontrolled observational studies, 931
Evolutionary theories of ageing, 227
Ewart's sign, 796
Excitation-contraction coupling, 681, 682–3, *683*
Excoriation, *1227*
Excretion urography (intravenous pyelography: IVP), 580
  chronic renal failure, 628
  urinary tract obstruction, 615, *615*
Exemestane, 463
Exenatide, 1039
Exercise
  and coronary artery disease risk, 746
  and depression, 1201
  diabetes mellitus, 1038
  heart failure, 738
  obesity, 230, 232
Exercise tests, *824*, 827
  asthma, 852
  electrocardiogram, 696–7, *697*
  heart failure, 737
Exercise-induced asthma, 848–9
Exfoliatin, 125
Exocytosis, 21
Exotoxins, 85, 895
Expanded Programme on Immunization, 103
Expert witnesses, negligence cases, 3
Express consent, 4
Expressed sequence tags, 36
Extracellular fluid (ECF), 649
  volume regulation, 649–50
Extracellular matrix, 500
Extracellular volume
  decreased, 659–61
    causes, 659–60, *660*
    clinical features, 659
    investigations, 660, *660*
    treatment, 660–1, *660*
  homeostasis, *651*, 653
  increased, 655–9
    causes, 655–6
    clinical features, 655
    treatment *see* Diuretics
  regulation, 650–1, *651*
    neurohumeral, 651–2
    oedematous conditions, 653
    vs osmoregulation, 654–5
Extracorporeal gas exchange, 916
Extracorporeal shock wave lithotripsy *see* Shock wave lithotripsy
Extradural haemorrhage, 1135–6
Extraocular muscle weakness, 1105
Extrapyramidal system, 1111, *1111*, 1112–13
  anatomy, 1113, *1113*
  connections, *1113*
  disorders of, 1113
  function, 1113
  neurotransmitters, *1113*
Extrarenal volume receptors, 652
Extrinsic allergic alveolitis, 876–8, *876*
  clinical features, 876
  differential diagnosis, 877
  investigations, 877
  management, 877
  pathogenesis, 876
Extrinsic apoptotic pathway, 28
Extubation, 917
Eye(s), **1085–94**
  anatomy and physiology, 1085–6, *1085*, *1086*
    see also Cornea; Lens; Retina
  blood supply, 1086
  cornea *see* Cornea
  dry, *see* Keratoconjunctivitis sicca
  innervation, 1086
  lens *see* Lens
  movements *see* Ocular movements
  refractive errors, 1086–7, *1086*
  retina *see* Retina
  specific conditions
    diabetes *see* Diabetic eye disease
    HIV, 188
    hyperthyroidism *see* Thyroid eye disease
    rheumatoid arthritis, 528
    sickle cell anaemia, 410
    systemic lupus erythematosus, 543
    see also entries beginning ocular; and specific disorders
Eyeball, *1085*
Eyelids, 1085, 1087, *1087*
  blepharitis, 1087, *1087*
  dacryocystitis, 1087, *1087*
  ectropion, 1087
  entropion, 1087, *1087*
Ezetimibe, *1066*, 1068

Fabry's disease, 597, 788, *789*, 1072
  diagnosis, 1072
Face masks, 912, *912*
Facet joint syndrome, 510
Facial appearance
  asymmetry, 1125
  mitral stenosis, 761
Facial nerve, 1107–8
Facial pain, 1138
  atypical, 1138
  giant-cell arteritis, 1138
Facial rashes, *1245*, **1245–8**
Facial weakness, 1107
  causes, 1107
Facioplegic migraine, 1137
Factitious fever, pyrexia, *91*
Factor assays, 435

Factor IX, 433
  concentrate, 429
  haemophilia B (Christmas disease), 439
Factor IXa, 433
Factor V, 442
  activated, 433
Factor V Leiden, 42, 442
Factor VII, 433
  and coronary artery disease, 747
Factor VIIa, recombinant, 428
Factor VIII, 433, *493*
  administration of, 439
  autoantibodies, 441
  concentrate, 429, 439
  lack of *see* Haemophilia A
  recombinant, 439
  synthesis, *438*
  vitamin K deficiency, *439*
  von Willebrand disease, *439*
Factor VIII gene, 438
Factor X, 433
Factor Xa, 433
Faecal chymotrypsin, 376
Faecal culture, 88, *89*
Faecal elastase, 376, 380
Faecal incontinence, 297, *297*, *298*
Faecal occult blood, *270*, 304
  colorectal carcinoma, 305
Faecal osmolar gap, 309
Faeco-oral infections, 82
  prevention, 83
Faeculent vomit, 242
Faget's sign, 114
Fainting *see* Syncope
Fallot's spells, 778
  treatment, 779
Fallot's tetralogy, 721, 768, *774*, 778–9, 893
  auscultation, *779*
  clinical features, 778–9
  clubbing, 690, 779
  investigations, 779
  pathophysiology, *779*
  right bundle branch block, *720*
  squatting posture, 774, 778
  treatment, 779
Falls, *1096*, 1097
Famciclovir, 101, *101*, *107*
  genital herpes, 181
  shingles, 1231
Familial adenomatous polyposis, 285, 301, *301*, *453*
  attenuated, 301, *301*
Familial amyloidotic polyneuropathy, 1072
Familial bilateral striatal necrosis, *37*
Familial cancer syndromes, *453*
Familial combined hyperlipidaemia, 1064, *1064*
Familial defective apoprotein B, *1064*
Familial fatal insomnia, 145
Familial hypercholesterolaemia, 746, 1064, *1064*, 1067
  gene therapy, 46
  rheumatological problems, 556
Familial hyperkalaemic hypertension *see* Gordon's syndrome
Familial hypertriglyceridaemia, *1064*
Familial hypocalciuric hypercalcaemia, 560, 1021
Familial hypomagnesaemia, hypercalciuria and nephrocalcinosis, 669
Familial iminoglycinuria, 1071
Familial isolated vasopressin deficiency, 1017
Familial Mediterranean fever, 556
  acute abdomen, *315*
  pyrexia, *91*
  treatment, 556
Family history
  and coronary artery disease, 745
  genetic, 45
  musculoskeletal disorders, 502
Family therapy, 1204
Famotidine, 252
Fanconi syndrome, 567, 575, 599, 1069–70
Fanconi's anaemia, 453, *453*
  leukaemia risk, 467
Fansidar, *156*
  side-effects, hepatotoxicity, *366*
Farmer's lung, *876*
Farnesol X, 322
Farr assay, 503
Fas, *28*
Fas ligand, *28*, 66, 451
Fas-associating protein with death domain, *28*
Fasciculi, 501
Fasciola hepatica, *163*, 170
  treatment, *170*
Fasciolopsis buski, *163*, 170
  treatment, *170*
Fasting, 271
  hormonal regulation, 968, 1030, *1032*
  lipids, 1062
Fat
  absorption, 273–4, *273*
  dietary requirements, 210–11
    diabetes mellitus, *1037*
  intake
    current recommendations, 211, 212
    reducing, 1065
  malabsorption, 277
  metabolism disorders, hyperlipidaemia *see* Hyperlipidaemia
  trans, 210
  see also Fatty acids; Lipid(s)
Fatal familial insomnia, 122
Fatigue, 1097
  chronic, syndrome *see* Chronic fatigue syndrome (CFS)
  cor pulmonale, 783
  Fallot's tetralogy, 778
  heart failure, 689
  mitral regurgitation, 763
  muscle *see* Muscle(s)
  pulmonary stenosis, 768
Fatty acids, 210, *210*, 321
  dietary sources, *210*
  essential fatty acids (EFAs) *see* Essential fatty acids (EFAs)
  cis-monounsaturated, 211
  non-esterified, *213*
  cis-polyunsaturated, 211
  polyunsaturated (PUFAs) *see* Polyunsaturated fatty acids (PUFAs)
  saturated, 211
  trans, 210
Fatty liver, *331*, 344–5, 358
  acute fatty liver of pregnancy, 363
  investigations, 359
  management, 359
Fatty streaks, 744
Favism, 413
Fc receptors, 415, 422
  see also Antibodies
Febrile neutropenia, *461*
Febuxostat, gout, 538
Feeding, hormonal regulation, 968
Feet *see* Foot
Feline leukaemia virus subgroup C, 393
Felodipine, 941
  hypertension, 805
Felty's syndrome, 423, *430*
  rheumatoid arthritis, 528
Females
  epilepsy, 1142–3
  gonorrhoea, 175
  hypothalamic-pituitary-gonadal axis, *995*
  idiopathic oedema, 656
  lower abdominal pain in, 242
  menopause, 462, 997–8
  musculoskeletal disease in, 501
  osteoarthritis, 520
  osteoporosis, 563
  puberty, 996
  reproductive disorders, **1001–6**
  reproductive physiology, 995–6, *995*
  rheumatoid arthritis, 523
  subfertility, 1006, *1007*
Femoral arteries, pistol shot sound, 767
Femoral head, avascular necrosis, 513
Femoral neck fracture, 512
Fenestration, 714
Fenfluramine
  side-effects
    mitral regurgitation, 763
    respiratory, *877*
Fenofibrate
  hyperlipidaemia, *1066*
  side-effects, hepatotoxicity, *366*
Fenoldopam, hypertension, 805
Ferrihaem, 403
Ferritin, 346
  haemochromatosis, 356
  iron storage, 394–5
  serum, microcytic anaemia, 396, *396*
Ferroportin, 356, 394–5
Ferrous sulphate, 396
  fixed dose prescribing, *925*
  side-effects, 396
    hepatotoxicity, *366*
Fertility problems, 1005
see also Subfertility
Festinant gait, 1097
α-Fetoprotein *see* Alpha(α)-fetoprotein (AFP)
Fetor hepaticus, 343, 353
Fetuin, 631
Fetus
  abnormalities, folate deficiency, 401
  circulation, 687–8, *687*
  haemoglobin, 406
  macrosomia, 1058
  maternal Graves' disease, 990–1
Fever, 86–7
  and epilepsy, 1141
  infective endocarditis, *770*, 773
  leucocytosis, 430
  non-haemolytic transfusion reactions, 427
Fever of unknown origin *see* Pyrexia of unknown origin (PUO)
Fexofenadine, rhinitis, 833
Fibrates, 746, 1068
Fibre *see* Dietary fibre
Fibreoptic bronchoscopy, 828–9, *828*
  bronchial carcinoma, 882, *882*
  tuberculosis, 866
Fibric acid derivatives, *1066*
Fibrillin-1 gene, 781
Fibrin, 433
  degradation *see* Fibrinolysis
  deposition in hypertension, 607
  see also Coagulation
Fibrin degradation products, 434, 440, *441*, 898
Fibrinogen, 432
  activation, 433
  and coronary artery disease, 747
  degradation, *see* Fibrinolysis
Fibrinolysis, 433–4, *434*
  excessive, 441
  myocardial infarction, 757
  pulmonary embolism, 786
  see also Fibrin degradation products (FDPs)
Fibrinopeptide(s), 433
Fibrinopeptide A, 782
Fibroblast growth factor, 451, 670
  rickets/osteomalacia, 567
Fibroblast growth factor receptor defect, 568
Fibroblasts
  synovial, 524
  systemic sclerosis, 546
Fibroglycan, 500
Fibromodulin, 500
Fibromyalgia *see* Chronic pain syndromes
Fibronectin, 432, 500, 557
  glomerulopathy, 592
Fibropolycystic disease, 360
Fibrosing alveolitis, 822
  in rheumatoid arthritis, 871
Fibrosis
  liver, 325, 345, *345*, 356
    congenital, 360
    markers of, 325
    perivenular, 358
  pulmonary *see* Pulmonary fibrosis
  syndrome *see* Chronic pain syndromes
Fibrositic nodulosis, 510
Fibrositis syndrome (fibromyalgia) *see* Chronic pain syndromes
Fibrosure test, 325
Fibrotest, 325
Fifth disease *see* Erythema infectiosum
Filaggrin, 1226, 1236
Filariasis, 164–5
  clinical features, 164
  diagnosis, 164
  lymphatic, *82*, 164
  treatment, *163*, 165
Filiform warts, 1231
Filoviruses, *110*, *114*, 121
Filtration pores, 582
Fimbria, 25, 85
Financial compensation, 2
Finasteride
  androgenic alopecia, 1267
  benign prostatic hypertrophy, 646, 998
  hirsutism, 1005
  prostate, cancer, 646
Fine-needle aspiration (FNA)
  breast cancer, 487
  goitre, 993
Finger clubbing, 329, *455*, 690, 820, *821*
  bronchial carcinoma, 881
  causes, *820*
  congenital heart disease, 774
  Fallot's tetralogy, 690, 779
  idiopathic pulmonary fibrosis, 872
  infective endocarditis, *770*
  tuberculosis, 865
Finkelstein's sign, 508
Fish oils, 211, 496
Fissures, *1227*
Fistulae
  in ano, 300, *300*
  arteriovenous, 635, *635*
Fitzhugh-Curtis syndrome, 175, 316
Fixed drug rash, *1265*
Flare, 57
Flat foot, 515
Flatulence, 242
Flaviviruses, *80*, *82*, *110*, *114*, **114–16**
  see also individual diseases
Flea-borne spotted fever, *147*
Fleas, *82*
Flecainide, 730, *731*
  genetic polymorphism in metabolism, *926*
  supraventricular tachycardia, 723
Flenley acid-base nomogram, *675*
Flexor tenosynovitis, 508
Flight of ideas, 1187

Flow-volume curves, *824*
Flow-volume loops, 816–17, *816*, 825–6
Flucloxacillin, 95, 125
  atopic eczema, 1238
  bronchiectasis, 843
  cellulitis, 1228
  echthyma, 1228
  folliculitis, 1228
  impetigo, 1228
  infective endocarditis, *773*
  osteomyelitis, 541
  septic arthritis, 539
  side-effects, hepatotoxicity, *366*
Fluconazole, 100, *100*
  blastomycosis, 152
  candidiasis, 183
  cryptococcosis, 151
5-Flucytosine, 100, *100*
Flucytosine, cryptococcosis, 151
Fludarabine, 460
  acute lymphocytic leukaemia, 475
  lymphoplasmacytic lymphoma, 481
  side-effects, immune haemolytic anaemia, 417
Fludrocortisone
  Addison's disease, *1012*
  potency, *1008*
  see also Mineralocorticoids
Fluid
  extracellular *see* Extracellular fluid (ECF)
  intake
  replacement, *see* Rehydration solution for malnutrition
  see also Water
Fluid balance, acute tubular necrosis, 623
Flukes *see* Trematodes (flukes)
Flumazenil, *937*
Fluocinolone acetonide, *1237*
Fluorescein angiography, 1052
Fluorescein dilaurate test, 376
Fluorescence in-situ hybridization, *41*
  chronic myeloid leukaemia, 473
Fluoride, 226
  recommended IV intake, 233
Fluorosis, 226
5-Fluorouracil, 460
  Bowen's disease, 1258
  breast cancer, *461*, 488
  colorectal carcinoma, *461*, 490
  gastric cancer, 266, *461*, 490
  oesophageal cancer, 489
  pancreatic carcinoma, 384
  pancreatic neuroendocrine tumours, 384
Fluoxetine
  genetic polymorphism in metabolism, *926*, *927*
  overdose, 949
Fluphenazine, schizophrenia, 1217
Flutamide, 463
  hirsutism, 1005
  prostate cancer, 492
Fluticasone
  asthma, *854*, 855
  rhinitis, 833
Fluvastatin, hyperlipidaemia, *1066*
Fluvoxamine, overdose, 949
Fms-tyrosine kinase 3, 388
Foam cells, 587, *587*
Focal adherent junctions, *25*
Focal segmental glomerulosclerosis, 586–7
  aetiology, 586
  clinical features, 586
  pathology, 586–7, *587*
  treatment, 587
  variants of, 587, *587*
Folates/folic acid, 211, 223, 400–1
  antagonists, 460
    see also Methotrexate
  deficiency, *218*, 397, 400–1
    causes, *400*
    clinical features, 401
    fetal abnormalities, 401
    investigations, 401
    treatment, 401
  dietary intake, 400
  liver disease, 325
  LRNI, *218*
  myelofibrosis, 421
  pernicious anaemia, 400
  recommended IV intake, *234*
  RNI, *218*
  structure, *400*
  supplementation, 223, 401
    thalassaemia, 408
  see also Tetrahydrofolate (THF)
Folinic acid, 460
  colorectal carcinoma, *461*, 490
  toxoplasmosis, 201
Follicle-stimulating hormone (FSH), *971*, 996
  secretion, 967
Follicular lymphoma, *480*, 481–2
  international prognostic index (FLIPI), 481
  management, 481
  prognostic factors, *481*
  treatment, 481–2
Follicular occlusion triad, 1246
Folliculitis, 1228
  HIV-related, 1269
  Pityrosporum, 1234
  pruritic, of pregnancy, 1269
  therapy-related, 1270
Fomepizole, *937*
  alcohol poisoning, 944
Fondaparinux
  arterial thrombosis, *443*
  venous thrombosis, 445
Fonsecaea, 153
Food
  energy from, *208*
  fortification of, 212
  intake in obese persons, 228
  see also Diet; Nutrition
Food allergy, 73, 237, 1194
Food intolerance, 237–8
  management, 238
Food poisoning (and food-borne infections), 82, *86*, 131–3, *132*, *133*
  bacterial causes, *133*
  oral rehydration solutions, *132*
  organic toxins, *133*
  prevention, 83
  see also Gastroenteritis
Food-sensitive enteropathy, 237
Foot
  arthritis, *515*
  diabetes mellitus *see* Diabetic foot
  flat, 515
  high-arched, 515
  ulcers
Foot pain, 515–16
  causes, *515*
  hallux rigidus, 515
  hallux valgus, 515
  metatarsalgia, 515
  rheumatoid arthritis, 526, *527*
  stress (march) fractures, 515
  tarsal tunnel syndrome, 515
Foramen ovale, 687
Forbes disease, *1069*
Forced expiratory volume (FEV), *824*, 825, *825*
Forced vital capacity (FVC), *824*, 825, *825*
Forchheimer spots, 113
Foreign bodies
  cornea, 1089, *1090*
  inhalation, 834–5, *834*
  nose, 1081
  pharynx, 1084
Foreign language, communication difficulty, 16
Forgetfulness
  communication difficulties, 11
  see also Amnesia
Formoterol, asthma, *854*
Forrestier's disease, 512
Fosamprenavir, *193*, *195*, 196
Foscarnet, 101, *101*
  cytomegalovirus, 107, 202
  nephrotoxicity, 189, 202
  oral hairy leukoplakia, 1269
Foxp3, Treg failure, 74
Fractures
  crush, of spine, 512
  femoral neck, 512
  foot, 515
  nose, 1082
  osteoporosis, 552, 561, *562*
  rib, 884
  scaphoid, 409
  skull, 1119, 1163
  vertebral, 486, 512, 561, 562, 564
Fragile X syndrome, *40*, *43*
Fragilitas ossium, 568
Frame-shift mutations, 41
Frank-Starling relationship, 683–4, *684*, *891*
Freckles (ephelides), 1265
Free androgen index, 1004
Free fatty acids (FFAs), 321
Free radicals, 20, 223–4
  and ageing, 227
  scavengers, 20
Freezing, 1147
Fresh frozen plasma, 428, 429
  thrombotic thrombocytopenic purpura, 437
Fridericia's cube root formula, 695, *695*
Friedreich's ataxia, 36, *43*, *789*, *1114*, 1170
  and diabetes, *1031*
Friedreich's sign, 793, 797
Frizzled related protein 4, 670
Frontal lobe disease *see* Dementia
Frontotemporal dementia, 1168
Frostbite, 955
Frovatriptan, migraine, 1137
Frozen shoulder, 507, 508
Fructosamine, 1044
Fructose, 211
  hereditary intolerance, 1068
  metabolism defects, 1068
Fructose-1,6-diphosphatase deficiency, 1068
Fructosuria, 1068
Frusemide *see* Furosemide
Fulminant hepatic failure, 342–3, *343*
  causes, *343*
  clinical features, 342–3
  course, 343
  investigations, 343
  prognosis, 343
  transfer criteria, *343*
  treatment, 343, *343*
Fulvestrant, 463
Fumes, 849
Functional disorders, *1192*, **1192–5**
  chest pain, 310–11
  chronic fatigue syndrome, 1192–3
  chronic fatigue syndrome (CFS), 516, 1192–3
  diarrhoea, 313
  dyspepsia, 31
  fibromyalgia, 1193
  gastrointestinal *see* Functional gastrointestinal disorders (FGID)
  irritable bowel syndrome, 237, 311–13, *312*, 1194
  management, *1193*
  menopause, 997–8, 1195
  multiple chemical sensitivity, 1194
  oesophagus, 310–11
  platelets, 437
  premenstrual syndrome, 1194–5
  urinary incontinence, 647
  vomiting, 311
  see also Chronic pain syndromes
Functional endoscopic sinus surgery, 1081
Functional gastrointestinal disorders (FGID), *309*, **309–10**
  acute abdomen, *315*
  bowel disorders, 311–14
    see also Irritable bowel syndrome
  brain-gut interactions, 310, *310*
  diarrhoea, 313
  gastroduodenal disorders, 311
  oesophageal disorders, 310–11
  pathophysiology, 310
  Rome III criteria, *310*
Functional MRI, brain, 1120
Functional residual capacity (FRC), 814, *816*
Fundus, 1124
Fungal infections, **149–53**, *150*
  HIV-associated, 200–1
  joints, 541
  skin, 1232–4
  subcutaneous, 152–3
  superficial, 153
  systemic, 149–52
  see also specific infections/fungi
Fungal pericarditis, 795
Furazolidone, *H. pylori* eradication, 261
Furosemide, *657*
  acute tubular necrosis, 622
  ascites, 353
  heart failure, *739*
  hypertension, 804
  nephrotic syndrome, 585
  side-effects, acute tubulointerstitial nephritis, *604*
Furuncles *see* Boils
Fusidic acid, 98
  impetigo, 1228
  osteomyelitis, 541
  septic arthritis, 539
Fusion genes, 49
Fusion inhibitors, *193*, 194
Fusobacterium, *123*
  skin and soft tissue infections, *124*

G-CSF, 389
  cancer therapy, 464
G-protein(s), 24
G-protein coupled receptors, *22*, 966
  antidiuretic hormone, 653
GABA (γ-aminobutyric acid), 1098, 1113
Gabapentin, *495*
  chronic widespread pain, 516
  diabetic neuropathy, 1055
  overdose, 940
  trigeminal neuralgia, 1107
Gados channel, 409
Gag reflex, 818, 1082
  poisoning, 934
Gaisböck's syndrome, 421
Gait
  abnormalities, 1096–7, *1096*
    cerebellar lesions, 1114
    Parkinson's disease, 1146
  broad-based, 1097
  festinant, 1097
  osteomalacia, 567
  stamping, *1096*, 1097, 1117
  waddling, *1096*, 1097
Gait apraxia, *1096*, 1097
Gait ataxia, 1111
GAL4 enhancer, 31
Galactokinase deficiency, 1068
Galactorrhoea
  chronic renal failure, 629
  drug-induced, *964*

Galactosaemia, 1068
d-Galactose, 424
Galactose-1-phosphate uridyltransferase, 1068
Galanin, *272*, 376
Galavardin's tachycardia, 729
Gall bladder, 320, 372
  adenocarcinoma, 374
  adenomyomatosis, 372
  cholesterolosis, 372
  chronic cholecystitis, 372
  motility, 367
  non-calculous cholecystitis, 372
  palpable, 331
  polyps, 373
Gall bladder disease, *314*
Gallstone colic, 368
Gallstones, 367–70
  cholesterol, *see* Cholesterol gallstones
  clinical presentation, 368, *368*
  complications, 372
  differential diagnosis, 369
  formation, 281
  investigations, 368–9
  management, 369–70
    cholescystectomy, 369
    stone dissolution and shock wave lithotripsy, 369–70
  and pancreatitis, 379
  prevalence, 367
  sickle cell anaemia, 410
  types, 367–8
  see also specific types
  see also Common bile duct stones
Gambierdiscus toxicus, 950
Gametogenesis, 36
Gamma camera, 245, 705
Gamma globulin
  see also Immunoglobulin(s)
γ-aminobutyric acid *see* GABA
γ-glutamyl transpeptidase, 324
  common bile duct stones, 370
γ-hydroxybutyric acid, poisoning, 944
Ganciclovir, 101, *101*
  cytomegalovirus, 107, 202
  human herpesvirus, 108
  myelotoxicity, 189, 202
  oral hairy leukoplakia, 1269
Ganglion, 509
Ganglion blockers, postural hypotension, *659*
Gangrene, *82*
Gap junctions, *25*, 26
Gardner syndrome, *301*, 302
Gardnerella vaginalis, 183
Gas gangrene, 125, *1179*
Gas transfer, *824*, 826
Gastric adenocarcinoma, 262, 263–6
  see also Gastric cancers/tumours
Gastric banding, 233, 1039–40
Gastric bypass surgery, 1039–40
Gastric cancers/tumours, **263–7**, 490
  adenocarcinoma, 262, 263–6
    clinical features, 265, *265*
    diagnosis, 265, *265*
    early, 264–5, *264*
    epidemiology and pathogenesis, 264, *264*
    gastric bleeding, 269
    incidence, 263–4
    pathology, 265
    screening, 264–5, *264*
    staging, 265–6, *266*
    treatment, 266
  chemotherapy, *461*
  epidemiology, *449*
  gastric polyps, 266
  gastrointestinal stromal tumours, 266
  lymphoma
    B cell, 262
    primary, 266, 483
  prognostic factors, 490
  treatment, 490
Gastric distension, 496
Gastric emptying, 258
  radionuclide studies, 245
  see also Gastrointestinal motility
Gastric lavage, 937
Gastric outlet obstruction, 262
Gastric polyps, 266
Gastric tumours *see* Gastric cancers/tumours
Gastric ulcers, 260, *261*
  H. pylori infection *see* *Helicobacter pylori*
  and NSAIDs, 263
  perforation, 261–2
  treatment, 263
  see also Peptic ulcer disease
Gastrin, *272*, 384
Gastrin-releasing peptide, *272*, 376
Gastrinoma, 384, 385
Gastritis, 262–3
  antral, 259
  autoimmune, 262, 263, 399
Gastro-oesophageal reflux disease (GORD), **250–3**
  clinical features, 251, *251*
  complications, 252–3, *253*
  diagnosis and investigations, 251–2, *251*
  outcome, *252*
  pathophysiology, 250–1, *250*
  surgery, 252
  treatment, 252
Gastro-oesophageal variceal haemorrhage, *267*, 349–51
  management, 268, 350–1, *350*
    acute episodes, 350–1
    acute rebleed, 351
    balloon tamponade, 350
    endoscopy, 350, *350*
    injection sclerotherapy, 350
    resuscitation, 350
    variceal banding, 350
    vasoconstrictors, 350
  prevention of recurrence, 351
  prophylactic measures, 351
Gastroenteritis, *80*, 128–31, *128*
  antibiotics, *132*
  Bacillus cereus, 131
  Campylobacter jejuni, 129
  clinical syndromes, 129
  clostridial infections, 131
  eosinophilic, 283
  Escherichia coli, 130
  management, 129
  mechanisms, 128–9, *128*
  Salmonella, *128*, 129
  Shigella, *128*, 129–30
  Staphylococcus aureus, 130–1
  travellers' diarrhoea, 131, *131*
  Vibrio, 130
  yersiniosis, 130
  see also Diarrhoea
Gastrografin, 244
Gastrointestinal bleeding, 246, **267–70**
  acute lower GI tract, 269–70, *269*
  acute upper GI tract, 267–9
    aetiology, 267, *267*
    clinical approach, 267
    discharge policy, 269
    drug therapy, 269
    emergency management, 267–8, *268*
    prognosis, 269
    repeat bleeding, 269
    risk assessment, 267, *268*
  chronic, 270
  Rockall score, *268*
Gastrointestinal complications
  see also Gastrointestinal symptoms
Gastrointestinal disease, **241–318**
  cancer, *241*
  clinical examination, 243–4
    abdomen, 243, *243*
    rectum and sigmoid colon, 243–4, *244*
  clinical investigations, 244–6
    barium contrast *see* Barium contrast studies
    computed tomography, 244, *245*
    magnetic resonance imaging, 244
    plain X-rays, 244
    positron emission tomography, 244
    ultrasound, 244
  functional *see* Functional gastrointestinal disorders (FGID)
  infections *see* Gastrointestinal infections
  malignancy incidence, *see* Gastric cancers/tumours
  rheumatological problems, 554–5
  symptoms and signs, 241–3
    abdominal pain, 242–3
    anorexia and weight loss, 243
    diarrhoea and constipation, 242
    dyspepsia and indigestion, 241
    flatulence, 242
    vomiting, 242, *242*
  see also Gastrointestinal tract; specific diseases
Gastrointestinal infections, **128–33**
  developing/tropical countries, 143–5
    cholera *see* Cholera
    enteric fever, 144–5
  food poisoning, 131–3, *132*, *133*
  gastroenteritis, 128–31, *128*
    Bacillus cereus, 131
    Campylobacter jejuni, 129
    clinical syndromes, 129
    clostridial infections, 131
    Escherichia coli, 130
    management, 129
    mechanisms, 128–9, *128*
    Salmonella, 129
    Shigella, 129–30
    Staphylococcus aureus, 130–1
    travellers' diarrhoea, 131, *131*
    Vibrio, 130
    yersiniosis, 130
  protozoal, 160–2, *161*
Gastrointestinal motility
  abnormal, 307
  gall bladder, 367
  oesophagus, 253–4
  small intestine, 271
  see also Gastric emptying
Gastrointestinal stromal tumours, 266
Gastrointestinal symptoms
  cancer-related, 495–6
  HIV-related, 189
  systemic lupus erythematosus, 543
  see also Gastrointestinal complications; specific symptoms
Gastrointestinal tract
  diabetic neuropathy, 1056
  functional disorders *see* Functional gastrointestinal disorders (FGID)
  radiation effects, 958
  see also Gastrointestinal disease; specific regions
Gastroparesis, 1056
Gastropathy, 262–3
  portal, 262
Gastroscopy, 245, 246, *246*
  gastric cancer, 265, *265*
GATA-1, **33**
Gate theory of pain, 1118
Gatifloxacin, 98
Gaucher's disease, *789*, 1071
  splenomegaly, 423
Gaze
  deviation, 1134
  lateral, 1104–5
  vertical, abnormalities of, 1105
  see also Conjugate lateral gaze
GCA *see* Giant-cell arteritis
Gefitinib, *464*
Gel electrophoresis, 33
Gelatinases, 500
Gelofusin, 429, 906
Gels, *1270*
Gemcitabine
  breast cancer, 489
  lung cancer, *461*, 486
  ovarian cancer, 491
  pancreatic carcinoma, 384
Gemeprost, 1006
Gemfibrozil, hyperlipidaemia, *1066*
Gemifloxacin, 98
Gender, 1008
  and coronary artery disease, 745
Gender role disorders, 1221
Gene(s), 29
  defects, *see* Genetic disorders
  mapping *see* Gene mapping
  sequencing, *35*
  see also DNA (deoxyribonucleic acid)
Gene chips, 34–5
Gene defects, 40–4
  mutations, 40–2
  single-gene disease, 42–4
Gene expression, 30–3
  control regions/proteins, 31, 33
  DNA-binding proteins, 31, *32*
  transcriptional control, 31
Gene knock-out, 47
Gene mapping, 36, 230, 1170
Gene regulation theory of ageing, 226
Gene splicing, *31*
Gene therapy, **46–9**
  adenosine deaminase deficiency, 46
  animal models, 47–8
  cancer, 465
  cystic fibrosis, 46, *46*, 845
  familial hypercholesterolaemia, 46
  muscle-cell-mediated, 46–7
  somatic disease, 47
General anxiety disorder (GAD), 1206–7, *1207*
see also Anxiety
General Medical Council (GMC)
  legal redress, 1
  professional regulation, 2
    good medical practice, 2
  standards of conduct, 1
General paralysis of the insane, 1157
General transcription factors, 31
Genetic anticipation, 45
Genetic counselling, 44–5
  congenital heart disease, 774
  Marfan's syndrome, 781
  see also Prenatal diagnosis
Genetic disorders, **38–51**
  avoidance and control, 46
  chromosomal *see* Chromosomal disorders
  clinical genetics, 44–5
  gene defects, 40–4
  gene therapy *see* Gene therapy
  and leukaemia risk, 467
  obesity, 228
  population genetics, 44
  prenatal diagnosis, 45
  prenatal testing *see* Prenatal diagnosis
  see also Cancer; Genetic factors (disease aetiology); Mutation(s); specific disorders
Genetic factors (disease aetiology)
Genetic polymorphisms, 35–6, 925–6, *925*, *926*
Genetic screening, 45, 454
  see also Prenatal diagnosis

Genetics
  cancer *see* Cancer
  clinical, *see* Genetic counselling; Prenatal diagnosis
  ethical considerations, 48
  see also Genetic disorders
Genital herpes, *105*, **180–1**
  clinical features, 180–1, *180*
  diagnosis, 181
  management, 181
  in pregnancy, 181
  prevention and control, 181
  recurrent attacks, 181
  ulceration, *172*
  urethral discharge, *172*
  vaginal discharge, *172*
  see also Herpes simplex virus (HSV)
Genital ulceration, *172*
Gentamicin, *94*, 96, *96*
  infective endocarditis, *773*
  pneumonia, 863
  septic arthritis, 539
  side-effects, 1080
    acute tubular necrosis, 617
  therapeutic drug monitoring, *927*
  urinary tract infections, 602
Genu valgus, 502
Genu varus, 502
Geographic tongue, 247
Germ cell tumours *see* Ovarian cancer; Testicular cancer
German measles *see* Rubella
Germinoma, 974
Gerstmann-Straussler-Scheinker syndrome, 121, 122
Gestational diabetes, 1058–9
Gestodene, 1006
Ghon focus, 864
Ghrelin, 229, *272*, 376, 977
Giant congenital melanocytic naevi, 1258
Giant-cell arteritis, *549*, 550, 1138–9
  clinical features, 1138–9
  elderly patients, 552
  investigations, 550
  management, 1139
  pyrexia, *91*
  treatment, 550
Giant-cell myocarditis, 788
Giardia intestinalis, 161, *161*, 277, 282
  diarrhoea, *307*
Giardiasis, 161–2, *162*
  clinical features, 162
  diagnosis, 162
  management, 162
Gigantism, 972, 979–80
  treatment, 979–80
Gilbert's syndrome, 329
Gilles de la Tourette syndrome, 1150, 1210
Gingivae, 247–8
  swelling, 248
Gingivitis, 247–8
  acute (necrotizing) ulcerative, 247
  aplastic anaemia, 402
  chronic, 247
  desquamative, 247
GISSI Prevention Trial, 211
Gitelman's syndrome, 26, *658*, 665–6
Glanders, as biological agent, 960
Glands
  endocrine, 965
  sebaceous, 1227
  sweat, 1227
Glandular fever *see* Infectious mononucleosis (glandular fever)
Glanzmann's thrombasthenia, 26, 437
Glasgow Coma Scale, *1123*
Glasgow scoring system, 378
Glatiramer acetate, multiple sclerosis, 1153
Glaucoma, 1091–2, *1091*
  acute angle-closure, 1091–2, *1092*
  primary open-angle, 1091
Gleason score, 491
Gliadin, 277
Glibenclamide, *1038*
Glicentin, *272*
Gliclazide, *1038*
Glimepiride, *1038*
Glioblastoma multiforme, 1160
Glioma, 974, 1118, 1160, *1160*
  see also specific tumours
Glipizide, *1038*
  genetic polymorphism in metabolism, *926*
Glitazones, diabetes mellitus, 1039
Global Registry of Acute Coronary Events (GRACE) prediction score, 752
Globin chains, 390, 406
  alpha, 406
  beta, 406
  defects, 412
    beta-thalassaemia, 407
    production and structure, 412
  delta, 406
  epsilon, 406
  fetal, 406
  gamma, 406
  genes *see* Globin genes
  synthesis, 406
Globin genes, 406, *406*
  mutations
Globus, 310
Globus pallidus, 1113
Globus pharyngeus, 1084
  see also Dysphagia
Glomerular arterioles, 571
Glomerular filtration rate (GFR), 572, 574, 650–1
  depression/reduced
  elderly patients, 646
  measurement of, 574–5
  pressures controlling, *573*
Glomerulonephritis
  acute nephrotic syndrome *see* Nephrotic syndrome
  anti-GBM, 595–6, *595*
  autoimmune
  IgA nephropathy, 593–4, *593*
  with infective endocarditis, 593
  lupus, 590–1, *590*
  mesangiocapillary (membranoproliferative), 589–90, *590*
  non-streptococcal post-infectious, 592, *592*
  post-streptococcal, 592–3, *593*
  rapidly progressive, 584, 595–7
    types of, *595*
  with visceral abscesses, 593
  see also Glomerulopathies; Nephritic syndrome, acute; specific types
Glomerulonephropathies, mesangial proliferative, 590
Glomerulopathies, 583–98
  classification, 584
  fibronectin, 592
  haematuria *see* Haematuria
  idiopathic fibrillary, 592
  immunotactoid, 592
  inflammatory *see* Glomerulonephritis
  membranous *see* Membranous glomerulopathy
  nephrotic syndrome *see* Nephrotic syndrome
  pathogenesis, 584
  pathological terms, 584
  pre-eclampsia-related, 597–8
  see also Glomerulonephritis; Glomerulosclerosis; specific disorders
Glomerulosclerosis, 624, 1054
  focal segmental *see* Focal segmental glomerulosclerosis
  HIV-associated, 597
Glomerulus, 571–2
  basement membrane, 582
  Bowman's capsule, 572, 582
  diseases/disorders, 582–3
    glomerulonephritis *see* Glomerulonephritis
    glomerulopathies *see* Glomerulopathies
    glomerulosclerosis *see* Glomerulosclerosis
    HIV-associated nephropathy, 597
    investigation, *585*
    thin glomerular basement membrane disease, 594
    see also specific disorders
  filtration barrier (slit diaphragm), 582–3, *583*
  filtration pores, 582
  mesangium, 582
Glomus tumour, *1110*
Glossina, 157
  G. morsitans, 157
  G. palpalis, 157
Glossitis, 247
  folate deficiency, 401
  pernicious anaemia, 399
Glossopharyngeal nerve, 1110
  lesions, 1110
Glossopharyngeal neuralgia, 110
Glucagon, *272*, *937*, 1030
  calcium channel blocker overdose, 941
  release in shock, 895
Glucagon-like peptide, *272*, 276, 280
Glucagon-like peptide 1, 229, 232
Glucagon-like peptide-1, *272*
Glucagonoma, 385, *1031*
Glucocorticoid axis, **1008–16**
Glucocorticoid response element, 1008
Glucocorticoids, 76, 1008
  abnormalities, 1009
  actions, *1008*
  basal levels, 1009
  Crohn's disease, *289*
  production, 1008
  side-effects, *928*
    osteoporosis, 565
    renal calculi, 611
  uses, *1015*
  see also Corticosteroids; Cortisol
Glucokinase, *1035*
Gluconeogenesis, 213, 321, 1029
  see also Glucose, production/synthesis
Glucosamine supplements, 522
Glucose, 211, *660*
  blood, 1029
    Addison's disease, 1011
    home testing, 1044
    pancreatitis, *378*
  dialysate
    haemodialysis, *634*
    peritoneal dialysis, *637*
  fasting, 1034
    impaired, 1037
  homeostasis, 321, 1029
  metabolism, *413*, 1029
  production/synthesis, 1029
    see also Gluconeogenesis
  profile, *1040*
  transporters, 1030
    GLUT-1, 1030
    GLUT-2, 1030
    GLUT-3, 1030
    GLUT-4, 1030, 1031
  in urine *see* Glycosuria
  utilization, 1029–30
    see also Hyperglycaemia; Hypoglycaemia
Glucose tolerance, impaired *see* Impaired glucose tolerance (IGT)
Glucose tolerance test (GTT), 979
  see also Impaired glucose tolerance (IGT)
Glucose-6 phosphate dehydrogenase (G6PD) deficiency, 410, 412–14, *413*
  blister cells, *414*
  clinical syndromes, 413–14
  drug-induced haemolysis, 413, *414*
  gene localization, 413
  investigations, 414
  treatment, 414
Glucose-dependent insulinotropic peptide, *272*
Glucosylceramide lipidoses, 1071
Glue ear, 1079, *1079*
GLUT *see* Glucose, transporters
Glutamate, 210, 1098
Glutamic acid decarboxylase, 1033
Glutamine, 210, 898
Glutathione, 210, 224, 898
Glutathione peroxidases, 20
Glutathione S-transferase, 818
Gluten-free diet, 279, 1256
Gluten-sensitive enteropathy *see* Coeliac disease
Gluteus medius tendonopathy, 512
Glycaemic control *see* Glucose, blood
Glycaemic index, 1038
GlyCAM-1, 67
Glycerol, 213, 1029
  suppositories, *297*
Glycerophosphocholine, 655
Glyceryl trinitrate
  acute coronary syndromes, *754*
  angina pectoris, 750
  myocardial infarction, 757
Glycine, 210, 1098
Glycocalyx, 271
14C-Glycocholic acid breath test, 277
Glycogen, 1029
  breakdown *see* Glycogenolysis
Glycogen storage disease, *789*, 1068
  hepatomegaly, *331*
  liver, *1069*
  muscle, *1069*
Glycogenolysis, 213, 321
Glycopeptides, 99
Glycoprotein Ib, 432
Glycoprotein IIb-IIIa, 432
  receptors, 753
Glycoprotein IIb/IIIa antagonists, 687
  acute coronary syndromes, 753–4, *754*
  arterial thrombosis, 443, *443*, *444*
Glycoproteins, 500
Glycopyrronium, 496
Glycosaminoglycans, 500, 610
Glycosuria
  chronic renal failure, 627
  diabetes mellitus, 1036
  renal, 575
Glycosylated haemoglobin, 1044
Glycosylphosphatidylinositol, 418
GM-CSF, 58
  cancer therapy, 464
GM2 gangliosidoses, 1072
Goblet cells, *835*
Goitre, 225, 963, **992–3**
  assessment, 992
  causes, *992*
  clinical features, 992
  diffuse, 992
  endemic, 985
  euthyroid, 993
  fibrotic, 993

Goitre (*Continued*)
investigations, 993
malignancy, 993, *993*
multinodular, 992
simple, 992
solitary nodular, 992–3
toxic multinodular, 987
treatment, 993
types, *992*
see also Hyperthyroidism
Gold
rheumatoid arthritis, 531
side-effects
bone marrow aplasia, 402
membranous glomerulopathy, 588
rash, *1265*
Goldman perimetry, 972
Golfer's elbow, 508
Golgi apparatus, *19*, 23, *1030*
Gonadal dysgenesis, *1007*
Gonadal function tests, 998, *998*
Gonadal sex, 1008
Gonadotrophin
deficiency, 974
failure, *1002*
PCOS, 1004
see also Follicle-stimulating hormone (FSH); Luteinizing hormone (LH)
Gonadotrophin-releasing hormone (GnRH), 967, *971*
see also Luteinizing hormone releasing hormone (LHRH)
Gonadotrophin-releasing hormone (GnRH) agonists, 463
Gonococcal arthritis, 540
Gonococcal conjunctivitis, 1088
Gonococcus, *70*
Gonorrhoea, **175–6**
clinical features, 175, *175*
men, 175, *175*
women, 175
history, 175
screening for, **174**
treatment, 176
urethral discharge, *175*
see also Neisseria gonorrhoeae
Gonyanlax
G. catenella, 950
G. tamarensis, 950
Good medical practice (GMP), 2
European Convention on Human Rights, 2
GMC standards, 2
Goodpasture's syndrome, 75, *595*, 874
autoantigens, *74*
haemoptysis, 819
HLA system in, *60*
treatment, 874
Good's syndrome, 888
Gordon's syndrome, *658*, 667
Goserelin, 463, 967
prostate cancer, 491, 492, 646
Gottron's papules, 547, 1251
Gout, **536–9**
acute, 537–8
chronic polyarticular, 538
chronic tophaceous, 538–9, *539*
clinical features, 537–8
diet, 538
elderly patients, 553
epidemiology, 536–7, *537*
investigations, 538
pain, 501
pathogenesis, 537, *537*
precipitating factors, 537
chronic renal failure, 629
diuretics, 501
synovial fluid, 504
treatment, 538
uric acid levels, *537*
see also Hyperuricaemia
Gracile nuclei, 1115
Gradenigo's syndrome, 1106
Graft-versus-host disease, 70, 462, 463
skin changes, 1252
transfusion-associated, 426
Graft-versus-leukaemia effect, 473
Graft-versus-tumour effect, 462
Graham Steell murmur, 761
cor pulmonale, 783
Gram-negative bacteria
infections
pneumonia, 861
urinary tract, 579
protein secretion, *86*
Gram-positive bacteria, 56, 94–5, 98–9, 122, 123, 149
Grand mal fits, 1140
Granisetron, 461
Granular osmophilic material, 1127
Granulocyte concentrates, 429
Granulocyte-colony-stimulating factor *see* G-CSF
Granulocyte-monocyte colony stimulating factor *see* GM-CSF
Granulocytes, *56*
origins of, 388
survival, 387
see also Basophil(s); Eosinophil(s); Neutrophil(s)
Granuloma annulare, 1245, 1254
HIV-related, 1269
Granuloma inguinale, 180
Granuloma telangiectaticum, 1258
Granulomatosis, Wegener's *see* Wegener's granulomatosis
Granulomatous arteritis *see* Giant-cell arteritis
Granulomatous hepatitis, pyrexia, *91*
Granulomatous lung disease, 868–70
with vasculitis, 870–1
Granulomatous (primary cerebral) angiitis, 1159
Granulomatous tubulointerstitial nephritis (TIN), *606*
Granzyme B, 66
Grapefruit juice, cytochrome P450 inhibition, *925*
Graves' disease, 75, 541, *965*, 987
autoantigens, *74*
eye signs, 988, *988*
see also Thyroid eye disease
fetal effects, 990–1
skin manifestations, 1254
Grawitz tumour *see* Renal cell carcinoma
Greetings, medical interview, 11
Grey Turner's sign, 377
Grief, pathological, 1209
Griseofulvin, 100, *100*
Grommets, 1079
Ground itch, 167
Growing pains, 554
Growth, 977–9
assessment of, 977–8, *978*
factors affecting, 977
failure, 978–9
normal, 977–8
retardation, 978–9
sickle cell anaemia, 410
Growth factors
haemopoiesis, 388–9, *388*
insulin-like *see* Insulin-like growth factor (IGF-1)
osteoarthritis, 520
see also individual growth factors
Growth hormone (GH), 384, *971*, 1030
control, 977, *977*
deficiency, 974
excess, 972
see also Acromegaly
physiology, 977
recombinant, 978–9
replacement therapy, *976*, 978–9
secretion, 967
status, 978, 979
Growth hormone (GH) antagonists
acromegaly, 980
pituitary tumours, *973*
Growth hormone (GH) receptor antagonists, pituitary tumours, *973*
Growth hormone release inhibitory hormone (GHRIH) *see* Somatostatin
Growth hormone-releasing hormone (GHRH), *272*, 970, *971*
Growth retardation, congenital heart disease, 774
Growth spurt, 978
GTP-binding proteins *see* G-protein(s)
Guanethidine, chronic regional pain syndrome type 1, 517
Guanine, 29, *30*
Guanylate cyclase, 687
Guanylyl cyclase-linked receptors, *22*
Guillain-Barré syndrome, 108, 886, 916, **1172–3**
clinical features, 1172–3
respiratory failure, 911
ventilatory failure, 912–13
course and management, 1173
diagnosis, 1173
Gums, 211
mucilaginous, *297*
Gums *see* Gingivae
Gut decontamination, 937
Gut macrophages, 58
Gut microflora, 280
and inflammatory bowel disease, 286
Gut motility *see* Gastrointestinal motility
Gut-associated lymphoid tissue (GALT), 53, *274*, 275
Gymnopilus, 952
Gynaecological disorders, *314*
Gynaecomastia, 329, 1000–1
causes, *1000*
drug-induced, 352
older males, 1001
pubertal, 1001

H_1-receptor antagonists
polycythaemia vera, 420
rhinitis, 833
see also Antihistamines
H_2-receptor antagonists, 252, 263
pancreatitis, 381
polycythaemia vera, 420
see also Antihistamines
HAART, 188
HIV-associated nephropathy, 597
initial therapy, *195*
opportunistic infections, 199, *199*
patient monitoring, 196, *196*
see also Antiretroviral drugs (ARVs)
Haem, 390, *390*
Haem carrier protein 1, 393
Haemaccel, 429
Haemagglutin antigen, 834
Haemagglutinins, 116, 117, *133*
Haemangioblastoma, *1114*
Haemangioma, cavernous, 1267–8, *1268*
Haemarthrosis, 438
arthritis, 556
knee, 514–15
Haematemesis, 241, 242, 265, 267
Haematological disease, **387–447**
anaemia *see* Anaemia
malignancies *see* Haematological malignancies
see also Blood; specific diseases
Haematological malignancies, **466–86**
leukaemia *see* Leukaemia
lymphoma *see* Lymphomas
myeloma *see* Myeloma
Haematomas
epidural, 1164
intracerebral, 1133–4
subdural, 1135
Haematoxylin bodies, 542
Haematuria, 578
asymptomatic, 584
causes, *579*
chronic renal failure, 627
glomerulopathies
IgA nephropathy, 594
infective endocarditis, *770*
red-cell casts, *579*
ureteric colic, 612
with/without sub-nephrotic range proteinuria, 593
Haemobilia, 373
Haemochromatosis, 395
and diabetes, 1031
genetic analysis, 325
hereditary *see* Hereditary haemochromatosis (HH)
secondary, 395
skin manifestations, 329, 356, 1265
Haemodiafiltration, 938–9
continuous, 623–4
Haemodialysis, 634–6, 938–9
access, 634–5, *634*, *635*
adequacy of, 635–6
arteriovenous (radiocephalic) fistula, 635, *635*
basic principles, 634
complications, 635
dialysate buffer, 635
dry weight, 635
frequency and duration, 635
intermittent, 910
prescription, 635
see also Dialysis
Haemodynamics
anaphylaxis, *899*
cardiac tamponade, *899*
cardiogenic shock, *899*
hypovolaemic shock, *899*
pulmonary embolism, *899*
Haemofiltration, 636, *636*
acute tubular necrosis, 623
continuous arteriovenous/venovenous, 623
Haemoglobin, *389*, 826
abnormalities, 406–7
diagnosis, 412
gene therapy, 412
geographical distribution, *407*
in pregnancy, 412
prenatal screening, 412
see also specific diseases
alterations in relation to plasma, *392*
artificial, 428
conformations, 391
COPD, 838
deoxygenated, *391*
electrophoresis, *407*, 410
fate after haemolysis, *404*
function, 391, *391*
glycosylated, 1044
Gower 1, 406
Gower 2, 406
Hb A, 390, 391, *402*
Hb A2, 390, *406*
Hb Barts, *406*, 409
Hb C, *406*
Hb F, 390, *402*
Hb H, *406*
Hb S (sickle cell), *406*, 409
musculoskeletal problems, 503
oxygen binding, 391
see also Oxyhaemoglobin dissociation curve
oxygenated, *391*
Portland, 406
structure, 390, *390*

globin chains/genes *see* Globin; Globin genes
haem *see* Haem
synthesis, 390–1, *390*
Haemoglobin Constant Spring, 42
Haemoglobinaemia, 621
Haemoglobinopathies, *see* specific diseases
Haemoglobinuria, 427
march, 419
paroxysmal cold, 417
paroxysmal nocturnal, 402, 418–19, 435
Haemolysin, 599
Haemolysis, *391*, *404*
consequences, *see* Haemolytic anaemia
consequences of, 403, *403*
drug-induced, 413, *414*
evidence for, 403–4, *404*
extravascular, 403, *416*
fate of haemoglobin, *404*
hereditary spherocytosis, 405
intravascular, 403, 415
lead poisoning, 945
microangiopathic, 598
sickle cell anaemia, 410
Wilson's disease, 357
see also Haemolytic anaemia
Haemolytic anaemia, **403–19**
acquired, 414–19, *415*
alloimmune, 417–18
causes, 415, *415*
drug-induced, 417
mechanical, 419
non-immune, 418–19
see also Autoimmune haemolytic anaemia (AIHA)
inherited, 404–14
haemoglobin abnormalities, 406–7
metabolic red cell disorders, 412–14
red cell membrane defects, 404–6
sickle syndrome, 409–12
thalassaemias, 407–9
see also specific disorders
Haemolytic disease, compensated, 403
Haemolytic disease of newborn (HDN), 417–18, 424
clinical features, 417–18
investigations, 418
postnatal management, 418
prevention, 418
see also Rh blood group
Haemolytic jaundice, 329
Haemolytic transfusion reactions, 427
delayed, 427
diagnosis, 427
investigations, 427
Haemolytic uraemic syndrome, 598–9, *598*
diarrhoea-associated, 598
Pneumococcus-associated, 599
Haemoperfusion, 938–9
Haemopexin, 403
Haemophilia(s), 46
acute abdomen, *315*
see also Coagulation; specific types
Haemophilia A, 43, 438–9
antenatal diagnosis, 439
carrier detection, 439
clinical and laboratory features, 438–9, *439*
complications, 439
inheritance, 438
treatment, 439
see also Factor VIII
Haemophilia B (Christmas disease), 439
Haemophilus
H. ducreyi, 179
H. influenzae, *70*, *90*, *123*, 191
antibiotic resistance, *94*
bacterial meningitis, 134
bronchiectasis, 843
conjunctivitis, 1088
epiglottitis, 834
pneumonia, *858*, 860
septic arthritis, 539
type b, vaccine *see* Hib vaccine
Haemopoiesis, 387–8
at birth, 387
extramedullary, 387, 422
radiation effects, 958
see also Bone marrow
Haemopoietic growth factors, 388–9, *388*
cancer therapy, 464
therapeutic use, 389
Haemopoietic stem cell transplantation, **462–3**
allogeneic, 462–3
acute myeloid leukaemia, 470
chronic myeloid leukaemia, 473
follicular lymphoma, 482
non-myeloablative, 463
autologous, 462
Haemoptysis, 819
pulmonary embolism, 784
Haemorrhage
epidural, 1164
extradural, 1135–6
gastrointestinal *see* Gastrointestinal bleeding
intracerebral *see* Intracerebral haematoma/haemorrhage
intracranial *see* Intracranial haemorrhage
subarachnoid *see* Subarachnoid haemorrhage (SAH)
treatment, 660
variceal *see* Gastro-oesophageal variceal haemorrhage
see also Bleeding
Haemorrhagic disease of newborn, 440
Haemorrhagic intra-alveolar exudate, 918
Haemorrhoids, 243, 299
Haemosiderin, 395, 403
Haemosiderosis, transfusion-related, 408
Haemostasis, 431–5
platelets in, 432
vessel wall, 431–2
see also Coagulation; Thrombosis
Haemostatic plug, 431, *432*
Haemovigilance, 425
Hair, 1227
brittle, 395
excessive growth, 963, 1267
see also Hirsutism
hypertrichosis, 1003
loss, 1266–7
chemotherapy-induced, 461
see also Alopecia
normal, 1003
vellus, 1003
Hair cells, 1077
Hair disorders
dissecting cellulitis, 1267
hypertrichosis, 1267
telogen effluvium, 1267
Hairy cell leukaemia, 475
Haldane effect, 391
Halitosis, 246
Hallpike's test, 1079, 1109
Hallucinations, 1188, *1188*
Hallucinogens, 1215
Hallux rigidus, 515
Hallux valgus, 515
Haloperidol
behavioural disturbance, 1191
genetic polymorphism in metabolism, *927*
nausea, 461
sickle cell anaemia, *411*
side-effects
hepatotoxicity, *366*
long QT syndrome, *727*
Halothane, side-effects, hepatotoxicity, 365, *366*
Hamartoma, pulmonary, 884
Hamartomatous polyps, 302
Hamman-Rich syndrome, 872
Hand
foot and mouth disease, 111, *111*
rash, *88*
Handicap, 502
Hands
eczema, 1238, *1238*
osteoarthritis, *521*
pain, 508–9
causes, *508*
rheumatoid arthritis, 526
Hansen's disease *see* Leprosy (Hansen's disease)
Hantavirus, *110*, *114*, 116
Haplotype maps, 36
Hard ticks, *82*
Hard-water syndrome, 635
Hardy-Weinberg equation, 44
Harrison's sulcus, 567
Hartnup's disease, 222, 1071
Hasenclever score, 477
Hashimoto's thyroiditis, 541, 985
autoantigens, *74*
goitre in, 992
HAV *see* Hepatitis A virus (HAV)
Haversian system, 557
Hayfever *see* Seasonal rhinitis
Hazard ratio, 931
Head injury
headache, 1096
mechanical ventilation, 913
see also Brain damage/injury; Skull fractures
Head tilt, 710
Headache(s), 1096, 1136–8
chronic (benign), 1096
cluster, 1138
coarctation of aorta, 777
cough, 1137
following head injury, 1096
giant-cell arteritis, 1138
ice-cream, 1137
incidence, *1096*
low CSF volume, 1137
management, 1137
see also Analgesics
mechanisms, 1096
meningitis, *see* Meningitis
migraine *see* Migraine
post-traumatic, 1138
pressure, 1096
recurrent, 1096
scalp tenderness, 1096
sex-induced, 1138
subacute onset, 1096
tension, 1136
thunderclap, 1135
Health authorities, limited resources, 7–8
Health outcomes
adherence to treatment, 9
effect of communication, 9
failure of communication, 10–11
improving outcome, 10
Health promotion, 212
Healthcare
communication, 8–11
consumerism, 9
expenditure, *928*
outcomes, 9
resources, 7–8, 10
see also Clinical practice
Hearing
loss *see* Deafness
physiology of, 1077–8
tests of, 1078, *1078*, 1108
Hearing-impaired patients
communication problems, 16
see also Deafness
Heart
abnormal rhythm *see* Cardiac arrhythmias
anatomy, 681–8
apex beat *see* Apex beat
auscultation *see* Auscultation
block *see* Heart block
cells *see* Myocardial cells (myocytes)
conduction system, 681–3, *682*
atrial/ventricular myocyte action potentials, 682, *682*
atrioventricular node, 682
excitation-contraction coupling, 681, 682–3
His bundle, 682
Purkinje fibres, 682
sinus node, 681–2, *682*
disease *see* Heart disease
embryology, 681–8
failure *see* Heart failure
physiology, 681–8
rate variability (HRV), *see* Heart rate
rheumatoid arthritis, 528
size, 692–3
Starling's law, 683–4, *684*
systemic lupus erythematosus, 543
transplant *see* Cardiac transplantation
valves *see individual valves*
see also entries beginning cardiac
Heart beats
'missed' *see* Palpitations
Heart block, 717–20
atrioventricular *see* Atrioventricular (AV) block
bundle branch *see* Bundle branch block
Heart disease
arrhythmias *see* Cardiac arrhythmias
congenital *see* Congenital heart disease
elderly patients, 797–8
investigations
chest X-ray *see* Chest X-ray
ECG *see* Electrocardiogram (ECG)
echocardiography *see* Echocardiography
see also individual techniques
ischaemic *see* Ischaemic heart disease
symptoms, 688–9
central chest pain, 688
dyspnoea, 688
fatigue, 689
palpitations, 688–9
peripheral oedema, 689
syncope, 689
valvular *see* Heart valves
see also Cardiovascular disease; specific types of disease
Heart failure, **733–43**
acute, 742
clinical syndromes, *742*
decompensated, *742*
diagnosis, 742
hypertensive, *742*
mechanical assist devices, 743, *743*
pathophysiology, 742
treatment, 742–3, *743*
see also Cardiac arrest; Myocardial infarction
air travel, 738
apoptosis, 735
calcium homeostasis, 735
causes, *734*
clinical syndromes, 736, *736*
see also Cardiac cachexia
congestive (biventricular)
diagnosis, 736–7, *737*
diastolic, 736
driving, 738

Heart failure (*Continued*)
echocardiography, 704
endothelial function, 736
extracellular volume regulation, 653
heart sounds, 692
hepatomegaly, *331*
high output, *742*
infective endocarditis, *770*
inter-aortic balloon pumping, 714
investigations, 737
jugular venous pressure, 691
left heart failure *see* Left ventricle (ventricular)
myocardial gene expression, 735
natriuretic peptides, 735–6
New York Heart Association Classification, *736*
oedema, 655
pathophysiology, 733–5, *734*
myocardial contractility, 734
myocardial remodelling, 734–5, *735*
outflow resistance (afterload), 734
salt and water retention, 734
venous return (preload), 734
post-STEMI, 758
Killip classification, 758
right heart failure *see* Right ventricle (ventricular)
right-sided, *742*
mitral regurgitation, 763
mitral stenosis, 761
pulmonary stenosis, 768
tricuspid regurgitation, 768
sexual activity, 738
signs, 736
symptoms, 736
treatment, 737–41
drug management, 738–40, *739*, *740*
lifestyle advice, 737–8
monitoring, 738
non-pharmacological, 740–1, *741*
see also specific agents
vaccination, 738
Heart murmurs, 692
aortic stenosis, 766
Austin Flint, 767
Graham Steell, 761, 783
infective endocarditis, *770*
mitral regurgitation, 763
pulmonary stenosis, 768
Heart rate, 891
irregularities *see* Bradycardias; Tachycardias
variability, 697
ventricular tachycardia, 726
see also Cardiac output (CO: Q1)
Heart sounds, 692
first, 692
second, 692
third, 692
fourth, 692
aortic stenosis, 766
auscultation, 692, 761–2, *761*
see also Auscultation
murmurs *see* Heart murmurs
Heart transplants *see* Cardiac transplantation
Heart valves
disease *see* Valvular heart disease
magnetic resonance imaging, 707
prosthetic, 769, *769*
complications, 769
echocardiography, 704
mechanical vs tissue, 769
mitral valve, 763
para-prosthetic leak, 769
see also individual valves
Heart-lung transplantation, 846
Heartburn, 242, 249, 251
see also Gastro-oesophageal reflux disease (GORD)
Heat, **953–4**
acclimatization, 953
cramps, 953
see also Malignant hyperpyrexia
Heat illness, 953–4
management, 954
Heat injury, 954
management, 954
prevention, 954
Heat shock proteins, 20–1, 61
Heberden's nodes, *520*
Heel pain, *515*
Achilles tendon bursitis, *515*, 516
Achilles tendonosis, *515*, 516
anterior tibial syndrome, 516
calcaneal bursitis, 515
compartment syndromes, 516
enthesitis, 515
plantar fasciitis, 515, *515*
plantar spurs, 515, *515*
Sever's disease, 515, *515*
Heerfordt's syndrome, 248
Height
charts, *978*
prediction of, 978
Heimlich manoeuvre, 709, *834*
Heinz bodies, 415, 422
Helicobacter pylori, *123*, **258–62**, *263*
diagnosis, 260
biopsy urease, 260
13C urea breath test, 245, 260, *260*
serological tests, 260
stool antigen test, 260
epidemiology, 258
eradication therapy, 261
and gastric lymphoma, 483
pathogenesis, 258–9
results of infection, 259–60
gastric adenocarcinoma, 262, 264
gastric lymphoma, 262, 479, 483
gastritis/gastropathy, 263
peptic ulcer disease *see* Peptic ulcer disease
Helix-loop-helix motifs, *32*, **33**
Helix-turn-helix motifs, *32*, **33**
HELLP syndrome, 362–3
Helminthic infections, *163*, **163–71**
cestodes, 170–1
nematodes, 163–8
trematodes, 168–70
Helper T cells *see* T cell(s), helper
Hemaccel, 906
Hemi-bundles, 682
Hemianopia, *1101*, 1102, 1125
Hemiballismus, 1112, 1113, 1149
Hemicelluloses, 211
Hemidesmosomal proteins, 1227
Hemidesmosomes, 25, *25*, 26
Hemifacial spasm, 1108
Hemiparesis, *1096*, 1097, 1098, *1099*, 1112
infantile, 1169
Hemiparetic migraine, 1137
Hemisensory loss, *1128*
Henderson-Hasselbalch equation, 672
Hendra virus, *110*, 119
Henoch-Schönlein purpura, 435, *549*
acute abdomen, *315*
children, 554
clinical features, *550*
diagnosis, *550*
renal involvement, 592
Hepadna viruses, 109
Heparan, *58*
Heparan sulphate, 431
Heparin, 444
acute coronary syndromes, *754*
arterial thrombosis, *443*
low-molecular-weight, 444, 447
deep vein thrombosis, 810
pulmonary embolism, 786
Heparin sulphate, 687
Heparin-induced thrombocytopenia, 444–5
Hepatic artery, 320, *320*
Hepatic cysts, 642
Hepatic duct, 320
Hepatic fibropolycystic diseases *see* Fibropolycystic disease
Hepatic flexure syndrome, 370
Hepatic granuloma, 355
Hepatic mitochondrial injury syndromes *see* under Liver disease
Hepatic nuclear factor 4, 576
Hepatic stiffness, 326
Hepatic vein occlusion, *331*
Hepatitis, **332–3**
acute, 332, *332*, 333
cytomegalovirus, 342
herpes simplex, 342
infectious mononucleosis, 342
management *see individual causative agents*
pathology, 333
toxoplasmosis, 342
yellow fever, 342
see also viral below
alcoholic *see* Alcoholic hepatitis
autoimmune *see* Autoimmune hepatitis
chronic, 332, 333
causes, *333*
drug-induced, 344
pathology, 333
unknown cause, 344
see also individual viral hepatitis
fulminant, 339
hepatomegaly, *331*
Knodell Scoring System, 333
rheumatoid arthritis, *503*
viral, *83*, **333–42**, *334*
non-A-E, 342
see also individual hepatitis viruses
see also Fulminant hepatic failure (FHF)
Hepatitis A virus (HAV), *110*, 333–5, *334*
antibodies, 335
clinical features, 333
course and prognosis, 335
diagnosis, *69*
differential diagnosis, 335
epidemiology, 333
immunization, *102*, *103*
investigations, 333–5, *334*
prevention, 335
returning travellers, *90*
screening for, 173
treatment, 335
vaccination, 173, 335
virion, *334*
Hepatitis B virus (HBV), *104*, 182, *334*, **335–9**
acute
antibodies, *337*
antigens, *337*
clinical features, 336, *337*
antibodies, 627
arthritis in, 540–1
chronic, *337*, 338–9
active, 338–9
inactive, 338
transplantation, 347–8
clinical features, 336, *337*
course, 337, *337*
epidemiology, 335
HIV-associated, 201–2
immunization, *102*, *103*
investigations, 336–7, *337*
mutants, 336
pathogenesis, 336
prevention, 338
prognosis, 339
prophylaxis, 338
proteins, 336, *336*
screening for, 173
transfusion-transmitted, 428
transmission
horizontal, 335
vertical, 335
treatment, 337
see also Interferon(s) (IFNs); Lamivudine
virion, 335–6, *336*
Hepatitis C virus (HCV), *110*, *334*, 340–1
antibodies, 627
chronic, 340–1
clinical features, 340
diagnosis, 340–1
pathogenesis, 340
transplantation, 348
treatment, 341, *341*
clinical features, 340
course, 340
diagnosis, 340
epidemiology, 340
HIV-associated, 201–2
screening for, 173
transmission
treatment, 340, *341*
see also Interferon(s) (IFNs)
virion, 340, *340*
Hepatitis D virus (HAD), *334*, 339
chronic, 339
co-infection with HBV, 339
diagnosis, 339
superinfection, 339
Hepatitis E virus (HEV), *110*, *334*, 341–2
genome, *342*
Hepatocellular carcinoma, 336, 363–4
aetiology, 363
alcoholic cirrhosis, 359
clinical features, 363
geographical distribution, 451
haemochromatosis, 356
investigations, 363, *364*
prevention, 364
prognosis, 363–4
pyrexia, *91*
treatment, 363–4
Hepatocytes, 320
Hepatojugular reflex, 360
Hepatolenticular degeneration *see* Wilson's disease (hepatolenticular degeneration)
Hepatomegaly, 329, 331
AIDS, 362
alcoholic liver disease, 358
causes, *331*
haemochromatosis, 356
primary biliary cirrhosis, 355
sickle cell anaemia, 410
tricuspid stenosis, 768
Hepatopulmonary syndrome, 354
Hepatorenal syndrome, 354, 625
see also Cirrhosis; Renal failure
Hepatosplenomegaly
sarcoidosis, 869
schistosomiasis, 361
Hepatotoxicity, drugs, 365–7, *366*
Hepcidin, 395
Hepcidin gene, 356
Hepeviruses, *110*
Hephaestin, *394*, 395
HER2 gene, 453
Her2-neu gene, *453*
Herbicides, poisoning, *935*
Hereditary angioedema, 72
Hereditary chronic pancreatitis, 380
Hereditary coproporphyria, 1074, *1074*
Hereditary diffuse gastric cancer syndrome, *453*
Hereditary haemochromatosis, *789*
rheumatological problems, 555

Hereditary haemochromatosis (HH), 325, 356–7
  aetiology, 356
  clinical features, 356
  complications, 356
  investigations, 356–7
    genetic testing, 356
    heterozygotes, 356
    homozygotes, 356
    liver biopsy, 356
    magnetic resonance imaging, 357
  mechanism of damage, 356
  pathology, 356
  prevalence, 356
  screening, 357
  treatment, 357
Hereditary haemorrhagic telangiectasia *see* Osler-Weber-Rendu disease
Hereditary non-polyposis colorectal cancer, 50, 302, *453*
  diagnostic criteria, *302*
Hereditary telangiectasia, and gastric bleeding, *267*
Hermaphroditism, *1007*
Hernia
  diaphragm, 887
  hiatus, 250, *251*
Herpangina, 111, *111*
  rash, *88*
Herpes gestationis, 1270
Herpes simplex virus (HSV), *70*, *90*, 104–5, *104*, *105*, *191*
  conjunctivitis, 1088–9
  encephalitis, *105*
  genital *see* Genital herpes
  hepatitis, 342
  HIV-related, 202
  HSV-1, 104, 105, *105*
  HSV-2, 104, 105, *105*
  keratitis, 1090, *1090*
  rash, *88*
  skin involvement, 1230–1
Herpes zoster (shingles), *105*, 106–7, 1157, 1231, *1231*
  clinical features, 106, *106*, 1157
  geniculate ganglion, 1108
  incidence, *1096*
  postherpetic neuralgia, 1157
  prophylaxis and treatment, 106–7
  treatment, 1231
  vaccination, 1157
Herpes zoster virus, *70*, *90*
  genital ulceration, *172*
  rash, *88*
Herpesviruses, 104–8, *105*
  cytomegalovirus *see* Cytomegalovirus
  Epstein-Barr virus *see* Epstein-Barr virus
  herpes simplex virus *see* Herpes simplex virus
  human herpesvirus *see* Human herpesvirus
  varicella zoster virus *see* Varicella zoster virus
  see also specific viruses
Herpetic whitlow, 1231
Hers disease, *1069*
Heterochromatin, 33
Heterochromation, 51
Heterozygous familial hypercholesterolaemia, 1063–4, *1064*
Hexose monophosphate shunt, 221, 412, *413*
Heymann nephritis, 588
HFE gene, 356
Hiatus hernia, 250, *251*
  rolling/para-oesophageal, *251*
  sliding, *251*
Hib vaccine, 834
Hibernating myocardium, 740
Hic-1 gene, 51
Hiccups, 496, 887
Hidradenitis suppurativa, 1230
High altitude *see* Altitude
High dependency units (HDU), 889
  see also Intensive care unit (ICU)
High-density lipoprotein cholesterol, 211, 746, 1062
  reduced, in diabetes, 1034
High-density lipoprotein (HDL), 321, 1061, *1061*, 1062
  disorders of, 1064
Highly Active Antiretroviral Therapy *see* HAART
Highly transmissible infections, 91
Hilar lymph nodes
Hinge joints, 499
Hip pain, 512–13, *512*
  avascular necrosis (osteonecrosis) of femoral head, 513
  causes, *512*
  femoral neck fracture, 512
  gluteus medius tendonopathy, 512
  inflammatory arthritis, 513
  meralgia paresthetica, 512
  osteoarthritis, 512, 521
  polymyalgia rheumatica, 513
  rheumatoid arthritis, 527
  transient synovitis, 554
  trochanteric bursitis, 512
Hip protectors, 563
Hirschsprung's disease, 297
Hirsutism, 1003, 1267
  age and speed of onset, 1003
  causes, 1003
  iatrogenic, 1003
  idiopathic, 1003
  menstrual disturbances, 1003–4
  treatment, 1005
  and weight gain, 1004
Hirudin, 445
His bundle, 682
Histamine, *58*, 1098
  allergic response, *72*
Histamine provocation test, 852
Histidase deficiency, *1070*
Histidinaemia, *1070*
Histidine, 210
Histiocytoma, 1257
Histology
  breast cancer, *487*
  cancer, 455–6
  coeliac disease, 278, *279*
  Crohn's disease, *287*
  ulcerative colitis, *287*
Histones, 33
Histoplasma, *H. capsulatum*, *70*, *82*, 149, *191*
Histoplasmosis, 150–1, *150*
  African, 151
  clinical features, 150–1, *150*
  diagnosis, 151
  diarrhoea, *307*
  HIV-associated, 201
  management, 151
  returning travellers, *90*
Historical controlled trials, 931
History-taking
  medical interviews, 11
HIV, *110*, **184–205**
  antibodies, 627
  associated pathogens, *190*
  diagnosis, 185–7, *186*
    genome detection assays, 186
    IgG antibody to envelope components, 185–6
    IgG antibody to p24, 186
    viral p24 antigen, 186
    virus isolation in culture, 186
  epidemiology, 184
  group M, 185
  group N, 185
  group O, 185
  half-life, 185
  immunodeficiency, 189–90, *190*
  infection *see* HIV infection
  lymphoma, 483–4
    Castleman's disease, 484
    CNS, 484
    effusion, 484
  management, *see* Antiretroviral drugs (ARVs); HAART; specific drugs
  natural history, 185–7, *186*
  opportunistic infections, *190*
  pathogenesis, 185
  persistent generalized lymphadenopathy, 187
  routes of acquisition, 184
    contaminated blood products, 184
    contaminated needles, 184
    mother to child, 184
    sexual transmission, 184
  seroconversion, 186, *186*
  skin manifestations, 1268–9
    itchy folliculitis, 1269
  structure, *185*
  transfusion-transmitted, 428
  viral load, 190–1
HIV infection, 79, 80
  CDC classification, *187*
  clinical features, 187–8
    clinical latency, 187
    incubation, 187
    symptomatic infection, 188
  conditions associated with, 200–5
    bacterial infections, 203–4
    fungal infections, 200–1
    Kaposi's sarcoma, 204–5, *205*
    lymphoma, 205
    protozoal infections, 201
    squamous cell carcinoma, 205
    viral infections, 201–3
  diagnosis, *69*
  effects of, 188–9
    cardiac complications, 189
    endocrine complications, 189
    eye disease, 188
    gastrointestinal effects, 189
    haematological complications, 189
    mucocutaneous manifestations, 188–9, *188*
    neurological disease, 188
    renal complications, 189
    respiratory complications, 189
  focal segmental glomerulosclerosis, 587
  HLA system in, *60*
  incidence, 184
  initial assessment, 190, *190*
  management, 191–9, *191*
  monitoring, 190–1
    immunological, 190
    virological, 190–1
  mortality, *80*, 184
  musculoskeletal involvement, 541
  neurological disease, 1158
  opportunistic infections, 199
    chemoprophylaxis, 199
    immunization, *199*
  oral manifestations, 248
  polyneuropathy, 1176
  post-exposure prophylaxis, 199
  pregnancy, 198–9
  prevention and control, 205
  pyrexia, *91*
  seroconversion, 187
    acute, 198
  and tuberculosis, 865–6
  see also AIDS
HIV-associated focal glomerulosclerosis, 597
HIV-associated nephropathy, 189, 597
Hives *see* Urticaria
HLA system, **59–62**, *61*
  antigen presentation, 58, 62, *62*
  class I, 61, *61*
  class II, 61
  class III, 61
  in diabetes mellitus, 1033
  disease associations
    ankylosing spondylitis, *60*, 532–3
    rheumatoid arthritis, *60*, 523
    seronegative spondyloarthropathies, 532
  HLA type, 61
  in organ rejection, 75
  see also Immune system; individual markers; Major histocompatibility complex (MHC)
HLA-B27 gene
  ankylosing spondylitis, 532
  enteropathic arthritis, 536
  reactive arthritis, 535
HMB45, *493*
HMG-CoA reductase, inhibitors (statins) *see* Statins
HNF-1a, *1035*
HNF-1b, *1035*
HNF-4a, *1035*
Hnf4, 51
Hoarseness, 1082–3, 1110
  acute-onset, 1083, *1083*
  nodules, 1083, *1083*
  Reinke's oedema, 1083, *1083*
Hodgkin's lymphoma (HL), **476–9**
  aetiology, 476
  chemotherapy, *461*
  classification, 476, *476*
  clinical features, 476–7, *476*, *477*
  incidence, 476
  investigations, 477, *477*
  lymphocyte-depleted, 476
  lymphocyte-rich, 476
  management, 477–8, *478*
    advanced disease, 478–9, *478*
    early stage disease, 477–8
  minimal-change nephropathy, 586
  mixed cellularity, 476
  nodular sclerosing, 476
  pathology, 476
  pericarditis, 795
  prognostic factors, 477
  recurrent, 479
  Reed-Sternberg cells, 476, *476*
  staging, *477*
  survival, *478*
  see also Lymphomas
Holmes-Adie pupil, 1103–4
Holter monitoring, 697, *698*
  heart failure, 737
Homan's sign, 810
Home care
Homocysteine, 210, 224
  and coronary artery disease, 747
  metabolism, *225*
  pernicious anaemia, 400
  and venous thrombosis, 443
Homocystinuria, 1070
Homogentisic acid oxidase deficiency, *1070*
Homonymous field defects, *1099*
Homozygous familial hypercholesterolaemia *see* Familial hypercholesterolaemia
Honesty in clinical practice, 4
Honeycomb lung, *873*
Hookworm, 167
Hordeins, 277
Hormonal therapy, cancer, 463, *463*
Hormone replacement therapy, 965, *976*, 997
  osteoporosis, 563, 564
Hormones, 963, **965–8**
  actions of, 966–7
  basal levels, 968
  concentration measurement, 969, *969*
  control and feedback, 967, *967*
  deficiencies, 974
  and depression, 1200
  ectopic secretion, 1027

Hormones (*Continued*)
endocrine, 965
inactivation, 323
as neuromodulators, 965
as neurotransmitters, 965–6
paracrine, 965
plasma transport, 966
receptors, 966–7
release, 966
secretion, 967–8, *968*
storage, 966
stress-related, 968
synthesis, 966
as therapy, 965
see also Endocrinology; individual hormones; Steroids
Horner's syndrome, 811, 1103, *1104*, 1124
Hospital patients, **233–6**
Hospital-acquired disease, acute renal failure, 619
Hospitalized patients
Host defence, *54*
immunological defences, 86
infection, 86, *86*
natural defences, 86
see also Immune system
Host-organism interactions, 80
House-dust mite, 831–2, 848
Household products, poisoning with, 944
Housemaid's knee, 513
Howell-Jolly bodies, 278, 422, *423*
HTLV-1 *see* Human T cell lymphotropic virus (HTLV-1)
HtrA2 protein, 28
Human albumin solution, 429, 906
Human BK virus, 108
Human chorionic gonadotrophin (HCG), *493*, *976*
as tumour marker, *456*
Human embryonic stem cells, 47
Human genome, 33
Human herpesvirus type 6 (HHV6), *104*, *105*, 108
rash, *88*
treatment, 108
Human herpesvirus type 7 (HHV7), *104*, *105*, 108
Human herpesvirus type 8 (HHV8), *104*, *105*, 108
and non-Hodgkin's lymphoma, 479
transfusion-transmitted, 428
Human immunodeficiency virus (HIV) *see* HIV
Human leucocyte antigens *see* HLA system
Human metapneumovirus, *110*
Human papillomavirus, *104*, 108, **181–2**, *191*, 1231–2, *1231*
clinical features, 181–2
diagnosis, 182
HIV-related, 203
immunization, *103*
prevention/vaccination, 182
treatment, 182, 1231–2
urethral discharge, *172*
warts, 181–2
Human platelet alloantigen, 437
Human Proteome Project, 48
Human rights, 2
Human Rights Act (1988), 2
Human T cell lymphotropic virus (HTLV-1), *110*
leukaemia association, 467
and non-Hodgkin's lymphoma, 479
Humidifier fever, *876*, 877–8, 961
Humoral immune response
Hungry bone syndrome, 1021
Hunner's ulcer, 602
Hunter's disease, *789*
Huntingtin, 1149
Huntington's disease (HD), *43*, *227*, 1113, 1122, 1149
genetic anticipation, 45
management, 1149
neurotransmitters, *1113*
pathology, 1149
Hurler syndrome, *789*, 1071
Hutchinson's sign, 1266
Hyaline cartilage, 500
Hyaluronan, 499
Hybridization, 33–4
Hydatid cyst, *331*, *362*
Hydatid disease, *80*, 171, 361–2, *362*, *1179*
Hydralazine
genetic polymorphism in metabolism, *926*
heart failure, *739*
hypertension, 805
side-effects
rash, *1265*
respiratory, *877*
retroperitoneal fibrosis, 616
systemic lupus erythematosus, 542
vasculitis, 596
vasodilation, 909
Hydrocephalus, 1114, *1114*, 1135, 1162, *1162*
adult, 1162
infantile, 1162
normal pressure, 1162
Hydrochloric acid, 959
Hydrocortisone, *1237*
Addison's disease, *1012*
replacement therapy, *976*
seborrhoeic eczema, 1239
Hydrogen
and pH, *672*
renal excretion, 673–4, *673*
Hydrogen breath test, *277*
Hydrogen peroxide, 20
Hydrogen sulfide, poisoning, antidotes, *937*
Hydronephrosis, 614
Hydrophobia, 120
Hydrops fetalis, 109, 417
Hydrostatic pressure, 649, *650*
Hydrotherapy, 522
Hydroxocobalamin, 401, *937*, 943
Hydroxyapatite, 557
Hydroxycarbamide (hydroxyurea)
myelofibrosis, 421
polycythaemia vera, 420
sickle cell anaemia, 411
thrombocythaemia, 421
Hydroxychloroquine
chronic discoid lupus erythematosus, 1252
dermatomyositis, 1251
rheumatoid arthritis, 531
side-effects, inclusion bodies, 597
Sjögren's syndrome, 548
systemic lupus erythematosus, 544
Hydroxydaunorubicin, diffuse large B cell lymphoma, 482
Hydroxydoxorubicin, non-Hodgkin's lymphoma, *461*
Hydroxyethyl starches, 906
Hydroxyl radical, 20
21-Hydroxylase deficiency, 1014
Hydroxymethylglutaryl coenzyme A, 1062
17α-Hydroxyprogesterone, PCOS, 1004
Hydroxyproline, urinary, Paget's disease, 565
5-Hydroxytryptamine, 686
allergic response, *72*
5-Hydroxytryptamine (5-HT; serotonin), 1098
see also Selective serotonin reuptake inhibitors (SSRIs)
Hydroxyurea *see* Hydroxycarbamide (hydroxyurea)
25-Hydroxyvitamin D (25-hydroxycholecalciferol)
bone disorders, 559
reference ranges, 559
rickets/osteomalacia, 567
see also Vitamin D
Hydroxyzine
rhinitis, 833
sickle cell anaemia, *411*
Hygiene hypothesis, *72*, 847, 1236
Hymenolepsis nana, *163*, 170
Hyoscine butylbromide, 496
Hyper IgE syndrome, 71
Hyper IgM syndrome, 70
Hyperacusis, 1107
Hyperaldosteronism
glucorticoid/dexamethasone-suppressible, 1024
primary *see* Conn's syndrome
secondary, 1024
Hyperbaric oxygen, 890
Hypercalcaemia, *455*, 668, 1019–21
Addison's disease, 1011
bone malignancy, 555
cancer-related, 458
causes, *1019*
differential diagnosis, 1020
drug-induced, *964*
familial hypocalciuric, 560, 1021
imaging, 1020
investigations, 1020
of malignant disease, 1027
myeloma, 485
pathophysiology, 1019–20
pruritus, 629
renal calculi, 610
sarcoidosis, 869
symptoms and signs, 1020
treatment, 1020, *1021*
see also Hyperparathyroidism
Hypercalcaemic nephropathy, 599
Hypercalciuria, 677
causes, 610
idiopathic, 610, 613
renal calculi, 610
sarcoidosis, 869
Hypercapnia, 837, 893
Hypercarbia
permissive, 916
Hyperchloraemic acidosis, 675, 1020
Hyperchlorhydria, 263
Hypercholesterolaemia, 21, 1063–4
familial *see* Familial hypercholesterolaemia
heterozygous familial, 1063–4
homozygous familial, 1064
hypothyroidism, 986
polygenic, 1064
prevention, 1067
risk prevention tables, 1067–8
Hypercoagulable states, 585
Hyperemesis gravidarum, 362
Hypereosinophilic syndrome, *875*, 876
Hyperfibrinolysis, 441
Hyperforin, cytochrome P450 induction, *925*
Hyperfractionation, 465
Hyperglycaemia, 1029
see also Glucose, blood
Hyperhomocysteinaemia, chronic renal failure, 631
Hyperkalaemia, **666–8**
causes, 666, *667*
acute tubular necrosis, 622
Gordon's syndrome, 667
hyperkalaemic periodic paralysis, 667
hyporeninaemic hypoaldosteronism, 667
pseudo-hypoaldosteronism type 1, 667
suxamethonium, 667
clinical features, 667–8, *667*
diuretic-induced, 353
transfusion-related, 906
treatment, 623, 632, 668
Hyperkalaemic periodic paralysis, 667, 1182
Hyperlinear palms, 1237
Hyperlipidaemia, **1063–8**
classification, 1063
clinical features, 1063
combined, 1068
and coronary artery disease, 746
investigation, 1063
nephrotic syndrome, 584
remnant, 1064–5
screening, 1065, 1067
secondary, 1063, *1063*
skin manifestations, 1254
treatment, 1065, *1066*
see also Hypercholesterolaemia; Hypertriglyceridaemia
Hypermagnesaemia, 668, 669, *669*
Hypermetabolism, 898
Hypermetropia, 1086, *1086*
Hypermobility syndrome, children, 554
Hypernatraemia, 663–4
causes, 663, *664*
clinical features, 664
investigations, 664
treatment, 664
Hypernephroma *see* Renal cell carcinoma
Hyperosmolar hyperglycaemic state, *1046*, 1048–9, *1049*
Hyperoxaluria, 613
renal calculi, 610
Hyperparathyroidism, 963
chronic renal failure, 631
primary, 1019
pruritus, 629
renal calculi, 610
rheumatological problems, 556
secondary, 1019
surgery, 1021
tertiary, 627, 1020
treatment, 1020–1
see also Hypercalcaemia
Hyperparathyroidism-jaw tumour syndrome, 1019
Hyperphosphataemia, 671–2
causes, *671*
pruritus, 629
treatment, 632–3
Hyperpigmentation, *455*, 1264–5
Hyperplasia, prostate gland *see* Benign prostatic hypertrophy
Hyperplastic polyposis, *301*, 302
Hyperpnoea, 820
Hyperprolactinaemia, 972, 980–2, *981*, *1002*
acromegaly, 979
causes, 981
chronic renal failure, 629
clinical features, 981
idiopathic, *981*
investigations, 981–2
physiological, in pregnancy, 981
symptoms, 974
treatment, 982
Hyperpyrexia, malignant, 1182
Hyperreactive malarial splenomegaly, 155
Hypersensitivity reactions
type I (immediate/reaginic/anaphylactic), **72–3**, *73*
mediators involved, *72*
treatment, 72–3
see also Allergic disease (type I hypersensitivity)
type II (cytotoxic) *see* Autoimmunity (type II hypersensitivity)
type III (immune complex) *see* Immune complex disease (type III hypersensitivity)
type IV (cell mediated) *see* Delayed hypersensitivity
Hypersomnia, 1198

Hypersplenism, haemolytic anaemia, 415
Hypertension, **798–805**
  accelerated, 607
  acute heart failure, 742
  assessment, 800–2, *801*
  associated diseases/disorders
    renovascular disease, 607–8
    systemic sclerosis, 608–9
  autosomal-dominant polycystic kidney disease, 642
  causes, 798–9
  chronic, 799
  chronic renal failure, 630
  complications, 800, *800*
  and coronary artery disease, 746
  definitions, 798, *798*
  and diabetes, 1034, 1050
  endocrine causes, *1022*
  essential *see* Essential hypertension
  examination, 801
  history, 801
  humoral mechanisms, 799
  idiopathic intracranial, 1162
  inheritance, *44*
  insulin resistance, 799
  investigations, 802
  kidney and *see* Renal hypertension
  left bundle branch block, *720*
  malignant, 607, 800
    management, 805
  pathophysiology, 799–800
  poisoning, 936
  portal *see* Portal hypertension
  porto-pulmonary, 354
  pregnancy, 799
    management, 805
  primary *see* Essential hypertension
  prognosis, 805
  prohibited drugs, *923*
  pulmonary *see* Pulmonary hypertension
  renal *see* Renal hypertension
  renin/angiotensin dependent, 1023
  secondary, 799, 1022–3
    congenital cardiovascular causes, 799
    drugs, 799
    endocrine causes, 799
    pregnancy, 799
    renal diseases, 799
  treatment, 802–5, *802*
  see also Antihypertensive drugs
  see also Blood pressure
Hypertensive encephalopathy, 1130
Hypertensive nephrosclerosis, 607
Hypertensive scleroderma renal crisis, 547
Hyperthermia, overdose/poisoning, 936
Hyperthyroidism (thyrotoxicosis), 963, **987–91**
  autoimmune *see* Graves' disease
  causes, *987*
    amiodarone-induced, 988
    de Quervain's thyroiditis, 988
    solitary toxic adenoma/nodule, 987
    toxic multinodular goitre, 987
  clinical features, 988, *988*
    diabetes, *1031*
  differential diagnosis, 988
  drug-induced, *964*
  goitre in, 992
  investigations, 988–9
  long-term consequences, 991
  and malabsorption, 282
  neonatal, *965*
  in pregnancy, 990
  pyrexia, *91*
  relapse, 989
  thyroid crisis, 990
  thyroid function tests, *984*
  treatment, 989, *989*
  see also Antithyroid drugs
  see also Goitre; Graves' disease
Hypertrichosis, 1003, 1267
Hypertriglyceridaemia, 746, 1067
  and diabetes, 1034
  hypothyroidism, 986
Hypertrophic cardiomyopathy *see* Cardiomyopathy
Hypertrophic pulmonary osteoarthropathy, *455*, 555, 668, 882
Hyperuricaemia, **536–9**
  aetiology, 536–7, *537*
  clinical features, 537–8
  nephropathy, 599, 606
  pathogenesis, 537, *537*
  renal stones, 610
  vitamin B_{12} deficiency, 401
  see also Gout
Hyperuricosuria, renal stones, 610
Hyperventilation, 688, 817, 820, 1145
  central neurogenic, 1124
  diabetic ketoacidosis, 1047
Hyperventilation syndrome, *1208*
Hyperviscosity, 458
  chronic renal failure, 627
Hypervolaemia, *661*
Hyphaema, 1090, *1090*
Hypnotherapy, 1197
Hypnotics, side-effects, elderly patients, *924*
Hypoadrenalism
  drug-induced, *964*
  primary *see* Addison's disease
  secondary, 1012
  symptoms, 974
Hypoalbuminaemia, 321
  in Crohn's disease, 289
  extracellular volume regulation, 653
  hypercatabolic states, 321
  interstitial oedema, 655
  nephrotic syndrome, 584
Hypoaldosteronism, 1024
  drug-induced, *964*
  hyporeninaemic, 667
Hypobetalipoproteinaemia, *1064*, 1068
Hypocalcaemia, 1021–2
  causes, *1021*, 1022
  Chvostek's sign, 1021, 1022
  clinical features, 1022
    papilloedema, *1102*
  investigations, 1022
  long QT syndrome, *727*
  myopathy, 1179
  pathophysiology, 1021
  and seizures, 1145
  transfusion-related, 906
  treatment, 632, 1022
  Trousseau's sign, 1021, 1022
Hypocalciuria, 665
Hypocarbia, 893–5
Hypochlorhydria, 189
Hypochondriasis, 1195
Hypocitraturia, 677
Hypogammaglobulinaemia, 69, 429
  and malabsorption, 282
Hypoglossal nerve, 1110
  lesions, 1110
Hypoglycaemia, 1059–61
  alcohol-induced, 1060
  drug-induced, *964*, 1060
  during insulin therapy, 1043
  endocrine causes, 1060
  factitious, 1061
  hepatorenal, 1060
  insulinoma *see* Insulinoma
  mild, 1043
  nocturnal, 1043
  postprandial, 1060
  recurrent severe, 1059
  and seizures, 1145
  severe, 1043
  tumour-related, 1060
Hypoglycaemic unawareness, 1043
Hypogonadism
  female, 1001, *1001*
  haemochromatosis, 356
  male, 998–1000
    causes, *999*
    clinical features, 998–9
    investigations, 999
    treatment, 999, *999*
  normosmic idiopathic hypogonadotropic, 1000
Hypokalaemia, *665*, **665–6**
  causes, 665–6, *665*
  clinical features, 666
  Conn's syndrome, 1024
  Cushing's syndrome, 1012
  drug-induced, *964*
  folate deficiency, 401
  long QT syndrome, *727*
  metabolic alkalosis, 679
  myopathy, 1179
  treatment, 666, *666*
  vitamin B_{12} deficiency, 401
Hypokalaemic periodic paralysis, 666, 1178, 1182
Hypolipidaemia, 1068
Hypomagnesaemia, 665, 669
  causes, *669*
  clinical features, 669
  familial hypomagnesaemia, hypercalciuria and nephrocalcinosis, 669
  and hypoparathyroidism, 1022
  isolated dominant, 669
  long QT syndrome, *727*
  suppression of parathyroid hormone response, 559
  treatment, 669
Hypomania, 1204–6
Hyponatraemia, 661–3, *662*
  acute tubular necrosis, 622
  clinical features, 662, *662*
  dilutional, 654, 662
  diuretic-induced, 353, 657
  with euvolaemia, *661*, 662–3
    clinical features, 662
    investigations, 663
    treatment, 663
  with hypervolaemia, *661*, 663
  hypothyroidism, 986
  with hypovolaemia, 661–2, *661*
  treatment, 662
Hyponatraemic encephalopathy, 662
Hypoparathyroidism, 1021–2
  clinical features, 1022
  idiopathic, 1022
  investigations, 1022
  treatment, 1022
Hypopharynx, 1082
Hypophosphataemia, 670–1, 1020
  causes, 670–1, *670*
    decreased phosphate reabsorption, 670–1, *671*
    Dent's disease, 671
    vitamin D-dependent rickets type I, 670
    vitamin D-dependent rickets type II, 670
  treatment, 671
Hypophysitis, 974
Hypopigmentation, 1264
  post-inflammatory, 1264
Hypopituitarism, 974–6
  causes, 974, *975*
  clinical features, 974–5
  genetics, 974
  investigations, 975, *976*
  pathophysiology, 974
  selective, *965*
  treatment, 975, *976*
Hypoplastic left heart, *774*
Hyporeninaemic hypoaldosteronism, 667, 676
Hyposmia, 1082
Hypotension, 900
  diabetic ketoacidosis, 1047–8
  poisoning, 935–6
  postprandial, 689
  postural (orthostatic) *see* Postural hypotension
  ventricular tachycardia, 726
Hypothalamic-pituitary axis
  function tests, *976*
  hormones, *971*
  receptors, *971*
Hypothalamic-pituitary-adrenal axis, *1009*
Hypothalamic-pituitary-gonadal axis, *995*
Hypothalamic-pituitary-thyroid axis, *967*
  control of, 983
  see also Thyroid hormones
Hypothalamus
  anatomy, 969–70, *969*
  disorders of, **969–76**
    presentation, 970, 972
    see also individual conditions
  hormones, *971*
  physiology, 970, *970*
  role in appetite control, 229, *229*
Hypothermia, 954
  at home, 954
  clinical features, 954
  cold water immersion, 954
  diabetic ketoacidosis, 1048
  diagnosis, 954
  management, 955
  outdoors, 954
  overdose/poisoning, 936
  sequelae, 954–5
  severe, 955
  therapeutic, 712
Hypothesis testing, 933
Hypothyroidism, 963, 985–6
  and amenorrhoea/oligomenorrhoea, 1001
  atrophic (autoimmune), 985
  borderline, 986–7
  causes, 986, *986*
  children, 985
  clinical features, 985, *986*
  drug-induced, *964*
  investigations, 985–6
  and obesity, *228*
  pathophysiology, 985
  rheumatological problems, 556
  screening, 987
  serum prolactin, *981*
  skin manifestations, 1254
  symptoms, 974
  thyroid function tests, *984*
  treatment, 986
  see also Myxoedema
Hypotonia, 1114
Hypovolaemia, 652, *661*
  arterial pressure variation, 902
  colloid therapy, 650, *650*
  postural hypotension, *659*
  see also Extracellular fluid (ECF); Volume depletion
Hypovolaemic shock, *894*
  clinical features, 898
  haemodynamic changes, *899*
Hypoxaemia, 817, 893
Hypoxia
  altitude, 956, *956*
  and epilepsy, 1141
  nocturnal, 839–41
Hypoxic vasoconstrictor resonse, 918
Hypromellose eyedrops, 991
Hysteresis, 817
Hysteria *see* Dissociative (conversion) disorders

Ibandronate
  breast cancer, 489
  osteoporosis, 563, *564*
  Paget's disease, 566

Ibuprofen
musculoskeletal problems, 517, *518*
pericarditis, 795
sickle cell anaemia, *411*
Ibutilide, 730
ICAM(s), 67, *68*, *686*
shock response, 896
ICAM-1, 67, 814
systemic sclerosis, 545
Ice-cream headache, 1137
Ichthyosis, malignancy-related, *1255*
Ichthyosis vulgaris, 1236, 1237
Icodextrin, 636
Icterus *see* Jaundice
Id reaction, 1238
IDDM *see* Diabetes mellitus
Identification, 1189
Idiopathic fibrillary glomerulopathy, 592
Idiopathic guttate hypomelanosis, 1264
Idiopathic intracranial hypertension, 1162
Idiopathic pulmonary fibrosis, 872–3, *872*
antinuclear antibodies, *504*
clinical features, 872
differential diagnosis, 873
investigations, 872–3, *873*
pathogenesis, 872, *872*
pathology, 872
prognosis, 873
treatment, 873
Idiopathic pulmonary haemosiderosis, 819
Ifosfamide, 459
side-effects, nephrotoxicity, 462
Ileo-anal anastomosis, 292, *292*
Ileocaecal valve, 294
Ileostomies
Ileum, 270
resection, 280–1, *281*
Iliac crest biopsy, 567
Iliac fossa
left, pain in, 242
right, pain in, 242
Illness behaviour, 1191
Illness beliefs, 1187
Illusions, 1188
Iloprost, pulmonary hypertension, 783
Imaging
brain tumours, 1161
cancer, 456–7
cirrhosis, 346–7
Crohn's disease, 289
digital vascular, 327
hypercalcaemia, 1020
kidney, 579–82
liver, 325–8
lung(s), 821–4
muscles, 1179
pancreas, 376–7
retroperitoneal fibrosis, 617
stroke, 1131
ulcerative colitis, 291
see also specific techniques and conditions
Imatinib, *464*
chronic myeloid leukaemia, 473
Imidazoles, 100, *101*
cytochrome P450 inhibition, *925*
Imiglucerase, 1071
Imipramine, side-effects, hepatotoxicity, *366*
Imiquimod
Bowen's disease, 1258
human papillomavirus, 182
Immediate hypersensitivity, *see* Allergic disease (type I hypersensitivity)
Immediate spin crossmatching, 425
Immotile cilia syndrome, 818
Immune complex disease (type III hypersensitivity), 435
Immune complexes
see also Immune complex disease (type III hypersensitivity)
Immune deposits
Immune deviation, *73*
Immune reconstitution inflammatory syndrome (IRIS), 198, 204
Immune response, 53
adaptive, 53, *55*, **62–6**
innate, 53, **55–9**
modulation, 687
Immune system, **53–77**
adaptive immunity, 53, *55*, **62–6**
see also Antibodies; B cell(s);
autoimmune disease, 73–5
cell migration, 66–8
cells, *see* B cell(s); Lymphocytes; T cell(s)
cellular *see* Cell-mediated immunity
chemokines, 55
cytokines, 54
HLA system, **59–62**
immune response, 53
innate immunity, 53, **55–9**
see also Inflammatory response; Phagocytosis
laboratory investigations, 69, *69*
organ rejection, 75, *75*
Immune thrombocytopenic purpura, 436
in adults, 436
in children, 436
clinical features, 436
investigation, 436
treatment, 436
Immune-based therapies, 75–7, *76*
immunosuppressive drugs, 76–7
monoclonal antibodies *see* Monoclonal antibodies
Immunization, 102–3
active, *103*, 338
in heart failure, 738
passive, *102*, 338
recommended schedules, *103*
travellers, 103
see also Vaccines
Immunocompromised patients
infection, 90, *90*
see also AIDS; HIV infection; Immunodeficiency; Immunosuppression
Immunodeficiency, **69–72**
acquired, 69–70, *69*
see also AIDS
clinical features, 69–70
HIV, 189–90, *190*
primary, 69–70, *69*, 70–2
antigen presenting cell defects, 71
B cell deficiency, *69*, 70–1
complement deficiency, *69*, 71–2
neutrophil defects, *69*, 71
T cell deficiency, *69*, 70
see also Opportunist infection(s); specific conditions
Immunodiagnostic tests, 88–9
Immunofluorescence
Immunoglobulin(s), 63–4, *63*, 429
Fab region, 63
Fc region, 63
non-Hodgkin's lymphoma, 481
serum, 325
structure, *63*
surface, 63
therapeutic uses
immunosuppression, 76–7
myasthenia gravis, 1180
see also Antibodies; individual immunoglobulins
Immunoglobulin A (IgA), 63, *63*
antibodies, 64
dimeric secretory, 818
nephropathy, 593–4, *593*
selective deficiency, 71
small intestine, 275
see also Mucosa-associated lymphoid tissue (MALT)
Immunoglobulin D (IgD), 63, *63*
Immunoglobulin E (IgE), 63, *63*
and asthma, 847
binding to basophils, 431
measurement, *69*
Immunoglobulin G (IgG), 63, *63*, 66
antibodies, *73*, 185–6
autoimmune haemolytic anaemia, 415
thrombocytic purpura, 436
Immunoglobulin M (IgM), 63, *63*
focal segmental glomerulosclerosis, 587
monoclonal antibodies, 417
nephropathy, 590
non-complement-binding antibodies, 416
primary biliary cirrhosis, 355
Immunoglobulin superfamily, *25*, 26–7
Immunology
Immunopathologies, 53
Immunoproliferative small intestinal disease, and small intestine tumours, 284
Immunoprophylaxis, 102–3
Immunosuppression
cancer-associated, 455
chemotherapy-induced, 461
and opportunistic infection, 463
transfusion-related, 427
Immunosuppressive agents, 76–7
alkylating agents, 76
anticytokines, 76
ciclosporin, 76
cytokines, 76
glucocorticoids, 76
intravenous immunoglobulin, 76–7
multiple sclerosis, 1153
myasthenia gravis, 1180
purine analogues, 76
rapamycin, 76
restoration of tolerance, 76
specific conditions
lupus nephritis, 591
relapsing polychondritis, 557
systemic lupus erythematosus, 544
tacrolimus, 76
see also Cytotoxic drugs
Immunotactoid glomerulopathy, 592
Immunotherapy, 102–3
allergen, 72–3, *73*
cancer treatment, 465
rhinitis, 833
Impaired communication, communication difficulties for patients, 10, 11
Impaired fasting glucose (IFG), 1037
Impaired glucose tolerance (IGT), 1037
acromegaly, 979
see also Glucose tolerance test (GTT)
Impairment, 502
Impedance, oesophageal, 250
Impetigo, *86*, 1228, *1228*
bullous, 1228–9
prevention, 1228
treatment, 1228
Implantable cardioverter-defibrillator (ICD), 732–3, *732*, *733*
heart failure, 741
Implantable loop recorder, 697
Implied consent, 4
Imprinting, 44
In silico cloning, 36
In situ hybridization, 34
Inborn errors of metabolism
amino acid metabolism, 1069–71, *1070*
see also Aminoaciduria
carbohydrate metabolism, 1068–9, *1069*
Incidentalomas, 965
adrenal, 1014
pituitary, 972
Inclusion body myositis, 548
Incompetent patients
best interests, 7
consent to treatment, 6
Incontinence
faecal *see* Faecal incontinence
urinary *see* Urinary incontinence
Incretins, 232, *233*, 1039
Incus, 1077
Indanediones, 445
Indapamide, *657*
Indian childhood cirrhosis, 364
Indigestion, 241
see also Dyspepsia
Indinavir, *193*
side-effects
paronychia, 1269
renal calculi, 611
Indometacin
gout, 538
patent ductus arteriosus, 777
pericarditis, 795
Infantile acne, 1246
Infantile hemiparesis, 1169
Infantile hydrocephalus, 1162
Infants
prescribing in, 924
see also Infantile
Infection, 79–84
definition, *899*
diabetes mellitus, 1057
highly transmissible, 91
immunocompromised patients, 90, *90*
injecting drug users, 90
metabolic response, 898
metabolic/immunological consequences, 86–7
muscle involvement, *1179*
renal transplantation, 641
returning travellers, 89–90, *90*
sources of, 81, *81*
suspected, 87–91
clinical examination, 87
general investigations, 87–8
history, 87
microbiological investigations, 88–9, *89*
transfusion-transmitted, 428
treatment, 89–91
Infectious agents, 80, *81*, *82*
and cancer, 450
see also Bacteria; Pathogens; Protozoal infections; Viruses
Infectious diseases, **79–84**
bacterial infections, **122–40**
developed world, 79
developing world, 80
fungal infections, **149–53**
helminthic infections, **163–71**
host-organism interactions, 80
immunization, 102–3
mortality from, *80*
pathogenesis, 85–6, *85*, *86*
colonization, 85
epithelial attachment, 85
host response, 86
tissue dysfunction/damage, 85–6
prevention and control, 83–4, *83*
principles and mechanisms, 84–7
protozoal infections, **153–62**
routes of transmission, 81–3, *82*
airborne spread, 82
consumption of infected material, 83
direct inoculation, 82–3
endogenous infection, 81

faeco-oral spread, 82
person-to-person spread, 82
vector-borne disease, 82
sites of infection, *84*
sources of infection, 81, *81*
specificity, 84–5, *84*
transmissible spongiform encephalopathies, **121–2**
viruses, **104–21**
see also individual diseases and systems
Infectious mononucleosis (glandular fever), *105*, 1083
antinuclear antibodies, *504*
hepatitis, 342
rheumatoid arthritis, *503*
see also Epstein-Barr virus (EBV)
Infective arthritis *see* Arthritis; Septic arthritis
Infective endocarditis, 540, **769–73**
aetiology, 770, *770*
clinical features, 770–1, *770*, *771*
culture negative, 770
diagnostic criteria, 771–2, *771*, *772*
echocardiography, 704
with glomerulonephritis, 593
prevention, 773
pyrexia, *91*
rheumatoid factor, *503*
surgery, 773
treatment, 772–3
Inferiority, communication difficulties for patients, 10, 11
Infertility
treatment, 1005
see also Subfertility
Infestations, 1234–6
Inflammatory arthritis *see* Arthritis
Inflammatory bowel disease (IBD), **285–94**
aetiopathogenesis, 285–6, *286*
cancer in, 293
Crohn's disease *see* Crohn's disease
differential diagnosis, 288
enteropathic arthritis, 536
environmental factors, 286
epidemiology, 285
genetics, 286
lifestyle, 286
microscopic inflammatory colitis, 294
mortality, 293–4
pathology, 287–8, *287*, *288*
extragastrointestinal manifestations, 287, *288*
macroscopic changes, 287
microscopic changes, 287, *287*
in pregnancy, 293
pyrexia, *91*
ulcerative colitis *see* Ulcerative colitis
Inflammatory diarrhoea, 307
Inflammatory diseases
see also specific diseases
Inflammatory mediators, 68
infective endocarditis, 772
shock, 895–7
see also individual mediators
Inflammatory response, 87
acute, **68**
asthma, 850–1, *851*
chronic, **68**
Infliximab, *76*
Crohn's disease, *289*, 290
rheumatoid arthritis, *530*, 531
side-effects, respiratory, *877*
Influenza, *110*, 116–17, 834
clinical features, 834
complications, 834
diagnosis and treatment, 834
immunization, *103*
pandemics, 834
prophylaxis, 834
returning travellers, *90*
Information
communication difficulties for patients, 10, 11
conflicting, 10, 11
medical interviews, 12–13
following bad news, 14
Informed consent *see* Consent
Inhalers, *853*
Inhibin, *971*
Inhibitor of apoptosis proteins, *28*, 451
Innate immune response, 895, *896*
Innate immunity, 53, *55*, **55–9**, 818–19
complement, 55
components of, *56*
dendritic cells, 55, *56*, 58–9, *58*, *59*
eosinophils, 56–7, *56*
macrophages, 58
monocytes, 58
neutrophils, 55–6
see also Inflammatory response; Phagocytosis
Inoculation, direct, 82–3
Inositol 3-phosphate, 24
Inotropes *see* Inotropic agents
Inotropic agents, 907–8, *908*
guidelines for use, 908
heart failure, *739*, 740
Inotropic state *see* Myocardial contractility
Insect bites, 951, 1235
Insect stings, 951
Insecticides, 183, 1235
Insertions, 41
Insight, 1187
Insomnia, 1145, 1197
causes, *1198*
management, 1198
psychophysiological, 1198
Insulin, 229, *272*, 1040–1
analogues, 1040, *1040*, 1042
complications of therapy, 1043
hypoglycaemia, 1043
lipohypertrophy, 1042, 1043
weight gain, 1043
dosage, 1042, *1042*
hormonal regulation, *1030*, *1031*
hyperkalaemia, 668
infusion devices, 1042
inhaled, 1040–1
injections, 1042
longer-acting, 1040
overtreatment, 1059
profile, *1040*
regimes, *1041*
resistance, chronic renal failure, 629
secretion, 1029, *1030*
in type 2 diabetes, 1034–5
short-acting, 1040
side-effects, *928*
endocrine disease, *964*
signalling, *1031*
structure, 1029, *1040*
synthesis, *1030*
type 1 diabetes, 1041
type 2 diabetes, 1041, *1041*
see also Diabetes mellitus
Insulin aspart, 1040, *1041*
Insulin detemir, 1040, *1041*, 1042
Insulin glargine, 1040, *1041*, 1042
Insulin glulisine, 1040
Insulin lispro, 1040
Insulin receptor, 1030–1, *1031*
mutations, *1035*
Insulin resistance, 898, 1034, 1043
hypertension, 799
Insulin tolerance test, *976*
Insulin-dependent diabetes mellitus *see* Diabetes mellitus
Insulin-like growth factor binding protein 3, blood levels, 978
Insulin-like growth factor (IGF-1), *971*, 977
asthma, 852
blood levels, 978
acromegaly, 979
control, *977*
osteoarthritis, 520
Insulin/glucagon ratio, 214
Insulinoma, 384, 1059–60, *1059*
treatment, 1060
Whipple's triad, 1060
Integrase inhibitors, *193*, 194
Integrins, *25*, 26, 67, 1227
Intellectual impairment, *1099*
Intensive care, **889–921**, *890*
acute disturbances of haemodynamic function, *see* Shock
acute kidney injury, 624
acute lung injury, 917–19
acute lung injury/adult respiratory distress syndrome, 917–19
acute renal failure, 910
adult respiratory distress syndrome, 917–19
brain death, 920–1
cardiorespiratory physiology, 890–4
acid-base balance, 893–4
oxygen delivery and consumption, 890–3
discharge, 890
indications for admission, *889*
management, 890
respiratory failure, 910–17
management, 911–17
monitoring, 911
types and causes, 910–11
shock, **894–910**
clinical features, 898–9
management, 904–10
monitoring of critically ill patients, 899–904
pathophysiology, 894–8
withholding and withdrawing treatment, 919–20
Intensive care units (ICUs), 889
Intention tremor, 1114
Intention-to-treat analysis, 930
Intercellular adhesion molecules *see* ICAM(s)
Intercellular connections, 26
Intercostal drainage, 827–8, *828*
Interferon(s), 54, 102, 818
cancer therapy, 463–4
hepatitis B, 339
hepatitis C, 340
pegylated
hepatitis B, 339
hepatitis C, 341
renal cell carcinoma, 644
side-effects, 464
see also individual interferons
Interferon α-2a, pegylated, hepatitis D, 339
Interferon-α, 54, *101*, 102
cancer therapy, 464
thrombocythaemia, 421
Interferon-β, 54
multiple sclerosis, 1153
Interferon-γ, 54, 66
Interleukin(s), 54, 323
cancer therapy, 464
response to shock, 895
see also individual interleukins
Interleukin-1, 68
osteoarthritis, 520
regulation of metabolism, 214
tissue necrosis, 430
Interleukin-2, 65, 76, 214
cancer therapy, 464
receptor *see* CD25
renal cell carcinoma, 644
Interleukin-3, 57
haemopoiesis, 388
Interleukin-4, 58, *65*
Interleukin-5, 57
Interleukin-6, *65*, 68
in diabetes, 1034
haemopoiesis, 388
osteoarthritis, 520
regulation of metabolism, 214
release from Kupffer cells, 323
Interleukin-7, haemopoiesis, 388
Interleukin-8, 55, 56
osteoarthritis, 520
release from Kupffer cells, 323
Interleukin-10, *65*, 66, 214
systemic lupus erythematosus, 542
Interleukin-11, haemopoiesis, 388
Interleukin-12, *65*
haemopoiesis, 388
Interleukin-23, *65*
Interleukin-35, 54
Interlobar arteries, 571
Intermediate filaments, 25
Intermediate-density lipoprotein (IDL) particles, 321, 1061–2, *1061*
Intermittent mandatory ventilation (IMV), 915
Intermittent positive-pressure ventilation (IPPV), 912
see also Mechanical ventilation
Internal capsule, hemiparesis, 1112
Internal carotid artery, endarterectomy, 1133
Internal jugular vein cannulation, *901*
International Classification of Psychiatric Disorders (ICD-10), *1190*
International Hapmap Project, 36
International Normalized Ratio, 324, 445, 926
International Sensitivity Index, 445
Interphalangeal joints
distal (DIP)
proximal (PIP)
Interstitial cells of Cajal, 271
Interstitial fluid, 649
electrolyte composition, *650*
see also Extracellular fluid (ECF)
Intervertebral discs, 499
disease *see* Disc disease, prolapse/protrusion
Interview *see* Medical interview
Intestinal enzyme inhibitors, 1039
Intestinal ischaemia, 283
acute, 283
chronic, 283
Intestinal juices
Intestinal lymphangiectasia, 282, 283
Intestinal obstruction, *314*, 316–17
cancer-related, 496
causes, *316*
clinical features, 316
management, 316–17
Intestinal resection, 280–1
ileal resection, 280–1
jejunal resection, 281
short-bowel syndrome, 281
Intestine
large *see* Colon; Rectum
small *see* Small intestine
see also entries beginning Intestinal; Gastrointestinal tract
Intra-aortic balloon counterpulsation, 714–15, *715*, 909
acute heart failure, 714
unstable angina pectoris, 714–15
Intracellular fluid (ICF), 649
electrolyte composition, *650*
Intracellular signal inhibitors, 464, *464*

Intracerebral haematoma/haemorrhage, 1133–4
  aetiology, 1133–4
  management, 1134
  recognition, 1134
  see also Cerebral haemorrhage
Intracranial aneurysm, 643
Intracranial epidural abscess, 1159
Intracranial haemorrhage, 1133–5
  cerebral, 1134
    see also Intracerebral haematoma/haemorrhage
  subarachnoid *see* Subarachnoid haemorrhage
Intracranial mass
  see also Brain tumours
Intracranial pressure
  elevation, 1161
    intercerebral metastases, 458
Intracranial tumours *see* Brain tumours
Intraductal papillary mucinous tumour, 384
Intrahepatic cholestasis of pregnancy, 362
Intraocular pressure (IOP), 1086
  normal, 1091
  raised, *see* Glaucoma
Intrarenal volume receptors, 651
Intrauterine growth retardation, 977
Intravenous fluids, *660*
Intravenous immunoglobulin *see* Immunoglobulin(s)
Intravenous pyelography (IVP) *see* Excretion urography (intravenous pyelography: IVP)
Intravenous urography (IVU) *see* Excretion urography (intravenous pyelography: IVP)
Intrinsic factor, 399
Intrinsic pathway, 28
Introns, 33
Intussusception, 300
Inversion, 39
Involucrin, 1226
Iodine, 225, 984
  deficiency, 225, 984, 985
  dietary intake, 984
  radioactive, 989–90
  recommended IV intake, 233
  RNI, 225
  side-effects, endocrine disease, *964*
Iohexol, 575
Ion channelopathies, 792
Ion channels, 21–2
  Gados channel, 409
  see also specific channels
Ion-exchange resins, hyperkalaemia, 668
Ionizing radiation, 958–9
  dosage, 958
  late exposure effects, 958–9
  radiation dermatitis, 958
  radiation sickness
    acute, 958, *958*
    CNS syndrome, 958
    gastrointestinal syndrome, 958
    haemopoietic syndrome, 958
    mild acute, 958
  see also Radiation
Ipecacuanha, syrup of, 937
IPEX syndrome, 66, 74
IPF-1, *1035*
Ipratropium
  asthma, *854*, 855
  COPD, 839
Irbesartan, hypertension, 804, 1023
Irinotecan, 460
  colorectal carcinoma, 490
  genetic polymorphism in metabolism, *926*
  lung cancer, 487
  oesophageal cancer, 489
IRIS *see* Immune reconstitution inflammatory syndrome (IRIS)
Iris (of eye), 1086
  diabetic eye disease, 1052
Iritis, 188, 1092–3, *1092*
Iron, 225, 393–5
  absorption, 393–5, *394*
    factors affecting, *394*
  bone marrow, *396*
  deficiency, 393
    see also Iron deficiency anaemia; Microcytic anaemia
  dietary intake, 393
  erythroblasts, *396*
  haem, 393
    absorption, 393
  non-haem, 393
    absorption, 393–5
  oral, 396
  overload, 225, 356
    cardiomyopathy, 410
    thalassaemia, 408
    see also Haemochromatosis
  parenteral, 397
  poisoning, 944
    antidotes, *937*
    clinical features, 944
    metabolic acidosis, *939*
    treatment, 944
  recommended IV intake, 233
  requirements, 395
  RNI, 225
  serum, 395
    haemochromatosis, 356
    microcytic anaemia, 396, *396*
  storage, 395
  transport, 395
Iron deficiency anaemia, 225, 393, 395–7
  causes, 395
  clinical features, 395–6
  differential diagnosis, 396
  investigations, 396
  treatment, 396–7
Iron dextran, 397
Iron sorbitol, 397
Irritability, 1187
  physical causes, *1190*
Irritable bowel syndrome, 237, 311–13, *312*, 1194
  diagnostic criteria, 312, *312*
  as multisystem disorder, 312, *312*
  non-gastrointestinal features, *312*
  treatment, 312–13, *313*
    probiotics and prebiotics, 312–13
  trigger factors, *312*
Irritant dusts, 849
Ischaemia
  cardiac *see* Ischaemic heart disease (IHD)
  intestinal, 283
  leg *see* Lower limbs, ischaemia
  myocardial *see* Myocardial ischaemia
  transient, 711
  transient cerebral, 1144
  see also individual ischaemic conditions
Ischaemic colitis, 297–8
  treatment, 298
Ischaemic heart disease (IHD), 742, **743–60**
  coronary atherosclerosis, 744
  echocardiography, 704
  see also Angina pectoris; Atheroma; Coronary artery disease; Myocardial infarction
Ischaemic penumbra, 1129
Islet amyloid polypeptide *see* Amylin
Islet cells, 376
  autoantibodies, *1033*
  transplantation, 1043–4
Islets of Langerhans, 376, 384
Isocyanates, allergy, *848*
Isodesmosine, 500
Isoflurane
  side-effects, *928*
    hepatotoxicity, 365
Isoleucine, 210
Isoniazid, *93*
  genetic polymorphism in metabolism, *926*
  poisoning, *935*
  resistance, 866
  side-effects, 867, *928*
    hepatotoxicity, *366*
    neuropathy, *1174*
    pellagra, *222*
    rash, *1265*
    respiratory, *877*
    systemic lupus erythematosus, 542
  tuberculosis, 283, 866
  tuberculous meningitis, 1156
Isophane insulin, 1040
Isoprenaline, *908*
Isosorbide dinitrate
  heart failure, *739*
  vasodilation, 909
Isosorbide mononitrate, angina pectoris, 750
Isospora belli, *161*, *191*
  diarrhoea, 309
Isotretinoin, 247
  acne vulgaris, 1246–7
  rosacea, 1247
Ispaghula husk, 297
Itching *see* Pruritus
Itraconazole, 100, *100*
  allergic bronchopulmonary aspergillosis, 876
  blastomycosis, 152
  candidiasis, 183
  chromoblastomycosis, 153
  pityriasis versicolor, 1234
  ringworm, 1233
  sporotrichosis, 152
Ivabradine, angina pectoris, 750
Ivermectin, *166*
  filariasis, *163*
  loiaisis, *163*, 165
  onchocerciasis, *163*
  strongyloidiasis, 167

J-chain protein, 275
Jaccoud's arthropathy, 542
JAK *see* Janus kinase
JAK/STAT pathway, 977
Janeway lesions, infective endocarditis, *770*, *771*
Janus kinase, 388
  JAK2, mutations, 420
Japanese encephalitis, *81*, *110*, 115–16, 1156
Jarisch-Herxheimer reaction, 179, 1157
Jaundice, 328, 329–32
  acholuric, 329
  cholestatic *see* Cholestatic jaundice (acquired)
  haemolytic, 329
  pancreatic carcinoma, 384
  primary biliary cirrhosis, 355
  see also Congenital hyperbilirubinaemias (non-haemolytic); Haemolytic jaundice
Jaw thrust, 710
Jejunitis, ulcerative, 279
Jejunum, 270
  abnormal mucosa
  resection, 281
Jellyfish stings, 950
Jerk nystagmus, 1109
Jervell-Lange-Nielsen syndrome, *727*
  see also Torsades de pointes
Jet-lag, 960
Jewish populations
  cystic fibrosis, 844
  Gaucher's disease, 1071
  GM2 gangliosidoses, 1072
  inflammatory bowel disease, 286
  Niemann-Pick disease, 1071
Jiggers, 171–2
Job's syndrome, 71, *86*
Joint(s), **499–501**
  aspiration, 528
    see also Synovial fluid
  Charcot's, 540, 555–6, 1055, 1117
  deformities
  in diabetes mellitus, 1057
  examination, 502–3, *502*
  fibrocartilaginous, 499
  fibrous, 499
  infections, **539–41**
    actinomycetes infection, 540
    brucellosis, 540
    fungal, 541
    gonococcal arthritis, 540
    infective endocarditis, 540
    leprosy, 540
    Lyme arthritis, 540
    meningococcal arthritis, 540
    septic arthritis *see* Septic arthritis
    syphilitic arthritis, 540
    tuberculous arthritis, 540
    viral arthritis, 540–1
    see also Arthritis, infective
  neuropathic, 1165
  pain, *see* Arthralgia
  sensation, 500
  synovial, 499
  systemic lupus erythematosus, 542
  X-rays *see* X-ray
  see also specific joints
Joint aspiration, 504
  septic arthritis, 539
Joules, 208, 958
Jugular veins, mitral stenosis, 761
Jugular venous pressure, 691–2, *691*
  cor pulmonale, 837
  low, 659
  measurement, 691
  raised, 655, 691
Jugular venous pressure wave, 691–2, *691*
  a wave, 691
    cor pulmonale, 783
    tricuspid stenosis, 769
  c wave, 691
    tricuspid regurgitation, 768
  v wave, 691–2
    tricuspid regurgitation, 768
  x descent, 691
  y descent, 692
Jugular venous pulse, 783
Jugular venous thrombosis, *1110*
Jumper's knee, 513
Junctional adhesion molecules, 67
Junctional naevi, 1257
Juvenile idiopathic arthritis, 553–4
  antinuclear antibodies, *504*
  polyarticular, 553
  prognosis, 554
  rheumatoid arthritis, *503*
  systemic onset, 553
  treatment, 554
Juvenile nephronophthisis, 644
Juvenile polyposis, *301*, 302
Juxtaglomerular apparatus, 572, 895

K-*ras* oncogene, 382, *453*
Kala-azar *see* Leishmaniasis
Kallikrein-kinin system, 799
Kallmann's syndrome, 974, 1000
Kanamycin
  resistance, 866
  tuberculosis, 866

Kaposi's sarcoma, 204–5, *205*, 1260–1
  oesophagus, 257
  treatment, 1261
Kaposi's sarcoma virus, 75, *105*
Kaposi's sarcoma-associated herpesvirus *see* Human herpesvirus type 8 (HHV8; Kaposi's sarcoma virus)
Kaposi's varicelliform eruption, 1237
Karyotyping, *41*, 978
Kasabach-Merritt syndrome, 1268
Katayama fever, 169
Kawasaki's disease, *549*, 551
Kayser-Fleischer rings, 357
Kearns-Sayre syndrome, 39
Kell blood group, 424
Keloid scars, 1263–4, *1264*
Kennedy's syndrome, 1166
Kent bundles, 721
Keratan, 500
Keratan sulphate, 519
Keratin, *25*, 1226
Keratinocytes, 1226
Keratitis, *105*, 1090, *1090*
  contact lens-related, 1090, *1090*
  herpes simplex, 1090, *1090*
  marginal, *1090*
Keratoacanthoma, 1258, *1258*
Keratoconjunctivitis sicca
  primary biliary cirrhosis, 355
  sarcoidosis, 869
  see also Sjögren's syndrome
Keratoconus, 1086
Keratoderma blenorrhagica, 536
Keratolysis, pitted, 1230
Keratolytics, acne vulgaris, 1246
Keratomalacia, 218
Keratosis pilaris, 1237
Kernicterus, 417–18
  see also Jaundice
Kernig's sign, 1135
Keshan disease, 226
Ketoacidosis, 1046
  dehydration, *1046*
  diabetic *see* Diabetic ketoacidosis
Ketoconazole, 100, *100*
  blastomycosis, 152
  Cushing's syndrome, 1013
  mycetoma, 153
  pityriasis versicolor, 1234
  seborrhoeic eczema, 1239
  side-effects
    endocrine disease, *964*
    hepatotoxicity, *366*
Ketogenesis, *1047*
Ketonaemia, 1047
Ketone bodies, 214, 321, 1047
Ketones, 1036
Ketonuria, 1036, *1046*
Ketosis, *1046*
Kidd blood group, 424
Kidney
  autocrine function, 577–8
  blood flow *see* Renal blood flow
  blood supply, 571
  calculi *see* Urinary tract stones
  calyx, 571
  disease *see* Renal disease
  endocrine function, 576–8
  functional anatomy, 571–2, *572*
  functional impairment *see* Renal disease
  haemodynamics *see* Haemodynamics
  hypertension and *see* Renal hypertension
  interstitial inflammation *see* Tubulointerstitial nephritis (TIN)
  investigations, 578–82
    blood and quantitative tests, 579
    imaging, 579–82
    urine examination, 578–9
  obstruction *see* Urinary tract obstruction
  physiology, 572–4, *573*
  rheumatoid arthritis, 528
  systemic lupus erythematosus, 543
  transplantation donation, *see* Renal transplantation
  see also entries beginning nephro-; entries beginning renal; Renal tubules; specific diseases/disorders; Urinary tract
Kikuchi disease, 248
Killer cell immunoglobulin-like receptors *see* KIR
Killian's dehiscence, 1084
Killip classification, 758
Kimmelstiel-Wilson lesion, 589, 1054
Kinesin, 24
KIR, *66*
KIT proto-oncogene, 266
Kivexa, 195
Klebsiella, *123*
  K. aerogenes, urinary tract infection, *600*
  K. granulomatis, 180
  K. pneumoniae, *70*, *90*, 950
    pneumonia, 861
    ventilator-associated pneumonia, 914
  septicaemia, *136*
Kleihauer test, 418
Klinefelter's syndrome, 39, *40*, 542, 1000, *1007*
Knee(s), 513–15
  cruciate ligament tear, 514
  effusions, 514–15
  haemarthrosis, 514–15
  meniscus tear, 513–14, *514*
  osteochondritis dissecans, 514
  pain, *513*
    anterior, 513
    medial, 513
    osteoarthritis, *519*, 521
    rheumatoid arthritis, 526–7
Knee jerk reflex, *1115*
Knodell Scoring System, 333
Knowledge, lack, clinical communication difficulties, 10
Knudson multi-hit hypothesis, 51
Köbner phenomenon, 1232, 1241
Koilonychia, 1266
  iron deficiency anaemia, 392, 395
Konzo, 1165
Koplik's spots, 117
Korean haemorrhagic fever, 116
Kostmann's syndrome, *430*
Krebs cycle, 221
Kugelberg-Welander disease, 1167
Kulchitsky cells, 811, 881
Kupffer cells, 320, 322, 323
Kuru, 122, 1158
Kussmaul respiration, 667, 691, 1047, 1124
Kussmaul's sign, 793, 797
Kwashiorkor, 215–16, *215*
Kynureninase, 222
Kynurenine hydroxylase, 222
Kyphoscoliosis, 884
Kysanur Forest disease, haemorrhagic manifestations, *114*

L-myc, 49
Labetalol, hypertension, 805
Lactate, 1029
  dialysate, peritoneal dialysis, *637*
Lactate dehydrogenase (LDH)
  Hodgkin's lymphoma, 477
  myeloma, 485
  non-Hodgkin's lymphoma, 481
  pancreatitis, *378*
  polymyositis/dermatomyositis, 547
  thrombotic thrombocytopenic purpura, 437
Lactation, serum prolactin, *981*
Lactic acidosis, 677, *677*, 894, 907
  antiretroviral drugs, 198
  diabetes mellitus, *1046*, 1049
  see also Metabolic acidosis
Lactoferrin, *57*, 430, 818
Lactose intolerance, 237
Lactose tolerance test, 277
Lactulose, *297*
  sickle cell anaemia, *411*
Lacunar infarction, 1130
Lafora body disease, 1150
Lambert-Eaton myasthenio-myopathic syndrome, *455*, 1181
  pathophysiology, 1178
Lamina propria, 270, 275
Laminin, 1227
Lamivudine, *101*, *192*, *195*
  hepatitis B, 337, 339
  myelotoxicity, 189
Lamotrigine
  epilepsy, *1143*
  mania, 1206
  overdose, 940
  side-effects, *928*, *1143*
Langerhans' cell histiocytosis, 874
Langerhans' cells, 1226
Langhans' giant cells, 864
Lanreotide, 285
  acromegaly, 980
  insulinoma, 1060
  pancreatic neuroendocrine tumours, 384
  pituitary tumours, *973*
Lansoprazole, 252
  genetic polymorphism in metabolism, *926*
Lanthanum carbonate, 633
Lanugo, 1227
Lapatinib, breast cancer, 488, 489
Laplace's law, 734
Large intestine *see* Colon
Larva migrans
  cutaneous, *80*, *88*, 168
  visceral, 168
Laryngeal mask airway, 913
Laryngitis, 833
Laryngotracheitis, 117, 844
Laryngotracheobronchitis, 833
Larynx, 1082
  squamous cell carcinoma, 1083, *1083*
LASEK, 1087
Laser therapy
  bronchial carcinoma, 883
  eye surgery, 1087
LASIK, 1087
Lassa fever, *81*, 91, *110*, *114*, 120–1
  clinical features, 121
  diagnosis, 121
  treatment, 121
Latanoprost, glaucoma, 1091
Late phase response, 72
Latent autoimmune diabetes of adults (LADA), *see* Diabetes mellitus
Lateral medullary syndrome, 1129, *1129*
Latex allergy, *848*
Lathyrism, 1165
Laurence-Moon-Biedl syndrome, *228*
Lauric acid, *210*
Laws *see* Legal issues
Laxatives, 296–7, *297*
  bulking-forming, *297*
  osmotic, 296, *297*
  stimulant, *297*
Lead poisoning, 944–5, *1114*
  acute abdomen, *315*
  antidotes, *937*
  clinical features, 944–5
  renal tubular acidosis, *676*
  treatment, 945
  uraemia, 619
Leber's hereditary optic neuropathy, *37*, 39, 1102, *1102*
Lecithin, 321
Lecithin-cholesterol acyltransferase, 321, 1062
Lectins, 85
  antimicrobial, 275
Leflunomide, rheumatoid arthritis, *530*, 531
Left atrium (atrial)
  dilatation, 693
  pressure, 902, *902*
Left ventricle (ventricular)
  contrast echo for opacification, 704
  diastolic diameter, 702
  failure, 767
  function curve, *902*
  hypertrophy, 693, *764*, 799
    see also specific conditions
  non-compaction cardiomyopathy, *789*, 791–2
  outflow, 699
  systolic diameter, 702
  systolic dysfunction, 736
Leg(s)
  ischaemia, 805–7
  ulceration, 1261–2, *1261*
    arterial ulcers, 1261
    neuropathic ulcers, 1261–2
    sickle cell disease, 410
    venous ulcers, 1261
    see also Lower limbs
Legal issues, 1, 3–6
  complaints, 14–15
  confidentiality, 4–5
  consent *see* Consent
  discretion, 5
  duty of care – statutary/common law, 2
  ethical/legal boundaries, 6–8
  European Convention on Human Rights, 2
  evidence of negligence, 3
  Human Rights Act (1998), 2
  the law, 2
  lawsuits, 15
  medical mistakes, 3
  Mental Health Act (1983), 6
  poor practice, 3
  professional regulation, 2
    see also General Medical Council (GMC)
  psychiatry, 6
  rational self-interest, 2
  report writing, 15
  standards, 2
  vulnerable patients, 5–6
  see also Ethics
Legionella, *123*
  infective endocarditis, 770
  L. pneumophila, *70*, *82*, 128
    pneumonia, *858*, 860
Legionnaires' disease, *82*, 128, 961
Leigh's syndrome, *37*
  maternally inherited, *37*
Leishman-Donovan bodies, 159
Leishmania, *82*, 158
  L. aethiopica, *159*
  L. amazonensis, *159*
  L. braziliensis, *159*
  L. chagasi, *159*
  L. donovani, *159*, *191*
  L. garnhami, *159*
  L. guyanensis, *159*
  L. infantum, *159*
  L. major, *159*
  L. mexicana, *159*
  L. panamanensis, *159*
  L. peruviana, *159*
  L. pifanoi, *159*
  L. tropica, *159*
  L. venezuelensis, *159*

Leishmaniasis, *80*, *82*, 158–60
  cutaneous, 159–60, *159*
  HIV-related, 201
  mortality, *80*
  mucocutaneous, *159*, 160
  post-kala azar dermal, 159
  splenomegaly, 423
  visceral (kala azar), 159, *159*
Lenalidomide
  myelodysplasia, 422
  myeloma, 486
Lenegre's syndrome, 719, *719*, 792
Lens, 1086
  cataracts *see* Cataracts
Lentiginosis, *789*
Lentigo maligna, 1258–9
  melanoma, 1259
Lentigos, 1264
Lentiviruses ('slow' viruses), *120*
  see also HIV-1; HIV-2
LEOPARD syndrome, 1265
Lepidoglyphus
  L. destructor, 877
  L. domesticus, 877
Lepirudin
  arterial thrombosis, *443*
  venous thrombosis, 445
Leprechaunism, *1035*
Lepromatous leprosy, 140
Leprosy (Hansen's disease), *83*, 140–1, *140*, *141*
  borderline, 141, *141*
  Charcot's joints, 555
  clinical features, 141–2
  diagnosis, 141
  hypopigmentation, 1264
  indeterminate, 1230
  joint involvement, 540
  lepra reactions, 141
  lepromatous, 140, 141, 1230
    borderline, 141
  management, 141, *141*
  multibacillary, 140
  paucibacillary, 140
  prevention, 141
  rheumatoid factor, *503*
  skin involvement, 1230
  tuberculoid, 140, 1230
    borderline, 140–1
Leptin, 228, 229, 367, 376
  resistance, 230
  in starvation, 230
Leptospira, *123*
  L. interrogans, *80*, 137
  L. interrogans canicola, 137
  L. interrogans hardjo, 137
  L. interrogans icterohaemorrhagiae, 137, 138
  L. interrogans pomona, 137
Leptospirosis, *80*, *83*, 137–8
  clinical features, 138
  diagnosis, 138
  management, 138
Leriche syndrome, 805
Letrozole, 463
Leucine, 210
Leucine zipper, *32*, **33**
Leucocystoclastic vasculitis, 1263
Leucocyte adhesion deficiency, 71
Leucocyte alkaline phosphatase, 430
  myelofibrosis, 421, 430
  polycythaemia, 430
Leucocyte common antigen, *493*
Leucocyte count, pancreatitis, *378*
Leucocyte function antigen, *68*
Leucocyte function associated antigen-1, 67
Leucocytes, **430–1**
  antibodies, 427
  transfusion incompatibility, 427
  in urine, 579
  see also individual cells; Lymphocytes
Leucocytosis, septic arthritis, 539
Leucoerythroblastic anaemia, 430
Leuconychia, 1266
Leucopenia
  alcohol-induced, 325
  drug-induced, *86*
Leucoplakia, *247*
Leukaemia, **467–75**
  acute, 468–70, *468*
    active therapy, 469–70
    classification, *468*
    clinical features, 468
    curative therapy, 469
    investigations, 468–9, *469*
    management, 469
    palliative therapy, 469
    specific treatment, 470
    see also various types below
  acute lymphoblastic (ALL), 467, *468*, 471–2
    cytogenetic aberrations, 471
    immunophenotype, 471
    minimal residual disease, 471
    prognosis, 472
    prognostic factors, 471
    survival, *472*
    time to response, 471
  acute myeloid (AML), 467, *468*, 470–1, *470*, *471*
    with multilineage dysplasia, *468*
    and myelodysplastic syndromes, therapy-related, *468*
    with recurrent genetic abnormalities, *468*
    survival, *471*
  acute promyelocytic (APML), 468, 471
  aetiology, 467
  arthritis, 556
  cell surface markers, 468
  chronic lymphocytic (CLL), 467, 473–5
    clinical features, 473
    investigations, 473–4
    lymphomatous transformation, 475
    management, 474–5
    prognostic factors, 474
    staging systems, *474*
  chronic myeloid (CML), 49, 467, 472–3
    clinical features, 472–3
    investigations, 473
    management, 473
    Philadelphia chromosome, 472, *472*
  classification, 467
  epidemiology, *449*
  genetic abnormalities, 467–8, *467*
  gingival swelling, 248
  hairy cell, 475
  hepatomegaly, *331*
  prevalence, 467
  prolymphocytic, 475
  pyrexia, *91*
  remission, 470
Leukaemoid reaction, 430
Leukotriene(s), 56
  B4, *58*, *68*
    allergic response, *72*
  C4, 57, *58*
    allergic response, *72*
  D4, 57, *58*
    allergic response, *72*
  E4, allergic response, *72*
Leukotriene antagonists
  asthma, 856
  rhinitis, 833
Leuprorelin, 463
  prostate cancer, 492
Levamisole, *166*
  minimal-change nephropathy, 586
Levetiracetam
  epilepsy, *1143*
  overdose, 940
Levobutanolol, glaucoma, 1091
Levodopa, 1113, 1147
  side-effects, 1147
Levofloxacin, 98
Levomepromazine, nausea, 461
Levothyroxine, *976*, 986, 989
Lev's disease, 719, *719*
Lewy bodies, 1146
Lewy body dementia, 1168
Lhermitte-Duclos disorder, *301*
Lhermitte's sign, 1116–17
  multiple sclerosis, 1152
Li-Fraumeni syndrome, *453*
  leukaemia risk, 467
Libido, *995*
  lack of, 1000
Libman-Sacks syndrome, 543
Lice
  body lice, 1235
  head lice, 1235
Lichen planus, 1244–5, *1244*
  clinical features, 1244
  HIV-related, 1269
  oral, 247
  pathology, 1244
  prognosis, 1244
  treatment, 1245
Lichen sclerosus, 1245
Lichen simplex, 1240, *1240*
Lichenification, *1227*
Lichenoid rash, *1265*
Liddle's syndrome, *659*, 666
Lidocaine, 730, *731*
  genetic polymorphism in metabolism, *926*
  ventricular tachycardia, 726
Life-saving intervention, legally valid consent, 3–4
Life-support, medical ethics, 1
Lifestyle, 8
  avoidance of obesity, 233
  heart failure, 737–8
  inflammatory bowel disease, 286
  Marfan's syndrome, 781
  patient's rights, 8
  post-myocardial infarction, 759
Ligaments, 500
  see also Joint(s)
Ligase, 35
Light chain cast nephropathy, 599
Light chain deposition disease, 599
Lightning pains, 1116
Lignin, 211
Limb apraxia, *1099*
Limbic, hypothalamic, brain stem encephalitis, *455*
Limbs
  examination, *502*
  see also Leg(s); Lower limbs; Upper limbs
Lindane, 183, 1235
Linear IgA disease, 1256–7
Linezolid, 99, 124
Linitis plastica, 265
Linkage disequilibrium, 36
Linoleic acid, 210, *210*
Linolenic acid, 210, *210*
Lipaemia retinalis, 1063
Lipase, 375
  raised levels, pancreatitis, 378
  supplements, 382
Lipid(s)
  abnormalities
  breakdown *see* Lipolysis
  measurement, 1062
  metabolism, 321
    chronic renal failure, 629
    disorders of, **1061–8**
    see also individual conditions
  physiology, 1061–2, *1061*, *1062*
    see also Fat; Lipoproteins; Phospholipids
Lipid-lowering diet, 1065
Lipid-lowering therapy
  angina pectoris, 750
  hyperlipidaemia, 1065, *1066*
Lipodermatosclerosis, 655
Lipodystrophy
  antiretroviral drugs, 198
  congenital, diabetes in, *1031*
  partial, *1035*
Lipohypertrophy, 1042
Lipolysis, 213
Lipoma, lung, 884
Lipopolysaccharide, *59*, 85–6, 87, 122, 323, 895
Lipopolysaccharide binding protein, 895
Lipoprotein (a), and coronary artery disease, 747
Lipoprotein lipase deficiency, 1063, *1064*
Lipoprotein receptor-related protein 5, 484
Lipoproteins, 321, 1061
  HDL *see* High-density lipoprotein (HDL)
  LDL *see* Low-density lipoprotein (LDL)
  VLDL *see* Very low-density lipoprotein (VLDL)
Liposuction, 233
Lips
Liquorice
  side-effects
    endocrine disease, *964*
    hypertension, 799
    sodium retention, 656
Liraglutide, 1039
Lisch nodules, 1253
Lisinopril
  acute coronary syndromes, *754*
  hyperaldosteronism, 1024
  hypertension, 804
Listening skills, medical interview, 11
Listeria, 287
  L. monocytogenes, *70*, 85, *123*, 138
    bacterial meningitis, 134
    food poisoning, *133*
Listeriosis, 138–9
Lithium
  mania, 1205
  poisoning, 945
  side-effects
    endocrine disease, *964*
    rash, *1265*
    renal tubular acidosis, *676*
  teratogenicity, *924*
  therapeutic drug monitoring, *927*
  toxicity, 1205–6
Lithium dilution, 904
Lithocholic acid, 322
Little's area, 1080, *1081*
Livedo reticularis, 543
Liver, *319*, **319–67**
  acinus, 320, *320*
  alcohol effect on, *see* Alcohol; Alcoholic cirrhosis
  arterial supply, 320, *320*
  biochemistry, 324, 332
    autoimmune hepatitis, 344
    cirrhosis, 346
    hepatitis A, 333–4
    intrahepatic cholestasis of pregnancy, 362
    musculoskeletal problems, 503
  biopsy, 328, *328*, 341
    autoimmune hepatitis, 344
    cirrhosis, 347
    haemochromatosis, 356
    non-alcoholic fatty liver disease, 345
    primary biliary cirrhosis, 355
    tuberculosis, 865
    Wilson's disease, 357
  caudate lobe, 320
  disease *see* Liver disease
  drugs and, 365–7, *366*
    drug metabolism, 365
    hepatotoxicity, 365–7, *366*
  enlargement *see* Hepatomegaly
  failure, fulminant, 342–3
  fatty, *331*, 344–5, 358

acute fatty liver of pregnancy, 363
investigations, 359
management, 359
functions of, 320–3
bile formation, 321–2, *322*
bilirubin metabolism, 322–3
carbohydrate metabolism, 321
glucose metabolism, 1029
hormone and drug inactivation, 323
immunological, 323
lipid metabolism, 321
protein metabolism, 320–1, *321*
investigations, 323–8
blood tests, 324–5, *324*
imaging techniques, 325–8
see also individual techniques
see also under individual diseases
reticuloendothelial system, 323
sinusoids, 320
structure, 319–20, *320*
transplantation *see* Liver transplantation
tumours *see* Liver tumours
venous drainage, 320, *320*
see also entries beginning hepato-
Liver abscess, 360–1
amoebic, 361
pyogenic, 360–1
Liver disease
acute
signs, 328
symptoms, 328
transplantation, 347
AIDS-related, 362
alcoholic *see* Alcohol; Alcoholic cirrhosis; Alcoholic liver disease
chronic
signs, 328–9, *328*
skin manifestations, 328–9, 1254
symptoms, 328
transplantation, 347
cirrhosis *see* Cirrhosis
coagulation disorders in, 440
coeliac disease, 365
cystic fibrosis, 365
drug prescribing, 366
fibrosis, 325, 345, *345*, 356
congenital, 360
markers of, 325
perivenular, 358
glycogenoses, *1069*
hepatic mitochondrial injury syndromes, 364
and hypoglycaemia, 1060
idiopathic adult ductopenia, 364
parenchymal *see* Hepatitis
polycystic, 360
pregnancy, 362–3
rheumatological problems, 554–5
signs, 328–9
solitary cysts, 360
symptoms, 328
veno-occlusive, 360
see also individual diseases; Liver transplantation
Liver enzymes, infection, *88*
Liver flukes, 170
Liver function tests, 323
Liver palms, 328
Liver transplantation, 347–8
contraindications, 348
indications, 347–8
primary biliary cirrhosis, 356
prognosis, 348
rejection, 348
surgical procedure, 348
Liver tumours, 363–4, *364*
benign, 364
cholangiocarcinoma, 364
hepatocellular carcinoma *see* Hepatocellular carcinoma
primary, 363–4
secondary, 363
Lixivaptan, *657*, 663
Loa loa, *163*, *164*, 165
Locked-in syndrome, *1123*, *1125*, 1126, 1129
Lod score, 36
Loeffler's endocarditis, *789*, 793
Loiasis, 165
treatment, *163*
Loin pain, 614
Lomustine, 460
side-effects, respiratory, *877*
Long QT syndrome, 727–8, *789*, *1078*
acquired, 728
causes, *727*
clinical features, 728
congenital, 727–8, *728*
ECG, *728*
management, 728, *730*
see also Torsades de pointes
Long-chain 3-hydroxylacyl-CoA-dihydroxyl (LCHAD) deficiency, 363
Loop diuretics, *657*
clinical uses, 657, *658*
side-effects, 657
Loperamide, 297, 307, 313
Lopinavir, *193*, *195*, 196
Loratadine, rhinitis, 833
Lorazepam
behavioural disturbance, 1191
convulsions, 936
Loricrin, 1226
Losartan
genetic polymorphism in metabolism, *926*
gout, 538
heart failure, *739*
hypertension, 804, 1023
Lotions, *1270*
Louse, *82*
Louse-borne relapsing fever, *82*, 146–7
Low back pain, 509–12, 1177
ankylosing spondylitis, 512
causes, *509*
children, 554
diffuse idiopathic skeletal hyperostosis, 512
facet joint syndrome, 510
investigations, 509–10
lumbar disk prolapse, 510–11
lumbar nerve root entrapment, *511*
lumbar spondylosis, 510
management, *509*
mechanical, 510
treatment, 510
osteoporotic crush fracture of spine, 512
postural, 510
septic discitis, 512
spinal and root canal stenosis, 511
spondylolisthesis, 511–12
sway back of pregnancy, 510
treatment, 511
Low birthweight
and heart disease, 1034
and hypertension, 1034
Low-density lipoprotein (LDL), 21, *21*, 224, 1061, *1061*, 1062
disorders of, 1063–4
nephrotic syndrome, 584
and total cholesterol, 1062
see also Cholesterol
Low-molecular-weight heparin, 444
venous thromboembolism, 447
Lower limbs
examination, *502*
ischaemia, 805–7
acute, 806–7
chronic, 805–6
see also Leg(s)
Lower motor neurone (LMN) lesions, 1107, 1115
causes, 11115
loss of bladder control, 1119
signs of, 1115, *1115*
Lower reference nutrient intake (LRNI), 207
Lower respiratory tract infection, mortality, *80*
Lowe's syndrome, 1070
LRNI *see* Lower reference nutrient intake
Lubiprostone, *297*
sickle cell anaemia, *411*
Lumbar discs, prolapse/protrusion, 1176–7
Lumbar puncture, 1121, *1122*
brain tumours, 1161
Lumbar spine
back pain *see* Low back pain
Lumbosacral discs, protrusion, 1177
Lumiracoxib, gout, 538
Lundh meal, 376
Lung(s), 812, *813*
airways *see* Airways
collapse, 822, *822*
causes, *822*
consolidation, *821*, 822, *822*
disease *see* Lung disease
fibrosis, *821*, 822
imaging, 821–4
lobes, 812
lymphatics, 813–14
miliary mottling, 822–3
nerve supply, 814
parenchymal disease *see* Lung disease
rheumatoid arthritis, 527
shadows, 822, *822*
systemic lupus erythematosus, 543
vasculature, 813–14
ventilation, 814
see also Breathing
see also Pulmonary
Lung abscess, *86*, 863
Lung cancer/tumours, **486–7**
chemotherapy, *461*
epidemiology, *449*
geographical distribution, 451
metastatic, 883–4
non-small-cell, 486
prognostic factors, 486
screening, 884
small-cell, 486–7
see also Bronchial carcinoma
Lung compliance, 814
Lung contusion, 884
Lung cysts, 880
Lung disease
acute injury (ALI), 917–19
see also Adult respiratory distress syndrome (ARDS)
autoimmune rheumatic disease, 871–2
calcification
cancer *see* Lung cancer/tumours
diffuse parenchymal, *868*, **868–78**
granulomatous, 868–70
with vasculitis, 870–1
infection *see* Respiratory tract infections; specific infections
occupational, 878–80
respiratory bronchiolitis interstitial, 873
tuberculose *see* Tuberculosis
tumours *see* Lung cancer/tumours
see also individual diseases; Respiratory diseases; Respiratory tract infections
Lung function tests, 824–6
asthma, 852
blood gases, 826
COPD, 838
extrinsic allergic alveolitis, 877
idiopathic pulmonary fibrosis, 872
nitric oxide exhalation test, 826
sarcoidosis, 870
ventilatory function, 824–6, *825*, *826*
see also individual tests
Lung rupture, and diving, 957
Lung transplantation, 846
single lung, 941
Lung volume, *824*, 825, *826*
Lupus anticoagulant antibodies, 441, 1122
Lupus anticoagulant test, 544
Lupus erythematosus (LE), 1252–3
chronic discoid, 1252, *1252*
discoid, 543
drug-induced, *1265*
antinuclear antibodies, *504*
seronegative, 590
subacute, 1252
systemic *see* Systemic lupus erythematosus (SLE)
Lupus glomerulonephritis, 590–1, *590*
classification, *590*
management, 591
pathology, 591
prognosis, 591
Lupus pernio, 869, 1253
Lupus vulgaris, 1230
Lutein, 218
Luteinizing hormone (LH), *971*, 996
chronic renal failure, 630
secretion, 967
see also Luteinizing hormone releasing hormone (LHRH)
Luteinizing hormone releasing hormone (LHRH)
see also Gonadotrophin releasing hormone (GnRH)
Lutembacher's syndrome, 760
Lutzomyia, 159
Lycopene, 217, 218, 224
Lyme arthritis, 522, 540
Lyme disease, *80*, *82*, 139
clinical features, 139
diagnosis, 139
management, 139
pyrexia, *91*
rash, *88*
skin changes, 1252
Lyme pericarditis, 795
Lymph, 320
Lymph node(s), 53, 387
biopsy
Hodgkin's lymphoma, 477
non-Hodgkin's lymphoma, 481
cervical, *476*
enlargement
cervical *see* Cervical lymph node enlargement
see also Lymphadenopathy
movement of immune cells, *67*
rheumatoid arthritis, 528
sentinel, 457
tracheobronchial, 814
tuberculosis, 865
Lymphadenopathy
bilateral hilar, 869, *869*
peripheral, 480
persistent generalized, 187
Lymphangioleiomyomatosis, 874
Lymphangioma circumscriptum, 1263
Lymphangitis, 68
Lymphatic system
see also Lymph node(s)
Lymphocytes, 431
asthma, 850–1
atypical, *88*
passenger, 417
see also B cell(s); T cell(s)
Lymphocytic choriomeningitis, *110*, 121

Lymphocytic colitis, microscopic, 294
Lymphocytosis, *88*, 431
Lymphoedema, 1263
Lymphogranuloma venereum, 177
  clinical features, 177
  diagnosis, 177
  genital ulceration, *172*
  treatment, 177
Lymphoid cells, *56*
Lymphoid interstitial pneumonia, 873–4
Lymphoid interstitial pneumonitis, 189
Lymphoid organs
  primary, 53, *54*
  secondary, 53, *54*, 55
  see also Bone marrow
Lymphomas, **475–84**
  high-grade (non-Hodgkin's) *see* Non-Hodgkin's lymphoma (NHL)
  HIV-related, 205
  Hodgkin's *see* Hodgkin's lymphoma (HL)
  low-grade (non-Hodgkin's) *see* Non-Hodgkin's lymphoma (NHL)
  and malabsorption, 282
  non-Hodgkin's *see* Non-Hodgkin's lymphoma (NHL)
  primary effusion, 108
  pyrexia, *91*
  splenomegaly, 423
Lymphopenia, *88*
  HIV, 189
  musculoskeletal problems, 503
Lymphoplasmacytic lymphoma, *480*, 482
Lymphoproliferative disorders, post-transplantation, 641
Lynch syndrome, 50, 302
Lyonization, 37
Lysergic acid, side-effects, retroperitoneal fibrosis, 616
Lysine, 210
Lysosomal enzymes, 895–6
Lysosomal hydrolases, 323
Lysosomal storage diseases, 1071–2
Lysosomal trafficking regulator gene, 71
Lysosomes, *19*, 21, 23
Lysozyme, *56*, *57*, 274, 430, 818
Lyssavirus, *110*

M cells, 85
M mode echocardiography, 699
Mc4R mutation in obesity, 229
McArdle's syndrome, *1069*, *1122*, 1182
  pathophysiology, 1178
McCune-Albright syndrome, 996, 1026–7
Macroamylasaemia, *378*
Macrocytes, *398*
Macrocytic anaemias, *392*, 397–402
  megaloblastic *see* Megaloblastic anaemia
  pernicious *see* Pernicious anaemia (PA)
  without megaloblastic changes, 401–2
Macrocytosis, 325
  causes, 401
  without megaloblastic changes, 401–2
α_1-Macroglobulin, 433
Macrogols, *297*
Macrolide antibiotics, 97
Macrophage(s), 58, 431
  alveolar, 58, 812
  Bowman's space, 595, *595*
  gut, 58
  joints, 499
  specific conditions
    inflammatory bowel disease, 287
Macrophage colony stimulating factor (M-CSF), 557
Macroprolactinoma, *981*, 982
Macula, 1086
  examination, 1052
Macular amyloid, 1254
Macules, *1227*
  ash-leaf, 1253
Maculopapular rash, *1265*
Maculopathy, 1051–2, 1053, *1053*
Mad, **33**
Madura foot, 149, *150*, 153
Madurella mycetomi, 149
Magic mushrooms, 952
Magnesium
  concentration disorders *see* Hypermagnesaemia; Hypomagnesaemia
  dialysate
    haemodialysis, *634*
    peritoneal dialysis, *637*
  interstitial fluid, *650*
  intracellular fluid, *650*
  plasma/serum, *650*, 668
  renal handling, 668–9
Magnesium ammonium phosphate stones, *609*
Magnesium sulphate, 296, *297*
Magnetic resonance angiography (MRA)
  aortic coarctation, *778*
  renovascular disease, 608
  stroke, 1131
Magnetic resonance cholangiography (MRC), 370, *371*
Magnetic resonance cholangiopancreatography (MRCP), 327, 376, 380
Magnetic resonance imaging (MRI)
  abdominal, 244
  bone, 560
  brain, 1120
  cardiac
    angina pectoris, 749
    arrhythmogenic right ventricular cardiomyopathy, 791
    cardiomyopathies, 707
    congenital heart disease, 706–7
    dilated cardiomyopathy, 792
    heart failure, 737
    hypertrophic cardiomyopathy, *790*
    masses, 707
    mitral stenosis, 762
    pericardial disease, 707, *707*
    primary restrictive non-hypertrophic cardiomyopathy, 793
    valvular heart disease, 707
  cardiovascular, 706–9, *707*
    aortic disease, 707
    coronary artery disease, 707, *708*
    indications, *707*
    pulmonary vessels, 709
  enteroclysis, 277
    Crohn's disease, 289
  functional *see* Functional MRI
  kidney, 580
  lungs, 823
  pancreas, 376
  pituitary, 972, *973*, 981–2
  specific conditions
    ankylosing spondylitis, 534
    cancer, 457
    chronic renal failure, 628
    cirrhosis, 347
    Cushing's syndrome, 1013
    epilepsy, 1141
    haemochromatosis, 357
    liver disease, 327
    musculoskeletal problems, 505
    neck and shoulder pain, *506*
    osteoarthritis, 521
    pancreatitis, 378
    pericardial effusion, 704, *704*, 796
    pulmonary embolism, 785
    stroke, 1131, *1131*, *1132*
    urinary tract infections, 602
Magnetic resonance urography, 580
Mahaim fibres, 721
Major basic protein, 57, 850
Major histocompatibility complex (MHC), 59, 60–2
  see also HLA system
Malabsorption, 213, *277*, **277–82**
  bacterial overgrowth, 280
  coeliac disease, 277–9
  dermatitis herpetiformis, 279–80
  fat, 277
  intestinal resection, 280–1
  parasite infection, 282
  radiation enteritis, 282
  tropical sprue, 280
  Whipple's disease, 281–2
  see also Malnutrition; Nutrition; specific disorders
Maladie de Roger, 774
Malakoplakia, 604
Malar flush, 761
Malaria, 80, *82*, *83*, **153–7**
  airport, 153
  anaemia in, *154*
  cerebral, 155
  clinical features, 155, *155*
  diagnosis, 155–6
  epidemiology, 153, *154*
    holoendemic areas, 153
    hyperendemic areas, 153
    hypoendemic areas, 153
    mesoendemic areas, 153
  haemolytic anaemia, 415
  hyperreactive malarial splenomegaly, 155
  management, 156–7, *156*, *157*
  mortality, *80*
  parasitology, 153–5, *154*
  pathogenesis, 154–5, **154**
  prevention and control, 157
  prophylaxis, *157*
  returning travellers, *90*
  splenomegaly, 423
  see also Plasmodium species
Malarone, *156*, *157*
Malassezia, *150*, 153
  M. furfur, 1234
Malathion, 183, 1235
Males
  ageing, 998
  erectile dysfunction, 634, *995*, 1000, 1036, 1056, 1119
  hypothalamic-pituitary-gonadal axis, *995*
  musculoskeletal disease in, 501
  osteoarthritis, 520
  osteoporosis, 565
  puberty, 996
  reproductive disorders, 998–1001
  reproductive physiology, 994–5
  subfertility, 1006, *1007*
Malignancy/malignant disease, **449–97**
  in chronic renal failure, 631
  diagnosis, 454–7
  epidemiology, 449–51, *449*
    alcohol, 449
    diet, 450
    environmental/occupational, 450
    iatrogenic, 450
    infectious agents, 450
    tobacco, 449
  gall bladder, 374
  gastrointestinal, *241*
    inflammatory bowel disease, 293
  geographical distribution, 450–1
  goitre, 993
  liver, 363–4, *364*
  melanoma *see* Melanoma
  mouth, 247
  oesophageal, 256–8, *257*
  pancreas, 382–4
  prostate, 491–2, 646
  renal transplant-associated, 641
  rheumatological problems, 555
  salivary glands, 248
  secondary
    cytotoxic drugs, 462
    radiotherapy, 466
  sinonasal, 1081
  skin manifestations, 1255, *1255*
  see also Cancer; specific diseases; Tumours
Malignant hyperpyrexia, 1182
Malignant infiltration, 1177
Malignant melanoma *see* Melanoma
Malignant pericarditis, 795
Malleus, 1077
Mallory bodies, 358
Mallory-Weiss syndrome, *267*, 269
Malnutrition
  children, *216*
    care setting, 217
    prevention, 217
    prognosis, 217
    refeeding, 216, *217*
    rehabilitation, 217
  elderly people, 228
  protein-energy *see* Protein-energy malnutrition (PEM)
  rehydration solution for, 216
  risk factors, *213*
  treatment *see* Nutritional support
  see also Malabsorption; Undernutrition
Malnutrition Universal Screening Tool, 214, *215*
MALT lymphoma, *480*
Malta fever *see* Brucellosis
Maltworker's lung, *876*
Mammography, 487
Manganese, *226*
  recommended IV intake, 233
Mania, 1204–6
  aetiology, 1205
  clinical features, *1204*
  differential diagnosis, 1205
  epidemiology, 1205
  prognosis, 1206
  treatment, 1205–6
Manic-depressive disorder *see* Bipolar affective disorder
Mannan-binding lectin, 55, *56*, 59
Mannitol, 906
Mannose, acute tubular necrosis, 622
Manometry
  oesophageal *see* Oesophageal manometry
Mansonella perstans, *163*, *164*, 165
Mansonellosis, 165
Mantle cell lymphoma, *480*, 482
Mantoux test, 68, 865, 867, *867*
Maple syrup disease, *1070*
Marasmus, 215, *215*
Maraviroc, *193*
Marble bone disease, 568
Marburg virus, *110*, *114*, 121
March haemoglobinuria, 419
Marche à petits pas, 1097, 1130
Marfan's syndrome, *435*, **780–1**
  cardiac investigations, 781
  clinical features, 780–1, *780*
  genetic counselling, 781
  management, 781
  mitral regurgitation, 763
  mitral valve prolapse, 764–5
  renal tubular acidosis, *676*
Marijuana *see* Cannabis
Massive transfusion, 441
Masson's bodies, 873

Mast cells, *56*
in anaphylaxis, 73
asthma, 850
mediators, *58*
rheumatoid arthritis, 523
Mastocytosis
cutaneous, 1265
systemic, 1265
Maternal effect phenotype, 51
Maternally inherited diabetes and deafness, *1035*
Maternity blues, 1199
Matrix degradation markers, 325
Matrix deposition markers, 325
Matrix extracellular phosphoglycoprotein, 670
Matrix Gla protein, 219
Matrix metalloproteinase (MMP), 325, 345, 452, 500
osteoarthritis, 519–20
Matrix-assisted laser desorption, 48
Maturity-onset diabetes of the young (MODY) *see* Monogenic diabetes mellitus
Max, **33**
Maximum surgical blood ordering schedules, 425
MD2 protein, 895
Mean, 932
Mean corpuscular haemoglobin concentration (MCHC), *389*
Mean corpuscular volume (MCV), 389, *389*
high, *392*
low, *392*
Measles (rubeola), *83*, *110*, 117–18
clinical features, 117–18
diagnosis and treatment, 118
immunization, *102*, *103*, 118
prevention, 118
see also Measles-mumps-rubella (MMR) vaccine
rash, *88*, *117*
Mebendazole, *166*
threadworm, *166*, 167
Mechanical assist devices, 743, *743*
Mechanical haemolytic anaemia, 419
Mechanical ventilation, *915*
ALI/ARDS, 918–19
beneficial effects, 912
complications, 914
continuous positive airway pressure, 915
controlled, 912, 914–15
endotracheal intubation, 913, *913*
indications for, 912–13
institution of, 913, *913*
intermittent mandatory ventilation, 915
intermittent positive pressure, 912
positive end-expiratory pressure (PEEP), 912, 915
pressure controlled, 915
pressure support, 915
tracheostomy, 913–14, 1083, *1083*
complications, *914*
volume controlled, 915
weaning from, 916–17
see also specific techniques
Mechanic's hands, 547
Meckel's diverticulum, *269*, 270, 282
radionuclide studies, 245
Meconium ileus, 845
Median, 932
Mediastinoscopy, 828
Mediastinum, 887–8
dermoid cyst, *888*
subdivisions, *888*
Medical ethics *see* Ethics
Medical interview, **11–16**
breaking bad news, 13–14
amount of information, 14
emotional responses, 14
empathy, 14
patient's relatives, 13
positive information, 14
psychological effects, 14
closing, 13
following bad news, 14
complaints, 14–15
components, 11–13
difficult, 13–15
breaking bad news, 13–14
dealing with adversity, 16
emotional content, 14, 15
exploring/focusing
breaking bad news, 13–14
building a relationship, 11
empathy, 12, 14, 15
history taking, 11
idea/fear validation, 12
listening skills, 12
opening discussion, 11
patient facilitation, 11–12
questioning styles, 11
giving information, 13–14
aids to, 13
guidelines, 13
negotiation/collaboration, 13
planning/support, 14
research into, 12
summarizing, 13
lawsuits, 15
see also Legal issues
opening, 11
breaking bad news, 13
greeting, 11
patients with impaired communication, 16
cultural barriers, 15
hearing impaired, 16
impaired vision, 16
limited understanding/speech, 16
third parties, 16
phases, 11–13
Medical mistakes, 3
Medical profession, reputation, 1
Medical Research Council (MRC), 2
Medical treatment
age of consent, 5
competent, 6–7
patient lifestyle, 8
refusal
Mediterranean spotted fever, *147*
Medroxyprogesterone acetate
cancer therapy, 463
renal cell carcinoma, 644
Medullary cystic disease (juvenile nephronophthisis), 644
Medullary sponge kidney, 611, *611*, 644
Medulloblastoma, *1114*
Mefenamic acid, poisoning, *935*
Mefloquine, *157*
Megacolon, 297
toxic, 291, *291*
Megakaryocytes, 435
origins of, 388
Megalin, 588
Megaloblastic anaemia, 397–401
biochemical basis, 398, *398*
deoxyuridine suppression test, 398
haematological findings, 398
pernicious anaemia *see* Pernicious anaemia (PA)
prevention, 401
treatment, 401
see also Folates/folic acid; Vitamin B_{12}
Megaloblasts, *397*
Megestrol acetate, cancer therapy, 463
Meglitinides, 1039
Meibomian glands
Meiosis, 27
Melaena, 241, 267
Melanin, *493*
bronze skin, 356
Melanocortins, *229*
α-Melanocyte-stimulating hormone, 229
Melanocytes, 1226
Melanocytic naevi (moles), 1257
Melanoma, 1259–60
acral lentiginous, 1260
diagnosis, 1259, *1260*
epidemiology, *449*
familial, *453*
lentigo maligna, 1259
nodular, 1260, *1260*
oral, 247
superficial spreading, 1259
treatment, 1260
Melanonychia, 1266
Melanosis coli, 308
Melarsoprol, 158
MELAS, *37*, 39, 1182
Melasma, 1265
Melatonin, 967
Melioidosis, *82*, 148–9
Melphalan, 459
primary restrictive non-hypertrophic cardiomyopathy, 794
Membrane attack pathway, 55, *56*
Membrane biocompatibility, 624
Membrane cofactor protein, 598
Membrane inhibitor of reactive lysis, 418
Membranous glomerulopathy, 587–8
aetiopathogenesis, 588, *588*
clinical features, 587–8
treatment, 588
Memory, 1100
long-term, 1188
verbal, 1188
Memory disorders, 1100
loss, 1100
see also Amnesia
physical causes, *1190*
Men *see* Males
Menalocyte-stimulating hormone (MSH), 1008
menarche, *995*
Mendelson's syndrome, 861
Ménétrier's disease, 263, 282
Ménière's disease, 1080
Meningeal membranes (meninges)
see also Meningitis
Meningioma, 974, 1106, 1160, *1160*
Meningitis, *83*, *86*, *111*, **1153–6**, *1154*
aseptic, HIV-related, 188, 1158
bacterial, 134
acute, 1154
see also specific infections
chronic, 1155, 1158
clinical features, 1154
differential diagnosis, 1154–5
incidence, *1096*
infective causes, *1154*
malignant, 1156
meningitic syndrome, 1154
Mollaret's, 1159
mortality, *80*
pathology, 1153
prophylaxis, 1155
pyogenic, *1155*
treatment, 1155, *1155*
emergency, *1154*
tuberculous, *1155*
management, 1155–6
viral, 1154
Meningocephalocele, 1169
Meningococcal arthritis, 540
Meningococcal septicaemia, 80, *83*, 136
clinical features, 136
diagnosis, 136
management, 136
rash, *88*, *136*
Meningococcus, *70*
Meningogoccus, immunization, *103*
Meningomyelocele, 1169
Menkes' kinky hair syndrome, 225
Menopause, 997–8, 1195
clinical features and treatment, 997
hormone replacement therapy *see* Hormone replacement therapy
premature, 462, 997–8
Menstrual cycle, 967
hormonal and follicular changes, *995*
Menstruation, *995*
disorders of, 963
see also Polycystic ovary syndrome (PCOS)
iron loss, 395, 396
Mental Capacity Act 2008, 1222, *1223*
Mental disorders
Mental Health Act (1983), 6
Mental Health Act (2008), 1222
Mental State Examination (MSE), 1186–9
Mepacrine
chronic discoid lupus erythematosus, 1252
giardiasis, 162
Mephalan, side-effects, leukaemia, 467
Meptazinol, osteoporosis, 563
Meralgia paresthetica, 512
6-Mercaptopurine, 460
Crohn's disease, *289*, 290
genetic polymorphism in metabolism, *926*
Mercury poisoning, 945
antidotes, *937*
clinical features, 945
membranous glomerulopathy, 588
treatment, 945
uraemia, 619
Merkel cells, 1226
MERRF, *37*, 39, 1182
Mesalazine, ulcerative colitis, 292
Mesangial proliferative glomerulonephropathy, 590
Mesangiocapillary (membranoproliferative) glomerulonephritis, 589–90, *590*
management, 590
Mesangium, 582
Mesothelioma, 879–80, 885
Messenger RNA, *1030*
Meta-analysis, 931
Metabolic acidosis, 675–8, *675*, 894, 1047
acute tubular necrosis, 622
high anion gap, 677
lactic *see* Lactic acidosis
mixed, 678
normal anion gap, 675–7, *675*
and tissue perfusion, 900
transfusion-related, 906
Metabolic alkalosis, *675*, 678–9, *678*, 894
causes, *678*
chloride-resistant, 679
chloride-responsive, 679
clinical features, 679
transfusion-related, 906
treatment, 679
Metabolic disease(s)
diabetes mellitus *see* Diabetes mellitus
hyperlipidaemia, **1063–8**
inborn errors of metabolism *see* Inborn errors of metabolism

Metabolic disease(s) (*Continued*)
lipid metabolism, **1061–8**
porphyria(s), 1073–5
spinal cord, 1165
see also specific diseases
Metabolic myopathy, 1179–80
Metabolic neuropathies, 1174, *1174*
Metabolic syndrome (syndrome X), 231, *231*, 799, 1034
antiretroviral drugs, 198
see also Obesity
Metabolism
fasted state, *213*, 214
fed state, *213*, 214
regulation of, 214
Metabolomics, 48
Metacarpophalangeal joints (MCP), 509
osteoarthritis, 521
Metals *see* individual metals
Metamyelocytes, 397, 430
Metapneumovirus, 119
Metastases, 452, *452*
bone, 452
CNS, treatment, 457
lung, 883–4
osteoblastic, 452
osteolytic, 452, *452*
of unknown origin, 492–3, *493*
diagnosis, 493
prognosis, 493
treatment, 493
Metatarsalgia, 515
Metatarsophalangeal joints (MTP), 515
Metatarsus primus varus, 515
Metformin
diabetes, 1038, *1038*, 1041
infertility, 1005
menstrual disorders, 1005
overdose, 940
metabolic acidosis, *939*
Methacholine provocation test, 852
Methadone, side-effects, long QT syndrome, *727*
Methaemalbumin, 403
Methaemoglobin, 403
Methaemoglobinaemia, antidotes, *937*
Methanol
metabolism, *945*
poisoning, *935*, 945–6
antidotes, *937*
clinical features, 945–6
metabolic acidosis, *939*
treatment, 946
Methionine, 210
Methionine malabsorption syndrome, 1071
Methotrexate, 460
acute lymphoblastic leukaemia, 471
ANCA-positive vasculitis, 597
ankylosing spondylitis, 534
asthma, *854*, 855
breast cancer, *461*, 488, 489
bullous pemphigoid, 1256
Burkitt's lymphoma, 483
Crohn's disease, *289*, 290
juvenile idiopathic arthritis, 554
polymyositis/dermatomyositis, 548
psoriasis, 1242
psoriatic arthritis, 535
rheumatoid arthritis, 530–1, *530*
sarcoidosis, 1253
side-effects
hepatotoxicity, *366*
respiratory, *877*
Methoxycinnamate, 1248
Methylcellulose eye drops, 991
Methylcobalamin, 398, *398*
Methyldopa
hypertension, 805
side-effects
hepatotoxicity, *366*
immune haemolytic anaemia, 417
rash, *1265*
retroperitoneal fibrosis, 616
Methylene tetrahydrofolate, 398
Methylene tetrahydrofolate reductase deficiency, 223, *1070*
3,4-Methylenedioxymethamfetamine (MDMA) *see* Ecstasy
Methylmalonic acid, pernicious anaemia, 400
Methylnaltrexone, *297*
Methylphenidate, narcolepsy/cataplexy, 145
Methylphenyltetrahydropyridine, and Parkinson's disease, 1146
Methylprednisolone
multiple sclerosis, 1153
pemphigus vulgaris, 1256
thrombotic thrombocytopenic purpura, 437
Methylthioninium, *937*
Methysergide, side-effects, retroperitoneal fibrosis, 616
Meticillin, *94*
Metoclopramide, 252, 461, 495, 496
migraine, 1137
myocardial infarction, 757
poisoning, *935*
side-effects, endocrine disease, *964*
Metolazone, *657*, 658
heart failure, *739*
Metoprolol, 730
acute coronary syndromes, *754*
hypertension, 804
supraventricular tachycardia, 723
Metronidazole, 99, *99*
amoebiasis, 161
amoebic abscess, 361
bacterial vaginosis, 183
balantidiasis, 162
Crohn's disease, *289*
diverticulitis, 299
empyema, 863
gastroenteritis, *132*
giardiasis, 162
gingivitis, 247
H. pylori eradication, 261
rosacea, 1247
side-effects, neuropathy, *1174*
trichomoniasis, 182
Metyrapone
Cushing's syndrome, 1013
side-effects, endocrine disease, *964*
Mexiletine, *731*
diabetic neuropathy, 1055
Micelles, 322
mixed, 273
Miconazole, 100, *100*
ringworm, 1233
seborrhoeic eczema, 1239
Microalbuminuria, 579
diabetic nephropathy, 589, 1054
Microangiopathic haemolysis, 598
Microangiopathic haemolytic anaemia, 419
Microarrays, 34–5, *35*
Microcytic anaemia, *392*, 393–7
causes
iron deficiency, 393
see also Iron; Iron deficiency anaemia
see also Thalassaemia(s)
differential diagnosis, *396*
sideroblastic *see* Sideroblastic anaemia
Microembolism, transfusion-related, 906
Microfilaments, 25–6, *25*
β_2-Microglobulin, 576
myeloma, 485
non-Hodgkin's lymphoma, 481
Micrographia, 1146
Micropolyspora faeni, 877
Microprolactinoma, *981*, 982
Microsatellite DNA
instability, 50
somatic, 50
see also Triplet (trinucleotide) repeat expansion/mutation
Microsatellite instability, 302, 303
Microscopic inflammatory colitis, 294
Microscopic polyangiitis, *549*, 596, 871
ANCA antibodies in, 596
clinical features, *550*
haemoptysis, 819
diagnosis, *550*
Microscopy
Microsporea, 162
Microsporidia, *191*
Microsporidiosis, 162
HIV-associated, 201
Microsporum, 153
Microstomia, 546
Microtubule-organizing centre, 24
Microtubules, 24, *24*
Microvilli, 270
Micturating cystourethrography, 581
Micturition
neurological disorders, 1119
see also Urinary incontinence
Micturition syncope, 689, 1144
Mifepristone, 1006
Miglustat, 1071
Migraine, 1136–7
with aura, 1137
basilar, 1109, 1137
clinical patterns, 1137
differential diagnosis, 1137
facioplegic, 1137
food intolerance, 237
hemiparetic, 1137
management, 1137
mechanisms, 1136
ophthalmoplegic, 1137
prophylaxis, 1137
without aura, 1137
Migrating motor complex, 271
Mikulicz syndrome, 248
Milia, 1268
Miliary mottling, 822–3
Miliary tuberculosis, 822, 864–5
investigation, 865
Milk-alkali syndrome, 679
Millennium Development Goals, 80
Miller Fisher syndrome, 1173
Milrinone, 907
Milroy's disease, 1263
Mineralocorticoids, 1008
secretion, 1009
sodium retention, 656
Minerals, **225–6**
absorption, 274
adult requirements, 234
deficiency, 275–6
parenteral nutrition, 236
recommended IV intake, *233*, 234, *234*
storage in liver, 321
see also individual minerals
Mini-Mental State Examination, 1188–9, *1189*
Minimal-change nephropathy, 585–6, *586*
clinical features, 586
management, 586
Minimally conscious state, *1125*, 1126
Minocycline
acne vulgaris, 1246
bullous pemphigoid, 1256
leprosy, *141*
pyoderma gangrenosum, 1251
side-effects
rash, *1265*
vasculitis, 596
Minoxidil
androgenic alopecia, 1267
hypertension, 805
side-effects, hypertrichosis, 1003
Miosis, senile, 1103
Mirizzi's syndrome, 371
Mirtazapine, 1203
Misconceptions, communication difficulties for patients, 10, 11
Misoprostol, 263
Mitochondria, *19*, 23–4
chromosomal disorders, 39
chromosomes, 37, *37*
diseases *see* Mitochondrial disease
Mitochondrial diseases, 44, 1178, 1182
Mitochondrial toxicity, antiretroviral drugs, 198
Mitogen activated protein kinase, 24
Mitomycin C, 460
bladder tumours, 645
side-effects, haemolytic uraemic syndrome, 598
Mitosis, 27, 36
Mitoxantrone
acute lymphocytic leukaemia, 475
breast cancer, 489
multiple sclerosis, 1153
Mitral facies, 761
Mitral regurgitation, **763–5**
causes, 763
heart sounds, 692
investigations, 763–4
cardiac catheterization, 764
chest X-ray, 763
echocardiogram, 763–4
electrocardiogram, 763, *764*
pathophysiology, 763
post-STEMI, 759
signs, 763
symptoms, 763
treatment, 764
Mitral stenosis, 692, **760–3**, 891
complications, *760*
echocardiogram, 703, 762
heart sounds, 692
investigations, 762, *762*
pathophysiology, 760
rheumatic, 760, *760*
signs, 761
symptoms, 760–1
treatment, 762–3
Mitral valve, 684
disease, chest X-ray, *693*
prolapse, 764–5
autosomal-dominant polycystic kidney disease, 642
investigations, 764
pathophysiology, 764
signs, 764
symptoms, 764
treatment, 764–5
replacement, 763
rheumatic disease, *703*
Mittelschmerz, 316
Mivacurium, genetic polymorphism in metabolism, *926*
Mixed antagonistic response syndrome (MARS), *899*
Mixed metabolic acidosis, 30
Mixed venous oxygen tension/saturation, 893, *893*
Mobiluncus, bacterial vaginosis, 183
Mobitz type I block, *718*
Mobitz type II block, *718*
Moclobemide, 1203
Modafinil, narcolepsy/cataplexy, 145
Mode, 932
Molecular biology, **29–37**
DNA structure/function, **29–33**
see also Gene(s)
tools, **33–6**

Molecular genetics
applications, **46–9**
see also Genetic disorders
Molecular mimicry, 74, 532, 987
Molecular theories of ageing, 226
Moles *see* Melanocytic naevi (moles)
Mollaret's meningitis, 1159
Molluscum contagiosum, *104*, 1232
conjunctivitis, 1089
HIV-related, 1268
Molybdenum, *226*
recommended IV intake, 233
Mometasone, asthma, 855
Mongolian blue spot, 1268
Monitoring
HAART therapy, 196, *196*
heart failure, 738
HIV infection, 190–1
Holter, 697, *698*, 737
parenteral nutrition, 236
respiratory failure, 911
Monkeypox, *81*, *104*, 109
Monoamine A, reversible inhibitors (RIMAs), 1203
Monoamine oxidase inhibitors (MAOIs), 1203
overdose, 946
side-effects, hypertension, 799
Monobactams, 95
Monoclonal antibodies, 76, *76*
asthma, 856
cancer therapy, 464, *464*
IgM, 417
side-effects, respiratory, *877*
in treatment of allergy, 72
see also Antibodies; and individual drugs
Monoclonal gammopathy of unknown significance, 485
Monocyte-derived multipotential cells, 47
Monocytes, *56*, 58, 431
origins of, 388
Monocytosis, 431
Monogenic diabetes mellitus, 1035–6, *1035*
Monoglycerides, 273
Mononeuritis, 1056
Mononeuritis multiplex, 1056, 1171, 1172
Mononeuropathies, 1171–2, *1171*
Mononuclear cells, *56*
Monoparesis, *1099*
Monosodium glutamate, intolerance to, 237
Monosodium urate deposition *see* Gout
Monospot test, 108
Montelukast
asthma, *854*, 856
rhinitis, 833
Mood, 1186–7
disorders *see* Mood (affective) disorders
diurnal variation, 1186–7
incongruous, 1187
labile/fluctuating, 1187
reactive, 1187
Mood (affective) disorders, **1198–206**
classification, 1198
depression *see* Depression
Moral arguments, 1
Moraxella
M. catarrhalis, *123*, *191*, 838
pneumonia, 861
M. lacunata, conjunctivitis, 1088
Morbidity
obesity, 230, *230*, *231*
Morphine
acute coronary syndromes, *754*
cancer pain, 495
dose titration, 495
musculoskeletal problems, 517
sickle cell anaemia, *411*
side-effects, 495
Morphoea, 1252
Mortality
inflammatory bowel disease, 293
obesity, 230, *230*
Morton's metatarsalgia, 515
Morton's neuroma, *515*
Mosquitoes
disease transmission, *82*
see also Malaria
Motilin, *272*
Motion sickness, 960
Motor conduction velocity (MCV), 1121, *1121*
Motor control systems, 1111, *1111*
Motor cortex, hemiparesis, 1112
Motor neurone disease (MND), *1110*, 1166–7
clinical features, 1166
course and management, 1167
diagnosis, 1166–7
Motor neurone lesions
lower *see* Lower motor neurone (LMN) lesions
upper *see* Upper motor neurone (UMN) lesions
Moulds *see* Fungal infections
Mountain sickness
acute, 956
chronic, 956
deterioration, 956
Mouth, **246–7**
acute abdomen, 315
burning mouth syndrome, 246
dry, 248
functions of, 246
gums *see* Gingivae
HIV infection, 248
pigmented lesions
neoplastic, 247
non-neoplastic, 247
recurrent aphthous ulceration, 246–7
teeth *see* Teeth
tongue *see* Tongue
white patches, 247
see also Oral
Movement, 503
disorders *see* Movement disorders
spinal, 510
Movement disorders, 1145–51, *1146*
akinetic-rigid syndromes, 1145–9
dyskinesia *see* Dyskinesias
loss of skilled movement, 1111–12
Moxifloxacin, 98
tuberculosis, 866
Moxonidine, hypertension, 805
mRNA *see* Messenger RNA
MRSA, 124
community-associated, 124
detection, 124
Mucins, *25*
Mucociliary clearance, 843
Mucociliary escalator, 812, 818
Mucocutaneous leishmaniasis *see* Leishmaniasis, mucocutaneous
Mucopolysaccharidoses, 1071–2
Mucor, *82*, 152
Mucormycosis, *82*, *150*, 152
rhinocerebral, 152
Mucosa
duodenum, 258
nasal, 814
oesophagus, 250–1
stomach, 258
in ulcerative colitis, 287
Mucosa-associated lymphoid tissue (MALT), 53, 262
see also Immunoglobulin A (IgA)
Mucositis, chemotherapy-induced, 461
Mucus
gastrointestinal tract, 258
respiratory tract, 818, *818*
removal of, 812
see also Sputum
Müllerian duct, 994
Multi-infarct dementia, 1130
Multidisciplinary care, 890
Multidrug-resistance proteins, 22, 322, 330
Multifactorial inheritance, 44
Multiorgan failure *see* Multiple organ failure (MOF)
Multiple chemical sensitivity, 1194
Multiple endocrine neoplasia syndromes, 453, *453*, 1026, *1026*
Multiple endocrine syndrome type I, 384
Multiple gland failure, 1026
Multiple mononeuropathy *see* Mononeuritis multiplex
Multiple myeloma, *485*
renal involvement, 599
Multiple organ dysfunction syndrome (MODS) *see* Multiple organ failure (MOF)
Multiple organ failure (MOF), 909–10
management, 909–10
pancreatitis, 379
Multiple sclerosis, *1114*, **1151–3**
aetiology, 1151
familial, 1151
HLA linkage, *60*, 1151
infection, 1151
migration, 1151
autoantigens, *74*
clinical features, 1151–2
brainstem demyelination, 1152
optic neuritis, 1102
optic neuropathy, 1152
spinal cord lesions, 1152
diagnosis, 48
differential diagnosis, 1152
end-stage, 1152
investigations, 1152–3, *1152*
management, 1153
pathogenesis, 1151
pathology, 1151, *1151*
prevalence, 1151
prognosis, 1153
rehabilitation, 1153
Multiple system atrophies, 1148
Multisystem disorders, 39
Mumps, *83*, *110*, 118
clinical features, 118, *118*
complications, 118
diagnosis and treatment, 118
immunization, *103*, 118
prevention, 118
see also Measles-mumps-rubella (MMR) vaccine
Mupirocin, impetigo, 1228
Muscle(s)
atrophy *see* Muscle atrophy (wasting)
biopsy, 1122, 1178–9
breakdown (proteolysis)
diseases *see* Muscle diseases
fatigue, 1097
glucose utilization, 1029
infection, *1179*
jerking
non-organic, 1150
see also Myoclonus
pain *see* Musculoskeletal pain
power, *1098*
respiratory, 814
rigidity, 1111, 1112, 1146
skeletal, 501
spasm, 1106–7, 1108, 1148, 1150
tone, 1112
asymmetry, 1125
weakness, 1178
see also entries beginning myo-
Muscle atrophy (wasting), 1165
diabetic amyotrophy, 1056
marasmus, 215
progressive, 1166
rheumatoid arthritis, 527
spinal, 1167
see also specific muscular atrophies
Muscle diseases, 1178–82
definitions, 1178, *1178*
drug-induced, 1179
dystrophies *see* Muscular dystrophies
glycogenoses, *1069*
pathophysiology, 1178
recognition, 1178–9
see also Cardiomyopathy; Myopathy; specific diseases
Muscle dysfunction
chronic renal failure, 630
systemic lupus erythematosus, 542
Muscle weakness, 1111
extraocular muscles, 1105
Muscle-cell-mediated gene therapy, 46–7
Muscular atrophy *see* Muscle atrophy (wasting)
Muscular dystrophies, 1178, 1181
Becker's, 1181
Duchenne *see* Duchenne muscular dystrophy (DMD)
limb-girdle and facioscapulohumeral, 1181, *1181*
oculocerebrorenal *see* Lowe's syndrome
Musculoskeletal pain (myalgia), 501
drug treatment, *see* Analgesia
Musculoskeletal problems, **505–16**
analgesics/anti-inflammatories, 517–18
diagnostic imaging, 504–5
examination, 502–3
history, 501–2
investigations, 503–5
neck and shoulder pain, 505–7
serum autoantibodies, 503
see also specific conditions
Mushroom worker's lung, *876*
Mushrooms, 952
Mutation(s), 40–2
consequences
see also Cancer; Genetic disorders
insertion/deletion, 41
oncogene activation, 49
point, 40–1
splicing, 41
termination, 42
Myalgia, *1179*
epidemic, 885
systemic lupus erythematosus, 542
Myalgic encephalomyelitis *see* Chronic fatigue syndrome (CFS)
Myasthenia, 1178
incidence, *1096*
Myasthenia gravis, 75, 1180–1
antinuclear antibodies, *504*
autoantigens, *74*
clinical features, 1180
course and management, 1180
diagnosis, *69*
investigations, 1180
pathophysiology, 1178
treatment, 1180
ventilatory failure, 912–13
myc, **33**, 49
myc gene, *453*
Mycetoma, 149, *150*, 153
Mycobacteria
Mycobacterial infections
skin, 1230
Mycobacterial ulcer, 142

Mycobacterium, 85, *90*, 122, *123*
M. africanum, 145
M. avium-intracellulare, *70*, *146*, 189, *191*, 868
HIV-related, 204
M. bovis, 145
M. chelonae, *146*
M. fortuitum, *146*
M. haemophilum, *146*
M. kansasii, *146*, 868
M. leprae, 140, 1230
see also Leprosy
M. malmoense, *146*
M. marinum, *146*, 1230
M. paratuberculosis, 287
M. scrofulaceum, *146*
M. tuberculosis, *70*, 145, *191*, 282, *1179*
CAPD peritonitis, *637*
HIV-associated, 203–4
pneumonia, 857, *858*
see also Tuberculosis
M. ulcerans, 142
M. xenopi, *146*
Mycofungin, 100, *100*
Mycophenolate mofetil, 76
ANCA-positive vasculitis, 597
autoimmune hepatitis, 344
bullous pemphigoid, 1256
Crohn's disease, 290
lupus nephritis, 591
membranous glomerulonephropathy, 588
pemphigus vulgaris, 1256
polymyositis/dermatomyositis, 548
renal transplantation, 640
retroperitoneal fibrosis, 617
thrombocytic purpura, 436
Mycoplasma, *70*, 122
M. genitilium
urethral discharge, *172*
urethritis, 176
M. hominis, bacterial vaginosis, 183
M. pneumoniae, *123*
pneumonia, *858*
Mycoplasma pneumonia, 859
Mycoplasmosis, 795
Mycosis fungoides, 483, 1260
Mycotic aneurysm, infective endocarditis, *770*
Myelitis, 1157
transverse, 1165
Myeloblasts, 430
Myelodysplasia (MDS), 397, 402, 422, *422*
classification, *422*
clinical and laboratory features, 422
management, 422
Myelofibrosis (myelosclerosis), 421–2
clinical features, 421
differential diagnosis, 421
investigations, 421
prognosis, 422
treatment, 421–2
Myeloid bodies, 597
Myeloid cells, 53
Myeloid dendritic cells, *56*, 58, *59*
Myeloid differentiation factor 88, 895
Myeloid leukaemias *see* under Leukaemia
Myeloma, **484–6**
bone disease, 484
clinicopathological features, 484
cytogenetics, 484
diagnosis, 485
epidemiology, *449*
investigations, 484–5
life-threatening complications, *485*
multiple *see* Multiple myeloma
symptoms, 484, *485*
treatment, 485–6
Myelopathy, 1071
acute transverse, 1165
cervical, 1177
radiation, 1165
Myeloperoxidase, 56, *57*, 430, 431
ANCA, 870
musculoskeletal problems, 504
Myeloproliferative disorders, **419–22**
hepatomegaly, *331*
myelodysplasia (MDS), 422, *422*
myelofibrosis (myelosclerosis), 421–2
polycythaemia *see* Polycythaemia
Myelosclerosis *see* Myelofibrosis (myelosclerosis)
Myenteric plexus, 271
MYH-associated polyposis, 302
Myiasis, 172
Myocardial calcification, 693
Myocardial cells (myocytes), 681
action potentials, 682, *682*
apoptosis, 735
depolarization, 682
excitation-contraction coupling, 681, 682–3, *683*
'funny' current, 682
hyperpolarization, 682
pacemaker potential, 682
Myocardial contractility, 734, 891, *892*
inotropic agents, 907–8, *908*
Myocardial disease, **787–94**
atrial myxoma, 787, *787*
see also Cardiomyopathy; Myocarditis
Myocardial failure *see* Heart failure
Myocardial infarction
acute abdomen, *315*
left bundle branch block, *720*
non-ST-elevation (NSTEMI), 752
plaque thrombosis, *745*
radionuclide imaging, 705
right bundle branch block, *720*
second-degree heart block, 718
ST-elevation (STEMI), 752, 755–9
complications, 758–9
early management, 757–8
electrocardiogram, 755–7, *756*, *757*
investigations, 757
lifestyle modification, 759
pathophysiology, 755
symptoms and signs, 755
TIMI risk score, *756*
see also Cardiac arrest; Cardiac failure
Myocardial ischaemia, 743
causes, 744
ST segment depression, 696–7, *697*
see also Ischaemic heart disease; Myocardial infarction
Myocardial oxygen consumption, 891
Myocardial perfusion
echocardiography, 702–3
radionuclide imaging, 705, *705*
Myocardial remodelling, 734–5, *735*
Myocardial stunning, 740
Myocardial viability, 705, *705*
Myocarditis, *111*, 787–8, *787*
causes, *787*
clinical features, 787–8
giant cell, 788
investigations, 788
pathology, 787, *787*
systemic lupus erythematosus, 543
treatment, 788
Myocardium
calcification, 693
cells *see* Myocardial cells (myocytes)
hibernating, 740
nerve supply, 684
rupture, 758–9
see also Myocardial
Myoclonus, *455*, 1149–50
benign essential, 1150
in epilepsy, 1150
nocturnal, 1150
MyoD, 51
Myofibrils, 501, 682, *683*
Myoglobin, acute coronary syndromes, 752
Myoglobinaemia, 621
Myoinositol, 655
Myokymia, 1108
Myopathy, 1178
critical illness, 916
inflammatory, 1179
metabolic, 1179–80, 1182
sarcoidosis and rheumatoid disease, 1179
symptom pattern, *549*
Myophosphorylase deficiency *see* McArdle's syndrome
Myopia, 1086, *1086*
Myosin, *493*, 501, 682
Myositis, *111*, 885, 1178, *1179*
acute suppurative (tropical), *1179*
inclusion body, 548
necrotizing, *1179*
toxoplasma, *1179*
tuberculous, *1179*
viral/bacterial/parasitic, 1179
Myositis-specific antibodies, 547
Myotonia, 1178, 1181–2
Myotonia congenita (Thomsen's disease), 1182
Myotonic dystrophy *see* Dystrophia myotonica
Myristic acid, *210*
Myxoedema, *965*
pretibial, 988, 1254
see also Hypothyroidism
Myxoedema coma, 987
Myxoedema madness, 987
Myxoma, atrial, 787, *787*

N-myc, 49
Naevus
blue, 1257
cellular, 1257
compound, 1257
flammeus, 1268
giant congenital melanocytic, 1258
junctional, 1257
melanocytic, 1257
spider, 328
strawberry, 1267–8, *1268*
Nails, 1227
brittle, 395
in dermatomyositis, 1251
in psoriasis, 1242, *1242*
sampling, *1228*
Nail disorders
Beau's lines, 1266
dystrophy, 534, 536
koilonychia, 392, 395, 1266
leuconychia, 1266
melanonychia, 1266
nail fold infarcts, 528
onychogryphosis, 1266
onycholysis, *1242*, 1266
paronychia, 1234
drug-induced, 1269
periungual fibroma, 1253
pitting, 1266
splinter haemorrhages, 690, *770*
T. unguium, 1233, *1233*
yellow-nail syndrome, 1266
Nail-patella syndrome, 1266
Nairovirus (Bunyavirus), *82*
Nalidixic acid, gastroenteritis, *132*
Naloxone, *937*, 947
primary biliary cirrhosis, 355
Naltrexone
alcohol dependence, 1214
primary biliary cirrhosis, 355
NANC fibres, 249
Nappy rash, 1268
Naproxen, gout, 538
Naratriptan, migraine, 1137
Narcolepsy, 1145
Narcotics *see* Opiates
Nasal cannulae, 912, *912*
Nasal decongestants, 1081
Nasal discharge
Nasal packs, 1081, *1081*
Nasal polyps, 819, 1081
and rhinitis, 832
Nasal tumours, 819
Nasogastric suction, 379
Nasogastric tube feeding, *235*
Nasopharyngeal carcinoma, *105*, *1110*
Nasopharynx, 1082
Natalizumab, multiple sclerosis, 1153
Nateglinide, 1039
Natilizumab, Crohn's disease, *290*
National Health Service (NHS), 7
resources, 7–8
National Institute for Clinical Excellence (NICE), 8
Natriuresis, 653
Natriuretic peptides, 799
heart failure, 735–6
resistance to, 653
see also Atrial natriuretic factors/peptides (ANPs)
Natural disasters, 953
Natural killer cells, *56*, 66
activation, 66
receptors, *66*
Natural resistance-associated macrophage protein, 394
Nausea, 242
chemotherapy-induced, 461
opiate therapy, 495, 496
see also Vomiting
Naxos disease, 791
Nebivolol, heart failure, *739*
Necator americanus, *163*, 166
Neck pain, 505–7, *505*
elderly patients, 552
mechanical/muscular, 505–6
nerve root entrapment, 506, *506*
whiplash injury, 506–7
Necrobiosis lipoidica, 1254
Necrolytic migratory erythema, malignancy-related, *1255*
Necrosis
aseptic bone *see* Avascular (aseptic) necrosis of bone
kidney
acute tubular *see* Acute tubular necrosis (ATN)
Necrotic cell death, 27–8
Necrotizing fasciitis, 125
Necrotizing myositis, *1179*
Nedocromil
allergic conjunctivitis, 1089
asthma, 855
rhinitis, 833
Needle-stick injuries, 184, 340
Negative feedback, 967, *967*
Negative selection *see* Tolerance
Negligence *see* Clinical negligence
Negri bodies, 120
Neisseria
N. gonorrhoeae, *123*, *175*
antibiotic resistance, *94*
nucleic acid amplification test, 174–5
septicaemia, *135*
urethral discharge, *172*
vaginal discharge, *172*
see also Gonorrhoea
N. meningitidis, *90*, *123*, 136
septic arthritis, 539
Nelson's syndrome, 972, *972*, 1014
skin manifestations, 1265
Nematodes (roundworms), 163–8, *163*
human intestinal, 166–7
ascariasis, 166, *166*

hookworm, 167
strongyloidiasis, 167
threadworm *see* Threadworm (*Enterobius vermicularis*)
whipworm, 167
tissue-dwelling, 164–6
dracunculiasis, 166
filariasis *see* Filariasis
loiasis, 165
mansonellosis, 165
onchocerciasis, 165
zoonotic, 168
Neomycin, 96, *96*
Neonatal alloimmune thrombocytopenia, 437
Neonatal herpes, *105*
Neonatal lupus syndrome, 544
Neonatal tetanus, 134
Neonatal thyrotoxicosis, *965*
Neonates
heart disease *see* Congenital heart disease
prescribing in, 924
rashes in, 1267–8, *1268*
vitamin K deficiency in, 219–20
see also Children; Infants
Neoplasms/neoplastic lesions *see* Cancer; Malignancy/malignant disease; Tumours
Nephrin, 583
Nephritic syndrome
acute, 584, 592–3
diseases associated with, *592*
glomerulonephritis with infective endocarditis, 593
glomerulonephritis with visceral abscesses, 593
post-streptococcal glomerulonephritis, 592–3, *593*
Nephritis
Heymann, 588
tubulo-interstitial *see* Tubulointerstitial nephritis (TIN)
see also Glomerulonephritis
Nephroblastoma (Wilms' tumour), 39, *453*, 645
Nephrocalcin, 610
Nephrocalcinosis, 613–14, *613*, *614*
causes, *614*
renal tubular acidosis, *676*
see also Urinary tract stones
Nephrogenic diabetes insipidus, 575, 654, 1018
Nephrogenic fibrosing dermopathy, 1254
Nephrogenic systemic fibrosis, 629
Nephron, 571, *572*, *652*
Nephropathy
analgesic, 605–6
Balkan, 606
C1q, 590
Chinese herb, 606
chronic allograft, 641
contrast, 624–5
diabetic *see* Diabetic kidney
HIV-associated, 189, 587, 597
hyperuricaemic *see* Hyperuricaemia, nephropathy
IgA, 593–4, *593*
clinical presentation, 594
management, 594
pathogenesis, 593–4, *593*
IgM, 590
light chain cast, 599, *599*
minimal change *see* Minimal-change nephropathy
obstructive *see* Urinary tract obstruction
radiocontrast, 599
reflux, 600–1, *601*, *605*, 676
sickle, 597
see also Renal disease
Nephrosclerosis, hypertensive, 607
Nephrotic syndrome, 584–92
with 'active' urine sediments, 589–92
cryoglobulinaemic renal disease, 591, *591*
fibronectin glomerulonephropathy, 592
Henoch-Schönlein purpura, 592
idiopathic fibrillary glomerulopathy, 592
immunotactoid glomerulopathy, 592
lupus glomerulonephritis, 590–1
mesangial proliferative glomerulonephropathy, 590
mesangiocapillary (membranoproliferative) glomerulonephritis, 589–90, *590*
with 'bland' urine sediments, 585–9
amyloidosis, 588–9, *589*
congenital nephrotic syndrome, 586
diabetic nephropathy, 589, *589*
focal segmental glomerulosclerosis, 586–7, *587*
membranous glomerulopathy, 587–8, *588*
minimal-change glomerular lesion, 585–6, *586*
glomerular lesions, *585*
HIV-related, 189
hypoalbuminaemia in, 321
management, 585
oedema, 655–6
pathophysiology, 584
hyperlipidaemia, 584
hypoalbuminaemia, 584
proteinuria, 584
Nephrotoxicity, chemotherapy-induced, 462
Nerium oleander, 952
Nerve(s)
neural transmission
see also Nerve root; Nerve supply
Nerve agents, *935*, 946
antidotes, *937*
G agents, 946
V agents, 946
see also Bioterrorism/biowarfare
Nerve blocks, 1119
Nerve entrapment, 1116
Nerve fibres, *1171*
pain, 1119
Nerve root
entrapment, 506
lumbar nerve root, *511*
neck, 506, *506*
lesions, 1176–7, *1176*
see also Spinal root
Nerve supply
airways, 852
bladder, *1119*
eyes, 1086
lung(s), 814
myocardium, 684
oesophagus, 248–9, *249*
skin, 1227
small intestine, 271
Nervous system
autonomic *see* Autonomic nervous system (ANS)
central *see* Central nervous system (CNS)
diseases *see* Neurological disease
infections, **133–5**
bacterial meningitis, 134
botulism, 134
tetanus, 134–5
peripheral *see* Peripheral nervous system (PNS)
rheumatoid arthritis, 528
systemic lupus erythematosus, 543
Nesritide, 736
Netilmicin, 96, *96*
Nettle rash *see* Urticaria
Neural tube defects
folate deficiency, 401
see also Folates/folic acid
inheritance, *44*
maternal folate status, 401
Neuralgia, 1116
postherpetic, 1116, 1157
trigeminal *see* Trigeminal neuralgia
Neuralgic amyotrophy, 1177
Neuraminidase, 117
Neuraminidase antigen, 834
Neuraminidase inhibitors, 101–2, *101*
influenza, 834
Neurasthenia *see* Chronic fatigue syndrome (CFS)
Neuritis
mononeuritis, 1056
optic, 1101–2
retrobulbar, 1102, 1152
vestibular, 1109
Neuro-imaging, 1119–22
Neuroanatomy, 1100
Neurocardiogenic (vasovagal) syncope, 717
Neurocysticercosis, 1141, 1157
Neuroectodermal syndromes, 1169–70
Neuroendocrine peptides, 271
see also individual hormones/peptides
Neuroendocrine system
peptides *see* Neuroendocrine peptides
Neurofibrillary tangles, 1168
Neurofibroma (Schwannoma), *1110*, 1160
Neurofibromatosis, *789*
type 1 (von Recklinghausen's disease), 453, *453*, 1169–70, 1253
type II, 1170
Neurogenetic tests, 1178
Neurokinin A, 814
asthma, 852
Neurokinin B, 814
Neuroleptic malignant syndrome, 1217
Neuroleptics, 1217, *1217*
atypical, 1217–18
overdose, 946
side-effects, *928*, 1148, *1217*
elderly patients, *924*
long QT syndrome, *727*
teratogenicity, 1217
Neurological disease, **1095–183**
congenital and inherited, *see* Neural tube defects
examination and formulation, 1097–*8*
HIV-related, 1158
impact of, 1095
incidence, *1096*
malignancy, *see* Brain tumours; Spinal cord tumours
symptoms and signs, 1095–8
see also Neuropathies; specific diseases
Neurological symptoms/signs, sickle cell disease, 419
Neuromuscular junction conditions, 1180–1
Neurones, 1098, *1098*
Neuronal disease
degenerative, **1166–9**
gene therapy, 47
Neurone specific enolase, *493*
Neuropathic joints *see* Charcot's joints
Neuropathic ulcers, 1261–2
Neuropathies, 1171
autonomic, 1176
of cancer, 1175
chronic sensorimotor, 1173
diabetic *see* Diabetic neuropathy
diphtheritic, 1173
hereditary sensorimotor, 1175
idiopathic multifocal motor, 1172
metabolic, 1174, *1174*
mononeuropathies *see* Mononeuropathies
motor
optic *see* Optic neuropathy (ON)
polyneuropathies *see* Polyneuropathies
rheumatoid arthritis, 528
of systemic diseases, 1175
toxic, 1174, *1174*
trigeminal, 1107
tropical axial, 1165
vitamin deficiencies, 1174–5, *1175*
see also individual conditions; Paraparesis
Neuropathy, ataxia and retinitis pigmentosa, *37*
Neuropeptide Y, 214, 229, *229*, *272*, 376, 970
Neuropeptides, 1098
see also individual neuropeptides
Neurosarcoidosis, 1159
Neurosis, 1189
Neurosurgery, 1119
epilepsy, 1143
stereotactic, 1148
Neurosyphilis, 1157
Neurotensin, *272*
Neurotoxicity, chemotherapy-induced, 462
Neurotransmitters, 1098
extrapyramidal system, *1113*
hormones, 965–6
Huntington's disease, *1113*
Parkinson's disease, *1113*
see also Neuropeptides; specific transmitters
Neutral endopeptidase, heart failure, 736
Neutropenia, *70*, *88*, 430–1
acquired, 69
causes, *430*
clinical features, 431
febrile, *461*
HIV-related, 189
investigations, 431
treatment, 431
viral, 431
see also Agranulocytosis; Neutrophil(s)
Neutrophil(s), 55–6, *56*
defects/dysfunction, *69*, 71
neutropenia *see* Neutropenia
extrinsic allergic alveolitis, 877
functions, 55
defective, *70*
granules
contents, *57*
function, *57*
release, 56
hypersegmented, *398*
late phase response, 72
measurement, *69*
migration, *68*
opsonization, 56, *57*
origins of, 388
phagocytosis, 56
see also Chronic granulomatous disease (CGD)
Neutrophil count, aplastic anaemia, 403
Neutrophil elastase, 358, 837
Neutrophil extracellular traps, 819
Neutrophil leucocytosis, 430, *430*
Neutrophilia, *88*
musculoskeletal problems, 503
Neutrophils, 430
function, 430
half-life, 430
host defence, 600
secretion of mediators, 819
Nevirapine, *193*, 195, *195*
side-effects, rash, 1269

New drug evaluation, 932, *932*
Nexin, 811
NFκB *see* Nuclear factor kappa B (NFκB)
Niacin, 222
  deficiency, *218*
  hyperlipidaemia, *1066*
  hyperphosphataemia, 633
  LRNI, *218*
  recommended IV intake, *234*
  RNI, *218*
  side-effects, hepatotoxicity, *366*
  sources of, 222
Niacin equivalent, *222*
Niacytin, 222
Nickel, *226*
  allergy, *848*
Nicorandil, angina pectoris, 750
Nicotinamide *see* Niacin
Nicotinamide adenine dinucleotide (NAD), 222
Nicotinamide adenine dinucleotide phosphate (NADP), 222
Nicotine
  and Parkinson's disease, 1146
  replacement therapy, 830
Nicotinic acid *see* Niacin
NIDDM *see* Diabetes mellitus
Niemann-Pick disease, *789*, 1071
Nifedipine, 941
  angina pectoris, 750
  cytochrome P450 induction, *925*
  hypertension, 805
  mountain sickness, 956
  side-effects, oedema, 656
Nifurtimox, 158
Night blindness *see* Xerophthalmia
Night sweats
Night terrors, 1198
Nikolsky's sign, 1256
Nimorazole, 99
  trichomoniasis, 182
Nipah virus, *110*, 119, 1156
Nitazoxanide, 162
Nitrates
  angina pectoris, 750
  heart failure, 739, *739*
  side-effects
    elderly patients, *924*
    postural hypotension, *659*
  vasodilation, 909
  see also Glyceryl trinitrate
Nitric oxide, 1098
Nitric oxide exhalation test, 826
  asthma, 853
Nitric oxide (NO), 20, 320, 351, 431, 896, 970
  ALI/ARDS, 919
  inducible, 897
  renal effects, 577–8
  sickle cell anaemia, 411
  synthesis and biochemical action, *897*
  vasomotor control, 686, *686*
Nitric oxide synthase, 577, 686, 896
  constitutive/endothelial, 896
  neuronal, 896–7
Nitrofurantoin
  side-effects
    haemolysis in G6PD deficiency, *414*
    hepatotoxicity, *366*
    neuropathy, *1174*
    respiratory, *877*
  urinary tract infections, 602
Nitrogen
  enteral nutrition, *235*
  excretion, 321
  negative balance, 210
  parenteral nutrition, 236
  urinary, 210
Nitrogen dioxides, 830, *830*
Nitrogen narcosis, 957
Nitrogen oxides, 830
Nitroglycerine, vasodilation, 909
Nitroimidazoles, 99, *99*
Nitrosoureas, 460
Nizatidine, 252
NKCC2, *658*
Nocardia, 149, *191*
  N. asteroides, *70*, 149
    pneumonia, *858*
  N. brasiliensis, 149
  in organ rejection, 75
Nocturia
Nocturnal angina, 749
Nocturnal hypoglycaemia, 1043
Nocturnal hypoxia, 839–41
  treatment, 840–1
Nocturnal myoclonus, 1150
Nodal osteoarthritis, 509
Nodular prurigo, 1240
Nodules
  intrapulmonary, 527
  Lisch, 1253
  skin, *1227*
  subcutaneous, rheumatoid arthritis, 527
  vocal cords, 1083, *1083*
Noise, 959
Noise trauma, 1079
Non-alcoholic fatty liver disease, 344–5, 1034
  diagnosis, 345
  histological changes, 344–5
  management, 345
  risk factors, 345
Non-disjunction, 38
Non-gonoccal urethritis *see* Urethritis
Non-haemolytic (febrile) transfusion reactions, 427
Non-Hodgkin's lymphoma (NHL), **479–84**
  aetiology, 479
  anaplastic, *480*
  B cell *see* B cell lymphoma
  Burkitt's lymphoma *see* Burkitt's lymphoma
  chemotherapy, *461*
  chromosome translocations, *480*
  classification, 480, *480*
  clinical features, 480
  cytogenetic features, 479–80, *480*
  epidemiology, *449*
  follicular, *480*, 481–2
  immunophenotypes, 480
  investigations, 480–1
  MALT, *480*
  mantle cell, *480*, 482
  pathogenesis, 479
  primary extranodal, 483–4
    cerebral, 483
    cutaneous, 483
    gastric, 266, 483
    in HIV disease, 483–4
  small intestine, 284
  T cell *see* T cell lymphoma
Non-insulin-dependent diabetes mellitus *see* Diabetes mellitus
Non-invasive ventilation, 916, *916*
Non-nucleoside reverse transcriptase inhibitors, 192, *193*, 195
  cytochrome P450 induction, *925*
Non-small-cell lung cancer (NSCLC) *see* Bronchial carcinoma
Non-starch polysaccharide, 211
  daily intake, 211
  deficiency, 211
Non-steroidal anti-inflammatory drugs *see* NSAIDs
Noonan syndrome, *789*
Noradrenaline (norepinephrine), 907, *908*, 1098
  cardiac contraction, 684, *684*
  see also Epinephrine (adrenaline)
Noradrenaline reuptake inhibitors (NaRIs), 1203
Norepinephrine *see* Noradrenaline (norepinephrine)
Norfloxacin, 98
Normal-transit constipation, 295–6, *296*
Normoblast, 389–90
Normocytic anaemia, *392*, 397
  pyogenic abscess, 361
Norovirus, *110*, 112
Northern blotting, 34
Nose, 814, *1080*, **1080–2**
  anatomy and physiology, 1080, *1080*
  anosmia, 1082
  bleeds/bleeding *see* Epistaxis (nosebleeds)
  blocked, 819
  examination, 1080
  foreign bodies, 1081
  fracture, 1082
  functions, 814
  Little's area, 1080, *1081*
  obstruction, 1081
  rhinorrhoea, 819, 1081
  secretions, 814, 819
  septal deviation, 1081
  sinusitis, 1081–2
  see also entries beginning Nasal
Nosebleeds *see* Epistaxis (nosebleeds)
Nosocomial diseases (hospital-acquired) *see* Hospital-acquired disease
NOTCH3 gene, 1127
Notifiable diseases, *83*
NSAIDs, *495*
  and coronary artery disease, 747
  gastropathy, 262–3
  gout, 538
  musculoskeletal problems, 517–18
  overdose, 946
  pain management, *see* Analgesia
  pericarditis, 795
  renal colic, 612
  rheumatoid arthritis, 529
  sarcoidosis, 556
  side-effects, 518
    acute tubulointerstitial nephritis, *604*
    asthma, 849
    bone marrow aplasia, 402
    elderly patients, *924*
    gastric bleeding, *267*
    hepatotoxicity, *366*
    hypertension, 799
    membranous glomerulonephropathy, 588
    rash, *1265*
    renal, *618*, *676*
    respiratory, *877*
  sodium retention, 656
  specific conditions
    rheumatic fever, 137
  teratogenicity, *924*
Nuclear factor of activated T cells, 65
Nuclear factor kappa B (NFkB), 24, 65
  Paget's disease, 565
Nuclear imaging/nuclear medicine *see* Radionuclide imaging
Nuclear transcription factor kappaB, 895
Nucleic, smooth muscle (actin), liver/kidney microsomal antibodies, *324*, 325
Nucleic acid amplification tests, 174–5
Nucleic acid detection, 89
Nucleoside analogues, 100–1, *101*
nucleotide-binding folds, *46*
Nucleotides, 29
Nucleus, *19*, **27–9**
Null hypothesis, 933
Numbness, 501, 1116
Nummular eczema, 1238
Nut allergy, 73
Nutcracker oesophagus, 254
Nutrient goals, 212
Nutrients
  see also specific dietary substances
Nutrition, **207–40**
  and ageing, 226–8, *227*
  dietary requirements, **208–12**
  enteral, 234–5
  parenteral, 235–6, *236*
  water and electrolyte balance, 207
  see also Diet; Food; Nutritional support
Nutritional deficiencies
  see also Malnutrition; Undernutrition
Nutritional support
  home patients, **236–7**
  hospital patients, **233–6**
    adult nutritional requirements, 234, *234*
    enteral nutrition, 234–5
    parenteral nutrition, 235–6, *236*
    principles of, 234
  refeeding syndrome, 216, 234
Nystagmus, 1079, 1105, 1108–9, 1114
  down-beat, 1109
  jerk, 1109
  pendular, 1109
Nystatin, 100, *100*
  candidiasis, 183

Obesity, **228–33**, 737, 1219
  adipostat, 228
  appetite control, 229–30, *229*
  Barrett's oesophagus in, 253
  benefits of weight reduction, *230*
  BMI in, 228
  body mass index, 230, *231*
  central, 1034
  clinical features, 230–1, *231*
  and coronary artery disease, 746
  disease association, *228*
  energy expenditure, 230
  food intake, 228
  genetic/environmental factors, 228
  heart sounds, 692
  and hypertension, 798
  metabolic syndrome, 231, *231*
  morbid, 232–3
  morbidity and mortality, 230, *230*, *231*
  and osteoarthritis, 520
  physical activity in, 230, 232
  prevention, 233
  thermogenesis, 230
  treatment
    behavioural modification, 232
    dietary control, 231–2
    drug therapy, 232
    surgical treatment, 232–3
      see also Diet
  twin studies, 228
  and type 2 diabetes, 1034
  see also Food
Obidoxime, *937*
Observational clinical trials
Obsessional rumination, 1187
Obsessive-compulsive disorder, 1210–11
  aetiology, 1211, *1211*
  clinical features, 1210–11
  prognosis, 1211
  treatment, 1211
Obstructive cardiac disease, coronary arteries (CAD), *see* Coronary atherosclerosis
Obstructive megaureter, 616
Obstructive shock, *894*, 904
  clinical features, 899
Obstructive sleep apnoea, 842, *842*
Occipital cortex, 1102
Occupational diseases
  cancers, 450
  lung *see* Occupational lung disease

musculoskeletal disorders, 502
repetitive strain injury, 516–17
Occupational lung disease, 878–80
asbestosis, 878–80, *879*
asthma, 848, *848*
berylliosis, 880
byssinosis, 880
coal-worker's pneumoconiosis, 878
silicosis, 878
Occupational sensitizers, 848, *848*
Octreotide, 283, 285
acromegaly, 980
insulinoma, 1060
pancreatic neuroendocrine tumours, 384
pituitary tumours, *973*
Ocular bobbing, 1124
Ocular movements
diabetic eye disease, 1052
unconscious patient, 1124
yoked, 1104–5
see also Nystagmus
Oculocerebrorenal dystrophy *see* Lowe's syndrome
Oculocutaneous albinism, 1264
Oculogyric crisis, 1148
Oculomotor nerve, 1105
Odds ratio, 931
Odynophagia, 249
Oedema, 655–6
cerebral *see* Cerebral oedema
cirrhosis, 655
extracellular volume regulation, 653
heart failure, 655
idiopathic, of women, 656
local increase, 656
minimal-change nephropathy, 586
nephrotic syndrome, 655–6
peripheral, 655, 689
tricuspid stenosis, 769
pulmonary *see* Pulmonary oedema
sodium retention, 656
treatment *see* Diuretics
Oenanthe, 952
Oesophageal disorders, 249
cricopharyngeal dysfunction, 255
diverticulum, 254–5
functional, 310–11
GORD *see* Gastro-oesophageal reflux disease (GORD)
investigation, 250, *250*
motility, 253–4
achalasia, 253–4
diffuse oesophageal spasm, 254
systemic sclerosis, 254
rings, 255
signs of, 250
strictures, 252, 254, 255
webs, 255
see also specific disorders
Oesophageal infections, 255
Oesophageal manometry, 250, *250*
achalasia, 254
Oesophageal perforation, 255–6
Oesophageal reflux, 245
Oesophageal rupture, 256
Oesophageal sphincters, 248, 250
transient relaxation, 249
Oesophageal tumours, 256–8, 489–90
cancer of oesophagus, 256–7, *257*, 279
epidemiology, *449*
prognostic factors, 489
treatment, 489–90
Oesophageal varices *see* Gastro-oesophageal variceal haemorrhage
Oesophagitis, 251, *251*
eosinophilic, 255
reflux, *267*
Oesophagogastroduodenoscopy, 246
Oesophagoscopy, 254
Oesophagus, **248–58**
antireflux mechanisms, 250, *250*
Barrett's *see* Barrett's oesophagus
corkscrew, 254
disorders of *see* Oesophageal disorders
mucosal defence mechanisms, 250–1
nerve supply, 248–9, *249*
nutcracker, 254
rupture, 884
structure and function, 248–9
swallowing, 248, 249
swan-neck deformity, 254, *254*
Oestradiol, *971*
PCOS, 1004
Oestrogen(s)
deficiency, 1001, *1001*
effects, *1001*
genetic polymorphism in metabolism, *926*, *927*
hirsutism, 1005
pituitary feedback, 996
replacement therapy *see* Hormone replacement therapy
side-effects
endocrine disease, *964*
hepatotoxicity, *366*
sodium retention, 656
Ofloxacin, 98
gonorrhoea, 176
leprosy, *141*
tuberculosis, 866
Oil acne, 1246
Oily cream, *1270*
Ointments, 1270
OKT3, *76*
Olanzapine
schizophrenia, 1218
side-effects, long QT syndrome, *727*
Oleander poisoning, antidotes, *937*
Oleic acid, *210*
Olestra, 212
Olfactory nerve, 1100
Oligoarthritis
extended, 553
persistent, 553
Oligodendrocytes, 47
Oligodendroglioma, 1160
Oligomenorrhoea, *995*
chronic renal failure, 630
Oligonucleotide primers, 34
Oligospermia, *995*, 1000
Oliguria, 619, 910
Olivo-ponto-cerebellar degeneration, 1148
Olopatadine, allergic conjunctivitis, 1089
Olsalazine (5-aminosalicylic acid) *see* 5-Aminosalicylic acid
Omalizumab, 72, 76, *76*
asthma, *854*, 856
Omega-3 acid ethyl esters, *1066*
Omega-3 acid marine triglycerides, *1066*
Omeprazole, 252, 269
cytochrome P450 induction, *925*
genetic polymorphism in metabolism, *926*
H. pylori eradication, 261
Omi protein, 28
Omnipaque, 244
Omsk haemorrhagic fever, *114*, 1156
Onchocerca volvulus, *82*, *163*, *164*
visual loss, 1094
Onchocerciasis, *82*, 165
clinical features, 165
diagnosis, 165
pathogenesis, 165
prevention, 165
treatment, *163*, 165
Oncogenes, 49
activation, 49
see also Cancer; specific genes
Oncogenesis, 49
Oncology *see* Cancer
Oncoviruses, *120*
Ondansetron, 461
Onychogryphosis, 1266
Onycholysis, *1242*, 1266
O'nyong-nyong fever virus, 113
arthritis in, 541
Operator proteins/regions, 31
Ophthalmia neonatorum, *83*
Ophthalmic emergency
Ophthalmoplegia
complete external, 1106
internuclear, 1105
Ophthalmoplegic migraine, 1137
Opiates
endogenous *see* Endogenous opioid peptides
misuse, 1179, 1215
overdose, *935*, 947
antidotes, *937*
pain relief, 495
side-effects, elderly patients, *924*
withdrawal syndrome, *1215*
see also Morphine
Opioid peptides, 1098
Opioids *see* Opiates
Opisthorchis
O. felineus, *163*, 170
O. viverrini, 170
cholangiocarcinoma, 364
treatment, *170*
Opisthotonus, 134
Opportunistic infection(s), 199, *199*
see also specific infections
Opportunistic organisms, 70
Opsoclonus, *455*
Opsonins, *57*
defects, *70*
Opsonization, 56, *57*, 430
Optic atrophy, 1101, 1102
Optic chiasm, 969, 1102
Optic disc
pallor *see* Optic atrophy
swelling, 1101–2
see also Papilloedema
Optic nerve, 1100–4
glioma, *1102*
lesions, 1101, *1101*
Optic neuritis, 1101–2
Optic neuropathy (ON)
ischaemic, *1102*
Leber's hereditary, 1102, *1102*
multiple sclerosis, 1152
toxic, *1102*
Optic papilla *see* Optic disc
Optic radiation, 1102
Optic tract, 1102
Oral cancer
epidemiology, *449*
geographical distribution, 451
Oral cavity *see* Mouth
Oral contraceptives, 965, 1005–6
and anticonvulsants, 1143
drug interactions, *1005*
fixed dose prescribing, *925*
mechanism of action, 1006
side-effects, *1006*
haemolytic uraemic syndrome, 598
hepatotoxicity, *366*
hypertension, 799
rash, *1265*
thyroid hormone levels, 984
venous thrombosis, 442
Oral hairy leukoplakia, 1269
Oral rehydration solutions (ORS), *132*, 144, 216
diarrhoea, 307
heat illness, 954
see also Fluid; Rehydration
Oral tolerance, 275
Oral ulceration, 246–7
Orexin, 1145
Orf, *104*, 109, 1232
Organ rejection, 75, *75*
classification, *75*
Organic acids
interstitial fluid, *650*
intracellular fluid, *650*
plasma, *650*
Organic anion transporter, 576
Organic anion transporting polypeptide 2, 321
Organic reaction, acute *see* Delirium
Organic toxins, *133*
Organophosphorus insecticides, *935*, 947
antidotes, *937*
long QT syndrome, *727*
Orientation, 1188
Orientia tsutsugamushi, *147*
Orlistat, 232
diabetes mellitus, 1039
and malabsorption, 282
Oromandibular dystonia, 1150
Oropharynx, 1082
Orotic aciduria, 397
Oroya fever, 148
Orphenadrine, poisoning, *935*
Orthomyxoviruses, *110*, 116–17, 834
Orthopnoea, 688, 820
mitral regurgitation, 763
Orthopox virus, *80*
Oseltamivir, 101–2, *101*
influenza, 834
Osgood-Schlatter disease, 513
Osler-Weber-Rendu disease, 435, *435*
and gastric bleeding, *267*
Osler's nodes, infective endocarditis, *770*, *771*
Osmoreceptors, 653, 654
Osmoregulation, 654
vs volume regulation, 654–5
Osmotic diarrhoea, 305–6
Osmotic diuresis, *664*
Osmotic laxatives, 296, *297*
Osmotic pressure, 649–50, *650*
Ossicles, 1077
Osteitis deformans (Paget's disease) *see* Paget's disease
Osteoarthritis, 500, **518–22**
aetiology, 518–19
causes, *520*
clinical features, 520
crystal-associated, 521, *522*
epidemiology, 518
erosive, 521
factors predisposing to, 519
first carpometacarpal, 509
hip, 512, 521
investigations, 521–2
joint distribution, *519*
knee, *519*, 521
management, 522
nodal, 509, 520–1, *520*, *521*
pathogenesis, 519–20
primary generalized, 521
surgery, 522
symptomatic, 552
Osteoblastic metastases, 452
Osteoblasts, 557
vitamin K-dependent proteins in, 219
Osteocalcin, 219, 557, 560
Osteochondritis
children, 554
dissecans, 514
Osteochondromatosis, 556
Osteoclasts, 58, 557–8
Osteocytes, 557
Osteogenesis imperfecta, *435*, 568
Osteogenic sarcoma, 565
Osteolytic metastases, 452, *452*

Osteomalacia, **567–8**, 629
  aetiology, 567, *567*
  clinical features, 567
  investigations, 567
  myopathy, 1179
  pathology, 567
  treatment, 567–8
  tumour-induced, 670
Osteomyelitis, *86*, 541
  sickle cell anaemia, 410
  treatment, 541
  tuberculous, 541
Osteonecrosis, 565
Osteopenia, 561
  diabetes mellitus, 1057
Osteopetrosis, 568
Osteophytes, 518, 519
Osteopontin, 557
  as tumour marker, *456*
Osteoporosis, **560–5**, 963
  associated disease, 562–3
  bone biopsy, *560*
  calcium intake, 228
  clinical features, 561–2
  crush fracture of spine, 512
  definition, 560–1
  fractures, 552, 561, *562*
  glucocorticoid-induced, 565
  incidence, 560–1
  investigations, 562–3, *562*
  in men, 565
  pathogenesis, 561
  post-renal transplant, 642
  prevention, 563
  risk assessment, 563
  risk factors, 561, *561*, 562–3
  treatment, 563–5
    drugs, 563–4, *564*
Osteoprotegerin, *452*, 557, 565
Osteosclerosis, 627
Otalgia, 1078
Otitis externa, 1078
Otitis media, 1078–9
  acute, 1079
  secretory, 1079, *1079*
Otorrhoea, 1078
Otosclerosis, *1078*, 1079
Outflow resistance *see* Afterload
Outreach care, 889
Ovarian cancer
  epithelial, 453, 490–1
    aetiology and pathology, 490–1, *490*
    epidemiology, *449*
    familial, *453*
    investigation, 491
    prognostic factors, 491
    screening, 454
    symptoms and signs, 491
    treatment, 491
  germ cell, 492
    histology, *492*
    investigations and surgery, 492
    symptoms and signs, 492
    treatment, 492
Ovarian failure, *964*, *965*, *1002*
Ovarian hyperstimulation syndrome, 1005
Ovarian hyperthecosis, 1003
Ovary
  cancer *see* Ovarian cancer
  cyst, 243, 316
  polycystic, *see* Polycystic ovary syndrome
Overflow incontinence, 614
Overlap syndromes, 548
Oxalate
  urinary, renal calculi, 613
Oxalate stones, 281
Oxaliplatin, 460
  colorectal carcinoma, *461*, 490
  gastric cancer, 490
  oesophageal cancer, 489
Oxalosis, *1070*
Oxalplatin, gastric cancer, 266
Oxaluria, 281
Oxantel pamoate, *166*
Oxazolidinones, 99, *99*, 124
Oxidative phosphorylation, 37
Oxidative stress injury, 345
Oxitropium
  asthma, *854*, 855
  COPD, 839
Oxygen, 912
  acute coronary syndromes, *754*
  administration
  see also Mechanical ventilation; Oxygenation
  arterial content, 890
  blood gas analysis *see* Arterial blood gases
  COPD, 839, *839*, *840*, 841, *841*
  cystic fibrosis, 845
  domiciliary, *840*
  haemoglobin binding *see* Haemoglobin
  hyperbaric, 890
  myocardial infarction, 757
  pulmonary hypertension, 783
  toxicity, 912
  see also Breathing; Respiration
Oxygen consumption, supranormal values, 908
Oxygen delivery, 890–3, *891*, 912
  face masks, 912, *912*
  nasal cannulae, 912, *912*
  supranormal values, 908
Oxygen narcosis, 957
Oxygen partial pressure, 391, 814, 890, 892
  nocturnal hypoxia, 839–40
  pancreatitis, *378*
Oxygen pressure, high altitude, *955*
Oxygen saturation, 890, 893
  arterial, 911
Oxygenation *see also* Mechanical ventilation; Ventilation
Oxyhaemoglobin dissociation curve, 391, 892, *892*
  blood transfusion affecting, 906
Oxymetazoline, rhinitis, 833
Oxytetracycline, acne vulgaris, 1246
Oxytocin, 970
Ozone, 830, *830*

p16 gene, 382
p53 gene, 47, 50, 382
  mutations, 454
PABA test, 376
Pacemakers/pacing *see* Cardiac pacemakers; Cardiac pacing
Pacing *see* Cardiac pacing
Packed cell volume (PCV), *389*, 826, 905
  COPD, 838
Paclitaxel, 24–5, 460
  breast cancer, *461*, 489
  cytochrome P450 induction, *925*
  drug-eluting stents, 751
  lung cancer, 486
  ovarian cancer, 491
  side-effects, neuropathy, *1174*
Paget's disease, 557, 565–6, 1107, *1255*
  aetiology, 565
  clinical features, 565, *566*
  investigations, 565
  pathogenesis, 565
  surgery, 566
  treatment, 566
Pain, 494–5, 1116, 1118–19
  abdominal *see* Abdominal pain
  back *see* Back pain
  bone *see* Bone pain
  chest *see* Chest pain
  chronic *see* Chronic pain syndromes
  diagnosis, 1118
  elbow, 508
  facial, 1138
  foot and heel, 515–16
  gate theory, 1118
  hand and wrist, 508–9
  hip, 512–13
  joint *see* Arthralgia; Joint(s)
  knee *see* Knee(s)
  loin *see* Loin pain
  management, 494–5, 1118–19
    see also Analgesia; Analgesics
  musculoskeletal *see* Musculoskeletal pain (myalgia)
  neck, 505–7
  physiology, 1118
  prostatic, 603
  psychological aspects, 1118
  shoulder, 507–8
  spinal root(s), 1116
  total, 494
  see also Headache(s)
Pain fibres, 1118
Pain/gas/bloat syndrome, 313
Painful arc syndrome, 507
Palivizumab, *101*, 102, 119
Palliative care, 494–7
  gastrointestinal symptoms, 495–6
  non-malignant disease, 496–7
  non-physical issues, 496
  pain, 494–5
  respiratory symptoms, 496
  symptom assessment, 494
Palmar erythema, 329
Palmitic acid, *210*
Palmoplantar keratoderma, *37*
Palpation
  abdomen, 243, *243*, *244*, 314
  chest, 820
    aortic stenosis, 766
    mitral stenosis, 761
Palpitations, 688–9, 722
  mitral valve prolapse, 764
Palsy
  Bell's *see* Bell's palsy
  bulbar, 1110–11, 1166
  cerebral, *1114*, 1169
  cerebral athetoid, 1149
  common peroneal nerve, 1172
  progressive supranuclear, 1148
  pseudobulbar, 1110–11, 1129, 1166
  'shaking' *see* Parkinson's disease
Pamaquin, side-effects, haemolysis in G6PD deficiency, *414*
Pamidronate, 552
  breast cancer, 489
  chronic regional pain syndrome type 1, 517
  hypercalcaemia, 458
  osteoporosis, 563
  pain relief, *495*
Pan-acinar emphysema, 836
Panaeolina, 952
Panaeolus, 952
Pancoast's tumour, 811, 1177
Pancreas, **375–85**
  acinar cells, 375, *375*
  alpha-cells, 376
  beta-cells, 376, *1030*
    genetic defects, *1035*
    insulin secretion, *1030*
  cancer *see* Pancreatic carcinomas
  D cells, 376
  endocrine, 376
  enterochromaffin cells, 376
  enzymes, 375
  exocrine, 375–6
  investigation, 376–7
    exocrine function, 376
    function tests, 376
    imaging, 376–7
  islet cells, 376
    autoantibodies, *1033*
    transplantation, 1043–4
  islets of Langerhans, 376, 384
  PP cells, 376
  Starling curve, 1035
  structure, 375
Pancreas transplantation, 1043–4
Pancreastatin, *272*
Pancreatic carcinomas, 382–4
  body/tail of pancreas, 383
  clinical picture, 382, *382*, *383*
  differential diagnosis, 383
  epidemiology, *449*
  head of pancreas, 383, *383*
  incidence, 382
  investigations, 383
  management, 383–4
  physical signs, 383
  risk factors, 382
  symptoms, 383, *383*
  see also Insulinomas
Pancreatic duct, 320, 375
Pancreatic insufficiency, tests for, *277*
Pancreatic intraepithelial neoplasias, 382
Pancreatic islet cells
  tumours *see* Insulinoma
Pancreatic polypeptide, *272*, 375
Pancreatic pseudocyst, 379, 381
Pancreatic supplements, 382
Pancreatic tumours
  cystic, 384
  neuroendocrine, 384–5
Pancreatitis, **377–82**
  acute, 377–9
    assessment of severity, 378–9
    clinical features, 377
    diagnosis, 377–8, *378*
    late complications, 379
    long-term outcome, 379
    pathogenesis, 377
    severe, *378*
    treatment, 379
  autoimmune chronic, 380
  causes, *377*
  chronic, 380–2
    aetiology, 380, *380*
    clinical features, 380
    complications, 381–2
    and diabetes, 1031
    differential diagnosis, 380–1, *381*
    investigations, 380
    pathogenesis, 380, *380*
    treatment, 381
  classification, 377
  gallstone-related, 379
  hereditary chronic, 380
  tropical chronic, 380
Pancytopenia, 402
  causes, *402*
  HIV-related, 189
  hypersplenism, 325
  see also Aplastic anaemia
Paneth cells, 274
Pangastritis, 260
Panhypopituitarism, 974
Panic attacks, 1145
Panic disorder, 1207–8
  aetiology, 1207–8
Panretinal photocoagulation, 1052
Pantoprazole, 252
Pantothenic acid, 223
  recommended IV intake, *234*
Papaverine, erectile dysfunction, 1000
Papillary necrosis, *605*
Papilloedema, 1101–2
  causes, *1102*
Papovaviruses, 108–9
Pappenheimer bodies, 423, *423*
Papules, *1227*
  Gottron's, 547, 1251
Parabens, 1270
Paracellin-1, *658*, 669
Paracentesis, ascites, 353
Paracetamol
  migraine, 1137
  musculoskeletal problems, 517, *518*

overdose and self-poisoning, 934, 947–8
antidotes, *937*
clinical features, 947–8, *947*
metabolic acidosis, *939*
treatment, 948, *948*
see also Fulminant hepatic failure (FHF)
sickle cell anaemia, *411*
side-effects
hepatotoxicity, 365, *366*
rash, *1265*
Paradoxical pulse, 690
Paraesthesia, 1116
Paraffin, 1270
Paragonimiasis, 170
Paragonimus, *163*
P. westermani, 170
treatment, *170*
Parainfluenza, *110*, 117
Parallel group studies, 929
Paralysis
diaphragm, 886–7
general of the insane, 1157
hyperkalaemic periodic, 667, 1182
hypokalaemic periodic, 666, 1178
Todd's, 1140
see also Palsy
Paralytic shellfish poisoning, 950
Paramedian pontine reticular formation, *1104*
Paramyoclonus multiplex, 1150
Paramyxoviruses, *110*, 117–19
Paraneoplastic polyarthritis, 555
Paraneoplastic syndromes, 455, *455*, 1177–8
Paraparesis, 1112, *1112*
causes, 1164–5
hereditary spastic, with ataxia, 1170
see also Paraplegia; Spinal cord diseases/lesions
Paraplegia, 1165–6
rehabilitation, 1166
see also Paraparesis
Paraproteinaemia, 484, 485
Paraquat, respiratory effects, *877*
Parasites/parasitic infections
eosinophilia, *431*
and malabsorption, 282
see also specific parasites
Parasympathetic innervation
Parathyroid antibodies, 1022
Parathyroid hormone (PTH), 559, 670, 1019, *1019*
bone disorders, 559
osteoporosis, 563, *564*
measurement, 1019
suppression, 633
see also Hyperparathyroidism; Hypoparathyroidism
Paratyphoid fever, *83*
Parental consent, 6
Parenteral nutrition, 235–6, *236*
administration and monitoring, 236
central venous catheter, 235–6, *235*, *236*
complications, 236
energy sources, 236
home, 237
nitrogen source, 236
nutrients, 236
peripheral, 235, *236*
regimens, *236*
Parietal cell antibodies, 399
Parietal cortex, lesions, 1118
Parinaud's syndrome, 1105
Parkinsonism, drug-induced, 1148
Parkinsonism-plus, 1148
Parkinson's disease, *1096*, 1113, **1145–8**
causes, 1145–6
diagnosis, 48, 1147
gait, *1096*, 1097
genetic factors, 1146
incidence, *1096*
natural history, 1146–7
neuropsychiatric aspects, 1148
neurotransmitters, *1113*
pathology, 1146
postural hypotension, *659*
prohibited drugs, *923*
stem cell therapy, 47
symptoms and signs, 1146
treatment, 1147–8, *1147*
drug therapy, 1147–8
physiotherapy and physical aids, 1148
stereotactic neurosurgery, 1148
tissue transplantation, 1148
see also Dopamine
Paronychia, 1234
drug-induced, 1269
Parotid gland, 1107
Parotitis, *86*
Paroxetine
genetic polymorphism in metabolism, *926*, *927*
overdose, 949
Paroxysmal cold haemoglobinuria, 417
Paroxysmal dyskinesia, 1145
Paroxysmal hemicrania, 1138
Paroxysmal nocturnal dyspnoea, 688
Paroxysmal nocturnal haemoglobinuria, 402, 418–19, 435
acute abdomen, *315*
clinical features, 418–19
investigations, 419
prognosis, 419
treatment, 419
Paroxysmal tachycardias, 689
Partial anomalous venous drainage, *774*
Partial pressure
carbon dioxide, 893
oxygen, *378*, 391, 890, 892
Particulates, 830, *830*
Passenger lymphocytes, 417
Passivity
Pastes, *1270*
Pasteurella multocida, skin and soft tissue infections, 124–5, *124*
Pasteurellosis, 124–5
Patau's syndrome *see* Trisomy 13
Patch tests, *1228*
Patellar tap sign, 514
Patent ductus arteriosus, *774*, 777
Patent foramen ovale, 775
Paterson-Brown-Kelly syndrome, 395
Pathogen sensing, 58
Pathogen-associated molecular patterns (PAMPs), 55, 58, *59*
septic shock, 895
Pathogens
see also Infections/infectious disease; specific organisms
Patient facilitation, medical interview, 12
Patient satisfaction, 9–10
Patient-centred communication, 8–9
Patient's complaints, 16
Patients' rights, 1–2
autonomy, 1–2, 3–6
confidentiality, 4–5
consent, 3–5
vulnerable patients, 5–6
competent refusal, 6–7
equality, 2, 7–8
high standard of clinical care, 1, 2
human rights issues, 2
lifestyle, 8
Pattern recognition receptors (PRRs), 55, 58, 59, *59*
Paul-Bunnell test, 107, 108
PC1, 229
Peak expiratory flow rate (PEFR), *816*, *824*, 825, *825*
asthma, 852
diurnal variation, *825*
Peak flow meter, *825*
Pearson's syndrome, 39
PECAM(s), *686*
PECAM-1, 67
Pectins, 211
Pectus carinatum, 884–5
Pectus excavatum, 884–5
Pediculosis
P. capitis, 1235
P. pubis, 183
Pegaptanib, age-related macular degeneration, 1094
Pegvisomant
acromegaly, 980
pituitary tumours, *973*
Peliosis hepatis, 365
Pellagra, 222, *222*, *789*
clinical features, 222, *222*
diagnosis and treatment, 222
Pelvic examination
Pelvic floor dyssynergia, 296
Pelvic tumours, 645
Pelviureteric colic, acute abdomen, *315*
Pelviureteric junction obstruction, 616, *616*
Pemphigoid
gestationis, 1270
Pemphigus, *455*
drug-induced, *1265*
Pemphigus vulgaris, 26, 1255–6
autoantigens, *74*
clinical features, 1255–6
HLA system in, *60*
treatment, 1256
Penciclovir, 101
Pencil in cup appearance, 535, *535*
Pendred's syndrome, 985, *1078*
Pendrin, 985
Pendular nystagmus, 1109
Penicillamine
D-Penicillamine, *937*
cystine stones, 613
rheumatoid arthritis, 532
side-effects, 613
membranous glomerulopathy, 588
rash, *1265*
systemic lupus erythematosus, 542
vasculitis, 596
Wilson's disease, 357
Penicillin(s), *94*, **94–5**
allergy, 73
gonococcal arthritis, 540
indications for use, 94–5
infective endocarditis, *773*
interactions, 95
mechanism of action, 94
meningitis, *1155*
prophylaxis, 406
side-effects
acute tubulointerstitial nephritis, *604*
hepatotoxicity, *366*
immune haemolytic anaemia, 417
rash, *1265*
respiratory, *877*
structure, *94*
toxicity, 95
see also individual drugs
Penicillin G *see* Benzylpenicillin
Penicillin V *see* Phenoxymethylpenicillin (penicillin V)
Penoxymethylpenicillin, 95
Pentamidine
nephrotoxicity, 189
side-effects, hypoglycaemia, 1060
Pentosephosphate pathway *see* Hexose monophosphate shunt
Pentostatin, hairy cell leukaemia, 475
Pentosuria, 1069
Pentraxins, *56*
Pepperpot skull, 628
Peptic structure, 252
Peptic ulcer disease
chronic, 269
clinical features, 260
complications, 261–2
epidemiology, 260
Helicobacter pylori infections *see* *Helicobacter pylori*
investigation, 260–1
surgery, 262
upper GI tract bleeding, 267
see also Duodenal ulcers; Gastric ulcers
Peptide YY, 229, *272*, 276
Peptide-histidine methionine, *272*
Peptidoglycan, *59*, 86
Peptococci, *123*
Peptostreptococci, *123*
Per protocol analysis, 930
Perceptions, abnormal, 1188
Percussion
abdomen, 243
chest, 820
Percutaneous coronary intervention
angina pectoris, 750–1, *751*
bleeding after, 269
myocardial infarction, 757–8
Percutaneous endoscopic gastrostomy, 215
Percutaneous transhepatic cholangiography (PTC), 327
Percutaneous transluminal angioplasty, 806
Perennial rhinitis, 819, 831–2, *831*
allergic, 831–2, *831*
non-allergic with eosinophilia, 832
Perforin, 66
Perglycan, 500
Pergolide, side-effects, retroperitoneal fibrosis, 616
Peri-partum cardiomyopathy, 794
Periaortitis, 318
chronic *see* Retroperitoneal fibrosis
Pericardial aspiration *see* Pericardiocentesis
Pericardial biopsy, 796
Pericardial calcification, 693
Pericardial disease, **794–7**
magnetic resonance imaging, 707, *707*
see also Pericardial effusion; Pericarditis
Pericardial effusion, 655, 795, 796
echocardiography, 704, *704*
heart size, 692
heart sounds, 692
treatment, removal *see* Pericardiocentesis
Pericardiocentesis, 714, *714*, 796
Pericarditis, *86*, *111*, 794–6, *795*
aetiology, *795*
bacterial, 795
chronic renal failure, 631
clinical features, 795
constrictive, 797, *797*
see also Calcification
fungal, 795
investigations, 795, *796*
malignant, 795
pain, 688
post-myocardial infarction, 794
systemic lupus erythematosus, 543
treatment, 795–6
tuberculous, 795
uraemic, 794
viral, 794

Peridiverticulitis, 298
Perigolide, side-effects, constrictive pericarditis, 797
Perimysium, 501
Periodic breathing, 688
Periodic leg movement disorder, 1198
Periodic paralysis
  hyperkalaemia, 667, 1182
  hypokalaemia, 667, 1178
Periodontal disease, HIV-related, 204
Perioperative multiorgan failure, 908, *909*
Perioral dermatitis, 1247
Peripancreatic fluid, 379
Peripheral nerve disease, 1116, **1170–1**
  HIV-related, 1158
  mechanisms of damage, 1170–1
  see also Neuropathies
Peripheral nerves, 1115
  biopsy, 1122
  compression, 1171–2, *1171*
  conduction, 1121, *1121*
  disease *see* Peripheral nerve disease
  entrapment, 1171–2, *1171*
  infarction, 1171
  infiltration, 1171
  regeneration, 1171
Peripheral nervous system (PNS)
  chronic renal failure, 630
  nerves *see* Peripheral nerves
Peripheral sensory neuropathy, *455*
Peripheral tolerance *see* Tolerance
Peripheral vascular disease, **805–10**
  arterial, 805–9
  venous, 809–10
Peripheral vasoconstriction, 659
Perisplenitis, 421
Peristalsis
  colonic, 294–5
  oesophageal, 249
Peritoneal cavity/peritoneum
  dialysis *see* Peritoneal dialysis
  diseases, inflammation, *see* Peritonitis
  fluid, *see* Ascites
Peritoneal dialysis, 636–9, *636*, *637*
  acute tubular necrosis, 624
  adequacy of, 638
  complications, 637–8
    catheter site infection, 638
    peritonitis, 637, *637*
  continuous ambulatory (CAPD) *see* Continuous ambulatory peritoneal dialysis (CAPD)
  contraindications, 638
  nightly intermittent, 537
  Tenckhoff catheter, *637*
  tidal, 537
  see also Continuous ambulatory peritoneal dialysis (CAPD)
Peritoneum, 317–18, *317*
Peritonitis, 316, 317
  CAPD-related, 637, *637*
  clinical features and management, 316
  complications, 316
  generalized, 316
  localized, 316
  spontaneous bacterial, 353
  tuberculous, 318
Permethrin, 183, 1235
Permissive hypercarbia, 916
Pernicious anaemia (PA), 263, 399–400, *965*
  associations, 399
  autoantigens, *74*
  clinical features, 399
  differential diagnosis, 400
  investigations, 399–400
  parietal cell antibodies, 399
  pathogenesis, 399
  pathology, 399
Peroneal muscular atrophy *see* Charcot-Marie-Tooth disease
Peroxidase, 850
Peroxisomes, 23
Perseveration, 1187
Persistent generalized lymphadenopathy, 187
Persistent vegetative state, *1123*
Personal failures, clinical communication, 11
Personality change, *1099*
Personality disorders, 1222
Personalized prescribing, 926, 927
Perthes' disease, 554
Pertussis, *83*, 127
  clinical features, 127
  diagnosis, 127
  immunization, *103*
  prevention and control, 127
  treatment, 127
Pertussis (whooping cough), 819
  mortality, *80*
Petechiae, 434, *1227*
  infective endocarditis, *770*
  see also Rashes
Pethidine, pancreatitis, 379
Petrous temporal bone, 1106, 1107, 1109
Peutz-Jeghers syndrome, 247, 285, *301*, 302, 1265
Peyer's patches, 53, 85, *274*, *275*
Peyronie's disease, 509
pH
  and H+ concentration, *672*
  oesophageal, 250, *251*
  see also Acidosis; Acid(s); Alkalosis
Phacoemulsification, 1091
Phaeochromocytoma, 222, 1014, 1025, 1145
  clinical features, 1025, *1025*
    diabetes, *1031*
  diagnosis, 1025
  pathology, 1025
  treatment, 1025
Phagocytosis, 21, *416*, 818–19
  defects, 70
  spleen, 422
  see also Macrophage(s); Monocytes; Neutrophil(s)
Phagosomes, 21
Phalen's test, 508
Phallotoxins, *133*
Pharmacodynamics, 926–7, *927*
Pharmacokinetics, 925–6, *925*, *926*
  genetic alterations, 925–6, *926*
Pharyngeal pouch, 1084, *1085*
Pharyngitis, 833, 1083–4
Pharynx, 248, 1082
  foreign body, 1084
  see also Oesophagus
Phenacetin, side-effects, haemolysis in G6PD deficiency, *414*
Phenobarbital, epilepsy, *1143*
Phenolphthalein, side-effects, rash, *1265*
Phenophthalein, *297*
Phenothiazines, 1217
  poisoning, *935*
  side-effects
    hepatotoxicity, 365
    long QT syndrome, *727*
Phenotype, maternal effect, 51
Phenotypic sex, 1008
Phenoxyacetate, poisoning, *935*
Phenoxymethylpenicillin (penicillin V), *93*, 126
  atopic eczema, 1238
  cellulitis, 1228
  echthyma, 1228
  impetigo, 1228
  tonsilitis, 833
Phentermine, side-effects, respiratory, *877*
Phentolamine, erectile dysfunction, 1000
Phenylalanine, 210
Phenylalanine hydroxylase deficiency, *1070*
Phenylhydrazine, side-effects, haemolysis in G6PD deficiency, *414*
Phenylketonuria, 46, 237, 1070
Phenytoin
  cytochrome P450 induction, *925*
  epilepsy, *1143*
    dose, *1143*
    therapeutic range, *1143*
  genetic polymorphism in metabolism, *926*
  overdose, *935*, 940
  side-effects, *928*, *1143*
    acute tubulointerstitial nephritis, *604*
    bone marrow aplasia, 402
    hepatotoxicity, *366*
    hypertrichosis, 1003
    neuropathy, *1174*
    respiratory, *877*
  teratogenesis, *924*
  therapeutic drug monitoring, *927*
  trigeminal neuralgia, 1107
Phialophora, 153
Philadelphia chromosome, *41*
  leukaemia, 467, *467*, *472*
  myelofibrosis, 421
Phlebotomus, 159
Phobias, 1208, *1208*
Phonation, 1082
Phosphatase and tensin homolog (PTEN), *301*, 302
Phosphate, 226, 670
  concentration disorders *see* Hyperphosphataemia; Hypophosphataemia
  deficiency, 226
  dietary restriction, 633
  interstitial fluid, *650*
  intracellular fluid, *650*
  plasma/serum, *650*
    bone disorders, 559, *559*
    rickets/osteomalacia, 567
  RNI, 226
Phosphate binders, 633
Phosphatonin, 670, *671*
Phosphide poisoning, 948
Phosphodiesterase inhibitors, 907, 1000
Phospholamban, 684
Phospholipase(s), 375
Phospholipase A2, 274
Phospholipids, 321
  see also Lipid(s)
Phosphoric acid, 674
Phosphorylation of proteins, 24
Photocoagulation
Photodermatology, 1248–9
Photodermatoses, idiopathic, 1248–9
Photophobia
Photoprotection, 1248
Photosensitive epilepsy, 1141
Photosensitive rashes, 1248
  differential diagnosis, *1248*
  drug-induced, *1265*
Photosensitizers, 1248
Phototherapy, 1248
  haemolytic disease of newborn, 418
Phthiriasis pubis, 183, 1235
Physical abuse *see* Abuse
Physical activity *see* Exercise
Physical activity ratio, 208, *208*
  activities, **209**
Physical diseases/disorders
Physiotherapy, 1133
Phytanic acid, 1122
Phytomenadione, 440
  see also Vitamin K
Phytophotodermatitis, *1248*
Pick's bodies, 1168
Picornaviruses, 109–12, *110*
  coxsackievirus, 111–12
  echovirus, 111–12
  enterovirus, 111–12
  poliomyelitis, 109–11
  rhinovirus, 112
PIG-A gene, 418
Pigmentation
  disorders, 1264–5
    drug-induced, *1265*
    see also Hyperpigmentation; Hypopigmentation
Pigmented villonodular synovitis, 556
Pilar cyst, 1257
Pill-rolling, 1114
Pimecrolimus, atopic eczema, 1238
Pinealoma, 974
Pink puffers, 835
Pinocytosis, 21, 59
Pins and needles, 501
Pinta, 142–3
Pioglitazone, 1039
Piperaquine, *156*
Piperazine, *166*
  threadworm, *166*, 167
Piroxicam, side-effects, hepatotoxicity, *366*
Pistol shot femorals, 767
Pituitary apoplexy, 975
Pituitary disease, **969–76**
  hypersecretion syndromes, 979–82
  presentation, 970, 972
  see also Hypopituitarism; Pituitary tumours
Pituitary fossa, 969, *969*
Pituitary function tests
  acromegaly, 979
  hypopituitarism, 981
Pituitary gland
  anatomy, 969–70, *969*
  anterior, physiology, 970
  posterior, physiology, 970
Pituitary hormones, *971*
  see also Hypopituitarism; specific hormones
Pituitary tumours and space-occupying lesions, 972–4, *972*
  differential diagnosis, 974
  incidentaloma, 972
  investigations, 972–3
  serum prolactin, *981*
  treatment, 973–4, *973*
  see also Prolactinoma
Pityriasis rosea, 1244
Pityriasis versicolor, 1234
Pityrosporum
  P. orbiculare, 1234
  P. ovale, 1234
Pivmecillinam, 95
Pizotifen, migraine, 1137
Placental growth factor, 597–8
Plague, *80*, *82*, *83*, 146
  as biological agent, *959*, 960
  clinical features, 146
  diagnosis, 146
  management, 146
  prevention, 146
Plakophillin-2, 791
Plane warts, 1231
Planned discharge, 890
Plant poisons, 952
Plantar fasciitis, 515, *515*
Plantar spur, *515*
Plantar spurs, 515, *515*
Plantar warts, 1231
Plaques, *1227*
  atheromatous *see* Atherosclerotic plaques
Plasma, 387, 649
  drug concentration
  electrolyte composition, *650*
  fresh frozen *see* Fresh frozen plasma

loss of, 660
osmolality *see* Plasma osmolality
proteins, 321
transfusion incompatibility, 427–8
viscosity, 389
volume regulation, 649
Plasma cell membrane *see* cell membrane
Plasma cells, 63
Plasma exchange, 437
ANCA-positive vasculitis, 597
anti-GBM glomerulonephritis, 596
Plasma expanders, 660
Plasma osmolality, 654
Plasmacytoid dendritic cells, *56*, 58, *59*
Plasmacytoma, soft tissue, 484
Plasmapheresis
lymphoplasmacytic lymphoma, 481
myasthenia gravis, 1180
Plasmids, 35
Plasmin, 434, 898
Plasminogen, 434, 898
activators, 431
see also Tissue-type plasminogen activator
Plasminogen activator-inhibitor-1, *434*, 631
in diabetes, 1034
pulmonary hypertension, 782
Plasmodium, *82*, 85
P. falciparum, 153, 154, 155, *155*
protective effect of sickle cell trait, 411
treatment, *156*
P. malariae, 153, 154, 155
treatment, *156*
P. ovale, 153, 154, 155
treatment, *156*
P. vivax, 153, 154, 155
life cycle, *154*
resistance to, 155
treatment, *156*
Plateau pulse, 690
Platelet activating factor acetylhydrolase, 73
Platelet adhesion, 432
Platelet aggregation, 432, *432*
Platelet concentrates, 429
Platelet count, *389*, 393, *435*
aplastic anaemia, 403
chronic myeloid leukaemia, 438
essential thrombocythaemia, 438
myelofibrosis, 421, 438
polycythaemia vera, 438
post-splenectomy, 423, *423*
Platelet factors, 432
Platelet release, 432
Platelet-activating factor, 57, *58*, *68*
in anaphylaxis, 73
in shock response, 895
Platelet-derived growth factors (PDGFs), 345, 432
asthma, 852
Platelet-endothelial cell adhesion molecule-1 *see* PECAM-1
Plateletpheresis, 429
Platelets, 387
antibodies, 436
disorders, 435–8
acquired, 437
functional, 437
inherited, 437
thrombocytopenia *see* Thrombocytopenia
thrombocytosis, 437–8
thrombotic thrombocytopenic purpura, 437
see also Thrombosis
in haeostasis, 432
origins of, 388
survival, 387
transfusion incompatibility, 427
see also Coagulation
Platinum, allergy, *848*
Platinum compounds, 460
Pleiotropy, 54, 55
Pleura, 812
parietal, 812
visceral, 812
Pleural aspiration, 827
Pleural biopsy, 827, *827*, *828*
Pleural effusion, 655, 885–6, *885*
chest X-ray, 822, *885*
diagnosis, 885
exudates, 885
malignant, 886
physical signs, *821*
transudates, 885
Pleural rub, 821
Pleurisy, 885
Pleuropericardial cysts, 888
Plexogenic pulmonary arteriopathy, 783
Plexus lesions, 1176–7, *1176*
Plummer-Vinson syndrome, 255, 395
Pluripotent stem cells *see* under Stem cells
Pluripotentiality, 47
Pluteus, 952
PML-RAR gene, *453*
Pneumatosis cystoides intestinalis, 298
Pneumococcus *see* *Streptococcus*, *S. pneumoniae*
Pneumoconiosis, 822, 878
Pneumocystis
in organ rejection, 75
P. jiroveci, 70, *70*, *90*, *150*, 153, *191*
HIV-associated, 200
pneumonia, *858*, 861
Pneumocytes
type I, 812, *812*
type II, 812
surfactant secretion, 817
Pneumomediastinum, 914
Pneumonia, *86*, *822*, **857–63**
acute abdomen, *315*
acute interstitial, 873
aetiology, *858*
aspiration, 861–2
classification, 857
clinical features, 858
community-acquired, 862
complications, 863
cryptogenic organizing, 873
desquamative interstitial, 873
diffuse (bronchopneumonia), 861
hospital-acquired, 862–3
idiopathic interstitial, 872–4
investigations, 858–9, *858*
lymphoid interstitial, 873–4
management, *859*, 862–3
mycoplasma, 859
non-specific interstitial, 873
opportunistic infections, 861
organisms
Chlamydia pneumoniae, 860
Chlamydia psittaci, 860
Coxiella burnetii (Q fever), 860
Haemophilus influenzae, 860
Klebsiella pneumoniae, 861
Legionella pneumophila, 860
Moraxella catarrhalis, 861
Pneumocystis jiroveci, 861
Pseudomonas aeruginosa, 861
Staphylococcus aureus, 860
precipitating factors, 857–8
rare causes, 862
returning travellers, *90*
ventilator-associated, 914
viral, 859–60
Pneumonic plague, 146
Pneumoperitoneum, 914
Pneumothorax, 886, 914
aspiration, *886*
diving, 957
management, *887*
physical signs, *821*
tension, 914
Podocin, 583
Podocytes, 571, 583
Podophyllotoxin, human papillomavirus, 182
POEMS syndrome, 1176
Poikilocytosis, *394*, 396
myelofibrosis, 421
Point mutations, 40–1, *453*
Point source, 83
Poisoning, **933–52**
management, *935*, 936–7
antidotes, 937, **937**
increasing elimination, 938–9
reducing absorption, 937–8
metabolic acidosis, *939*
nature of, 933–4
non-toxicological investigations, 939, *939*
paracetamol *see* Paracetamol
physical signs, *935*
specific poisons
marine animals, 950
mushrooms, 952
plants, 952
venomous animals, 951–2
toxicological investigations, 939
unconscious patients, 935–7
cardiovascular support, 935–6
ventilatory support, 935
see also specific poisons, drugs and chemicals
Poliomyelitis, 109–11, *1110*
acute, *83*
clinical features, 109–11
abortive poliomyelitis, 109
bulbar poliomyelitis, 110–11
inapparent infection, 109
non-paralytic poliomyelitis, 110
paralytic poliomyelitis, 110
diagnosis, 111
immunization, *103*
prevention and control, 111
treatment, 111
Polyaromatic hydrocarbons, 830
Polyarteritis nodosa, *549*, 550–1
clinical features, 550
investigations, 550–1
renal involvement, 598
treatment, 551
Polyarthritis
seronegative symmetrical, 534
see also Arthritis
Polychromasia, 403, *405*
Polycystic disease of liver, 360
Polycystic kidney/renal disease (PKD), 42, 642–4
autosomal-dominant *see* Autosomal-dominant polycystic kidney disease (ADPKD)
medullary cystic disease (juvenile nephronophthisis), 644
medullary sponge kidney, 644
renal calculi, 611
Polycystic ovary syndrome, 963, 1003–5
clinical features, 1003–4
diagnosis, 1005
hirsutism, 1003
investigations and differential diagnosis, 1004–5, *1004*
serum prolactin, *981*
treatment, 1005
Polycystic ovary syndrome (PCOS), 1001
Polycythaemia, 390, 419–21, *455*
Chuvasch, 420
congenital heart disease, 774
COPD, 840
Fallot's tetralogy, 778–9
primary (polycythaemia vera), 419–20
acute abdomen, *315*
clinical features, 420
course and management, 420
diagnosis, 420, *420*
prognosis, 420
relative/apparent (Gaisböck's syndrome), 421
secondary, 420–1, *421*
causes, *419*, *421*
thrombocythaemia, 421
Polydipsia, 1016
primary, 1018
Polyenes, 100, *100*
Polyethylene glycols, 296, *297*, 1270
Polygelatin solutions, 906
Polygenic hypercholesterolaemia, 1064
Polygenic inheritance, 44, *44*
Polyglandular autoimmune syndromes, 1026
Polymerase chain reaction (PCR), 34–5, *34*, *35*, *772*
expression microarrays/gene chips, 34–5
real-time, 34
Polymorphic eruption of pregnancy, 1269
Polymorphic light eruption, 1248–9
Polymorphisms, 35–6, 925–6, *925*, *926*
Polymorphs *see* Neutrophil(s)
Polymyalgia rheumatica, 501, 549–50, *549*
elderly patients, 552
hip pain, 513
investigations, 550
pyrexia, *91*
Polymyositis/dermatomyositis, *455*, **547–9**, 1179, 1251
antinuclear antibodies, *504*
association with malignancies, 547
autoantigens, *74*
clinical features, 547
adult dermatomyositis, 547
adult polymyositis, 547
antisynthetase syndrome, 547
childhood dermatomyositis, 547
differential diagnosis, 1179
investigations, 547–8
malignancy-related, *1255*
rheumatoid arthritis, *503*
symptom pattern, *549*
treatment, 548, 1179
Polymyxin, 128
Polyneuropathies, 1071, 1172–6, *1173*
cranial, 1173
critical illness, 916, 1176
HIV-associated, 1176
sensorimotor
symmetrical mainly sensory, 1055
Polyneuropathy, *1110*, *1117*
Polyomavirus BK, *104*
Polyomavirus JC, *104*, *191*
HIV-related, 203
Polypeptides, metabolism, 576
Polyposis
familial adenomatous *see* Familial adenomatous polyposis
MYH-associated, 302
Polyps
colorectal, 300–2, *301*
see also Familial adenomatous polyposis
gall bladder, 373
gastric, 266
nasal, 1081
Polysomes, 30
Polyunsaturated fatty acids (PUFAs), 210, 211, 1065
benefits of, 211
GISSI Prevention Trial, 211
n-3, 211
n-6, 211
and vitamin E requirement, 220
Polyuria, 614, 1036

Polyvinyl chloride, 959
  skin changes, 1252
POMC, 229
POMC, 1008
POMC/CART, 229
Pompe's disease, 788, *789*, *1069*
Pompholyx, 1238, *1238*
Poncet's disease, 540
Pons, 1107, 1109
  hemiparesis, 1112
Pontine lesions, 1118
Popcorn cells, 476
Popliteal artery entrapment, 806
Popliteal (Baker's) cyst, 514
  rupture, 514–15
Population genetics, 44
Pores of Kohn, 812
Porphyria(s), 1073–5
  acute intermittent, 1073–4, *1074*
    acute abdomen, *315*
    investigations, 1074
    neuropathy, 1174
    screening, 1074
  clinical features, 1073, *1074*
  congenital erythropoietic, *1074*, 1075
  cutanea tarda *see* Porphyria cutanea tarda (cutaneous hepatic porphyria)
  erythropoietic protoporphyria, *1074*, 1075
  hepatoerythropoietic, *1074*, 1075
  hereditary coproporphyria, 1074, *1074*
  management, 1075
  variegate, 1074, *1074*
Porphyria cutanea tarda (cutaneous hepatic porphyria), 1074–5, *1074*, 1254
  chronic renal failure, 629
Porphyrins, 1073
  metabolism, *1074*
Port-wine stain, 1268
Porta hepatis, 320
Portal gastropathy, 262
Portal hypertension, 348–9, *349*
  ascites, 351
  causes, 349, *349*
    posthepatic, 349
    prehepatic, 349
  clinical features, 349
  intrahepatic, 348
  non-cirrhotic, 349
  pathophysiology, 349
  portosystemic encephalopathy *see* Portosystemic encephalopathy (PSE)
  posthepatic, 348
  prehepatic, 348
  in schistosomiasis, 349
  splenomegaly, 329, 423
Portal vein, 320, *320*
  hypertension *see* Portal hypertension
  thrombosis, 349, 360
Porto-pulmonary hypertension, 354
Portosystemic encephalopathy (PSE), 353–4
  clinical features, 353–4
  course, 354
  diagnosis, 354
  investigations, 354
  management, 354
  pathogenesis, 353, *353*
  prevention, 350–1
  prognosis, 354
Portosystemic shunts, 351
Posaconazole, 100, *100*
Positive end-expiratory pressure (PEEP), 912, 915
Positive selection *see* Tolerance
Positive-pressure ventilation, COPD, 840
Positron emission tomography (PET)
  abdominal, 244
  brain, 1120
  carcinoma of oesophagus, 257
  gastric cancer, 265
  heart, 709
  Hodgkin's lymphoma, 477, *478*
  lungs, 823–4
  musculoskeletal problems, 505
  see also Radionuclide imaging
Post-herpetic neuralgia, 106
Post-infective neuropathy *see* Guillain-Barré syndrome
Post-phlebitic syndrome, 442
Post-traumatic stress disorder (PTSD), 1209–10
  clinical features, 1210
  treatment and prevention, 1210
Postherpetic neuralgia, 1157
Postpartum psychosis, 1199
Postprandial distress syndrome, 311
Postprandial hypoglycaemia, 1060
Postprandial hypotension, 689
Postprandial insulin effects, *1032*
Postural drainage, 843
Postural hypertension, 1144
Postural hypotension, 659, *659*, 689
Postural orthostatic tachycardia syndrome, 717
Postural tremor, 1114
Posture
  Parkinson's disease, 1146
  simian, 1146
Postviral/chronic fatigue syndrome, 121
Potassium
  concentration disorders *see* Hyperkalaemia; Hypokalaemia
  deficiency
  dialysate
    haemodialysis, *634*
    peritoneal dialysis, *637*
  excretion, 664–5
  input, *295*
  interstitial fluid, *650*
  intracellular fluid, 649, *650*
  losses from gut, *662*
  output, *295*
  plasma/serum, *650*, 664–5
  uptake, 664
  urinary loss, 1024
  see also Hyperkalaemia; Hypokalaemia
Potassium chloride, *660*
Pott's disease, 541
Pouchitis, 292–3, *293*
Power calculations, 933
Poxviruses, 109
  cowpox, 109
  monkeypox, 109
  smallpox, 109
  vaccinia virus, 109
Prader-Willi syndrome, 39, 44, *228*, *1035*
Pralidoxime, *937*
Pralidoxime mesilate, 947
Pramipexole, Parkinson's disease, 1147
Pranlukast, asthma, 856
Prasugrel, arterial thrombosis, 443, *443*
Pravastatin
  acute coronary syndromes, *754*
  hyperlipidaemia, 746, *1066*
Praziquantel
  cysticercosis, 171
  schistosomiasis, *163*, 361
  trematode infections, *170*
Pre-eclampsia, 362–3, 799
  glomerulopathy, 597–8
Pre-syncope, 726
Prebiotics, irritable bowel syndrome, 312–13
Precocious puberty, 996–7
Precordium, examination of, 692
Prednisolone
  Addison's disease, *1012*
  allergic bronchopulmonary aspergillosis, 876
  ANCA-positive vasculitis, 597
  asthma, *854*
  autoimmune hepatitis, 344
  Bell's palsy, 1108
  bullous pemphigoid, 1256
  chronic discoid lupus erythematosus, 1252
  COPD, 839
  diffuse large B cell lymphoma, 482
  erythema nodosum, 1250
  focal segmental glomerulosclerosis, 587
  giant cell arteritis, 550
  Hodgkin's lymphoma, *461*
  idiopathic pulmonary fibrosis, 873
  IgA nephropathy, 594
  infertility, 1005
  juvenile idiopathic arthritis, 554
  membranous glomerulonephropathy, 588
  microscopic colitis, 294
  non-Hodgkin's lymphoma, *461*
  pemphigus vulgaris, 1256
  polymyositis/dermatomyositis, 548
  potency, *1008*
  primary restrictive non-hypertrophic cardiomyopathy, 794
  rhinitis, 833
  sarcoidosis, 870
  systemic sclerosis, 547
  thrombocytic purpura, 436
  thyroid eye disease, 992
  trypanosomiasis, 158
  ulcerative colitis, 292
Pregabalin, *495*
  chronic widespread pain, 516
  diabetic neuropathy, 1055
  overdose, 940
Pregnancy
  alcohol consumption in, 239
  antidepressants in, 1203
  bacteriuria, 603
  diabetes mellitus, 1058–9
    general management, 1058
    gestational diabetes, 1058–9
    metabolic control, 1058
    neonatal problems, 1058
    obstetric problems, 1058
  dietary requirements
  eclampsia, 362–3
  ectopic, ruptured, 315–16
  epilepsy, 1142–3
  genital herpes, 181
  HIV infection, 198–9
  hyperemesis gravidarum, 363
  hypertension, 799
    management, 805
    see also Eclampsia; Pre-eclampsia
  hyperthyroidism, 990
  inflammatory bowel disease in, 293
  liver disease, 362–3
    acute fatty liver of pregnancy, 363
    intrahepatic cholestasis, 362
  low back pain, 510
  Marfan's syndrome, 781
  nausea and vomiting, 242
  pre-eclampsia, 362–3
  prenatal diagnosis *see* Prenatal diagnosis
  prescribing in, 924
  serum prolactin, *981*
  sickle cell anaemia, 410
  skin diseases, 1269–70, *1269*
  systemic lupus erythematosus, 544
  teratogenesis *see* Teratogens
  thyroid hormones in, 984
Pregnancy-associated plasma protein-A, 45
Preload, 891, *891*
  myocardial dysfunction, 734
  optimization, 904–5
Premature ageing syndrome, 37
Premature beats, 689, 690
Premature infants
  prescribing in, 924
Premenstrual syndrome, 1194–5
Prenatal diagnosis, 45
  see also Genetic counselling; Genetic screening
Prerenal uraemia, 619–20
Presbycusis, 1079, *1079*
Presbyopia, 1086
Pressure sores (decubitus ulcers), 1262–3, *1262*
  prevention, 1262–3
  risk assessment, *1262*
  risk factors, *1262*
  treatment, 1263
Pressure support ventilation, 915
Priapism, sickle cell anaemia, 410
Primaquine
  overdose, 940
  side-effects, haemolysis in G6PD deficiency, *414*
Primary biliary cirrhosis
  autoantigens, *74*
  diagnosis, *69*
  renal tubular acidosis, *676*
  rheumatological problems, 555
  transplantation, 347
Primary biliary cirrhosis (PBC), 355–6
  aetiology, 355
  associations, 355
  complications, 356
  course, 356
  differential diagnosis, 355
  investigations, 355
  prognosis, 356
  treatment, 355
Primary hyperlipidaemias *see* Hyperlipidaemia
Primary lateral sclerosis, 1166
Primary lymphoid tissues, *see* Bone marrow; Lymphoid organs; Thymus
Primary restrictive non-hypertrophic cardiomyopathy, 792–4
Primary sclerosing cholangitis (PSC), 372–3
  clinical features, 373
  diagnosis, 373, *373*
  management, 373
Primary torsion dystonia (PTD), 1150
Primidone, epilepsy, *1143*
Prinzmetal's angina, 749
Prion diseases, *1114*
  CJD *see* Creutzfeldt-Jakob disease (CJD)
  Gerstmann-Straussler-Scheinker syndrome, 121, 122
  kuru, 122
  variant CJD *see* Variant Creutzfeldt-Jakob disease (vCJD)
  see also Transmissible spongiform encephalopathies (TSE)
Prion protein (PrP), 80, *80*, 121, 1158
Pro-opiomelanocortin *see* POMC
Proanthrocyanidins, 600
Probenecid, side-effects
  haemolysis in G6PD deficiency, *414*
  membranous glomerulopathy, 588
Probiotics
  irritable bowel syndrome, 312–13
  ulcerative colitis, 293, *293*

Procainamide
genetic polymorphism in metabolism, *926*
side-effects
respiratory, *877*
systemic lupus erythematosus, 542
Procaine penicillin, syphilis, 179
Procarbazine, Hodgkin's lymphoma, *461*
Prochlorperazine, sickle cell anaemia, *411*
Proctalgia fugax, 243
Proctitis, 292
Proctocolitis, left-sided, 292
Proctography, constipation, *296*
Proctoscopy, 243–4, *244*
Professional bodies, 2
see also specific organizations
Professional Conduct and Performance Committees (GMC), 1
Progenitor cells, 388
Progeria, 37, 226
Progessive supranuclear palsy, 1148
Progestogens
cancer therapy, 463
replacement therapy, *976*
Programmed cell death *see* apoptosis; Apoptosis
Progressive familial intrahepatic cholestasis syndromes, 330
Progressive hepatolenticular degeneration *see* Wilson's disease
Progressive multifocal leucoencephalopathy, 108, *203*, 1158, 1159
Progressive rubella encephalitis, 1158–9
Proguanil, *157*
Proinsulin, *1030*
Projection, 1189
Prokaryotes, 33, 122
Prolactin, 970, *971*
actions, 981, *981*
control, *981*
excess, *see* Hyperprolactinaemia
PCOS, 1004
serum levels, *971*, *981*
Prolactinoma, 972, *972*
Prolamins, 277
Promethazine, breathlessness, 840
Promoters, 31
Promyelocytes, 430
Pronormoblasts, 389
Propafenone, 730, *731*
supraventricular tachycardia, 723
Propionibacterium acnes, 1246
Propranolol, *731*
hypertension, 804
hyperthyroidism, 989, *989*
side-effects
hypoglycaemia, 1060
respiratory, *877*
supraventricular tachycardia, 723
variceal bleeding, 351
Proprioception, 1115
loss of, 1097
Propylthiouracil, 989, *989*
side-effects
agranulocytosis, 989
hepatotoxicity, *366*
vasculitis, 596
Prorenin, 576
Prostacyclin (PGI_2), 431, 577, 651, 896
ALI/ARDS, 919
inhibition of platelet aggregation, 687
vasomotor control, 686, *686*, 687
Prostadynia, 603
Prostaglandin analogues
Prostaglandins, 577
allergic response, *72*
PGD_2, 577
PGE_2, 87, 577
extracellular volume regulation, 651
PGI_2 *see* Prostacyclin
synthesis, *432*
Prostanoids, *686*
Prostate cancer, 491–2, 646
advanced, 492
epidemiology, *449*
geographical distribution, 451
prognostic factors, 491, *491*
screening, 454
treatment, 491–2
Prostate gland
carcinoma *see* Prostate cancer
diseases of, 645–6
benign prostatic hypertrophy *see* Benign prostatic hypertrophy
Prostate-specific antigen, *456*, 491
bone malignancy, 555
Prostatitis, bacterial, 603
Prosthetic heart valves *see* Heart valves, prosthetic
Protease inhibitors, 192, *193*, 194, 196
genetic polymorphism in metabolism, *926*
MM (PiMM), 358
ZZ (PiZZ), 358
Proteases, *58*, 375, 376
Protein(s)
absorption, 272–3
adult requirements, 234
animal, 210
dietary requirements, 209–10
diabetes mellitus, *1037*
intake, 210, 212, 585
interstitial fluid, *650*
intracellular fluid, *650*
loss, 210
metabolism, 321–2, *322*, 576
oxidation, 209–10
plasma, *650*
glycosylated, 1044
proteolytic degradation, resistant proteins, *see* Prion protein (PrP)
synthesis and storage, 321–2
urine *see* Proteinuria
vegetable, 210
see also specific proteins
Protein binding, in renal disease, 618
Protein C, 897
activated, 433, *433*, 434
concentrate, 441
deficiency, 443
Protein kinase C, 24
Protein phosphorylation, 24
Protein S, 219, 433, 897
deficiency, 443
Protein-calorie malnutrition *see* Protein-energy malnutrition (PEM)
Protein-energy malnutrition (PEM), 207, **212–17**
body mass index (BMI), 214
clinical features, 214
conditions associated with, *212*
developed countries, 212–15
developing countries, 215–17, *215*
essential fatty acid deficiency, 211
investigations, 216
risk factors, *213*
treatment, 214–15, 216–17
Protein-losing enteropathy, 246, 282
hypoalbuminaemia in, 321
tests for, 277
Protein 4.1, 404
Proteinase-3, 56, *57*, 504
ANCA, 870
Proteinuria, 578–9, *585*
antithrombin deficiency, 442–3
asymptomatic, 584
chronic renal failure, 627
IgA nephropathy, 594
nephrotic syndrome, 584
postural, 579, 593
without haematuria, 593
Proteoglycans, *58*, 500, 519
Proteolysis
Proteomics, 48, *48*
Proteus, *123*
P. mirabilis, *70*
urinary tract infection, 600
P. morgani, 950
P. vulgaris, pyogenic abscess, 361
septicaemia, *136*
Prothrombin gene, variants, 442
Prothrombin time (PT), 324, 433, 434–5
cirrhosis, 346
DIC, 441
haemophilia A, *439*
hepatitis A, 334–5
oral anticoagulant therapy, 445
vitamin K deficiency, 324, *439*
von Willebrand disease, *439*
Prothrombinase, 432
Proto-oncogenes, 49, *453*
Proton pump inhibitor test, 250
Proton pump inhibitors (PPIs), 252, 263
pancreatitis, 381
Protoporphyrins, *390*, 1073
Protozoal infections, **153–62**
blood and tissue, 153–60
gastrointestinal, 160–2, *161*
HIV-associated, 201
see also individual infections
Protrusio acetabulae, 527
Proximal interphalangeal joints (PIP), 509
nodal osteoarthritis, 520
Prurigo of pregnancy, 1269
Pruritic folliculitis of pregnancy, 1269
Pruritus, 1253, *1253*
chronic renal failure, 629
HIV, 188
hypercalcaemia, 629
hyperparathyroidism, 629
hyperphosphataemia, 629
primary biliary cirrhosis, 355
Pruritus ani, 299
Pseudo-acanthosis nigricans, 1251
Pseudo-hypoaldosteronism
autosomal dominant, *659*, 667
type 1 (autosomal recessive), *659*, 667
type 2 *see* Gordon's syndrome
Pseudo-pseudohypoparathyroidism, 1022
Pseudoacromegaly, *1035*
Pseudobulbar palsy, 1110–11, 1129, 1166
Pseudocholinesterase, genetic polymorphism, *926*
Pseudogout, 501, 539
elderly patients, 553
synovial fluid, 504
Pseudohallucinations, 1188
Pseudohypoparathyroidism, 1022
Pseudomembranous colitis, *132*, 307
Pseudomonas, *82*, 122, *123*
CAPD peritonitis, *637*
cystic fibrosis, 845
P. aeruginosa, *70*
antibiotic resistance, *94*
bronchiectasis, 843
pneumonia, *858*, 861
skin and soft tissue infections, *124*
ventilator-associated pneumonia, 914
P. ovale, 1228
septicaemia, *136*
Pseudoporphyria, chronic renal failure, 629
Pseudoseizures, 1144–5
Pseudoxanthoma elasticum, *435*, 1263
and gastric bleeding, *267*
Psilocybe, 952
Psilocybin, 952
Psittacosis (ornithosis), *80*, 127–8
Psoralen, 1248
Psoriasiform rash, *1265*
Psoriasis, **1240–3**
aetiology, 1240, *1241*
arthritis in *see* Psoriatic arthritis
chronic plaque, 1241, *1241*
clinical features, 1241–2, *1241*
erythrodermic/pustular, 1242
flexural, 1241
guttate, 1241, *1241*
HLA system in, *60*
nails in, 1242, *1242*
pathogenesis, *1241*
pathology, 1240–1
prognosis, 1243
treatment, 1242–3
Psoriatic arthritis, 534–5, 1242
children, 553
clinical features, 534–5, *534*
pencil in cup appearance, 535, *535*
prognosis, 535
treatment, 535
Psychiatric disorders, **1189–204**
causes, 1190
classification, 1189–98
cultural and ethnic background, 1185–6
epidemiology, 1185, *1185*
history, 1186
increased risk, *1191*
Mental State Examination (MSE), 1186–9
and physical disease, 1190–1, *1190*
physical examination, 1189
prevalence, *1185*
treatment, 1190–1
consent/competency, 6–7
see also individual disorders
Psychiatric interview, *1186*
Psychiatry
legal issues
consent/competency, 6–7
see also Legal issues
Psychological debriefing, 1210
Psychological effects, breaking bad news, 14
Psychological factors (disease aetiology)
Psychological problems, **1185–223**
Psychological symptoms, *1207*
Psychological treatments
alcohol dependence, 1213
anxiety, 1208–9
depression, 1204
obsessive-compulsive disorder, 1211
schizophrenia, 1218
Psychometric assessment, 1122
Psychosis, 1189–90
delirium *see* Delirium
drug-induced, 1216
postpartum, 1199
schizoaffective, 1217
see also Schizophrenia
Psychosocial history, 502
Psychosomatic disorders *see* Functional disorders
Psychosurgery, 1203, 1211
Psychotherapy, 1197
interpersonal, 1204
psychodynamic, 1201
PTEN hamartoma-tumour syndrome, 302
Ptyalism, 248
Puberty, 996–7, *996*
delayed, 997
precocious, 996–7

Pubic lice *see Pediculosis*
Pubic symphysis, 499
Puerperal affective disorders, 1199–200
Pulmonary alveolar proteinosis, 874
Pulmonary artery, 814
  bedside catheterization, 714
  enlargement, 693
  pressure, 902–3, *902*, *903*
    elevated *see* Pulmonary hypertension
Pulmonary artery balloon flotation catheter, 902–3, *902*, *903*
Pulmonary artery pressure, 781
Pulmonary atresia, *774*
Pulmonary capillary wedge pressure, 781
Pulmonary disorders *see* Lung disease
Pulmonary embolism, 442, 784–6
  air travel, 955
  cancer-related, 458–9
  clinical features, 784–5
  diagnosis, 785
  haemodynamic changes, *899*
  investigations, 785–6, *786*
  massive, 784, 785–6
  multiple recurrent, 784–5
  prophylaxis, 786
  Revised Geneva Score, *785*
  right bundle branch block, *720*
  risk, *445*
  risk factors, *785*
  small/medium, 784, 785
  treatment, 786
Pulmonary eosinophilia
  asthmatic bronchopulmonary, 875, *875*
  hypereosinophilic syndrome, *875*, 876
  prolonged, *875*
  simple, *875*
  tropical, 164, *875*, 876
Pulmonary fibrosis
  idiopathic *see* Idiopathic pulmonary fibrosis
  progressive massive, 878
Pulmonary gas exchange, 893
Pulmonary haemorrhage
Pulmonary hamartoma, 884
Pulmonary heart disease (cor pulmonale), **781–6**
  acute cor pulmonale *see* Pulmonary embolism
  chronic, 783–4
    clinical features, 783
    investigations, 783–4
    pathophysiology, 783
    treatment, 784
  COPD, 837
    see also Chronic obstructive pulmonary disease (COPD)
  right bundle branch block, *720*
  see also Pulmonary embolism; Pulmonary hypertension
Pulmonary hypertension, 781–2
  arterial, 694
    ALI/ARDS, 918
  atrial septal defect, 775
  capillary disorders, 782
  causes, 781–2, *781*, *782*
  clinical presentation, 783, *783*
  congenital heart disease, 773–4
  and COPD, 840
  definition, 781, *781*
  management, 782
  mechanisms, 782
  mitral regurgitation, 763
  postcapillary, 782
  precapillary, 782
  primary, 782–3
  prognosis, 783
  treatment, 783
  venous, 694
Pulmonary infiltration, 869
  with eosinophilia, 874–6, *875*
Pulmonary oedema, 655, *694*
  acute, *742*
  disease associations, acute tubular necrosis, 622, 623
  high-altitude, 956, *956*
  management, 918–19
  non-cardiogenic, 917–18
Pulmonary oligaemia, 694
Pulmonary plethora, 693–4
Pulmonary regurgitation, 769
Pulmonary rehabilitation, 840
Pulmonary stenosis, 768–9
  congenital, *774*
  echocardiography, 703
  investigations, 768–9
  right bundle branch block, *720*
  symptoms and signs, 768
  treatment, 769
Pulmonary tuberculosis *see* Tuberculosis
Pulmonary valve, 684
Pulmonary vascular resistance (PVR), 781
Pulmonary veins, 814
  magnetic resonance imaging, 709
Pulpitis, 248
Pulse, 690–1, *690*
  aortic regurgitation, 767
  aortic stenosis, 765
  carotid, 690
  character of, 690–1, *690*
  jugular *see* Jugular venous pulse
  mitral stenosis, 761
  rate, 690
  rhythm, 690
Pulse oximetry, *824*, 911
Pulse wave velocity, 799
Pulsed-field gel electrophoresis, 33
Pulseless electrical activity, 711
Pulsus alternans, 690
Pulsus bigeminus, 690
Pulsus bisferiens, 691
Pulsus paradoxus, 690, 797
Pump failure *see* Cardiogenic shock
Punctate basophilia, 945
Pupil
  afferent defect, 1103
  Argyll Robertson, 1104
  constriction, 1103
  convergence reflex, 1103
  dilatation, 1102, 1124
  fixed but not dilated, 1124
  fixed and dilated, 1124
  Holmes-Adie, 1103–4
  Horner's syndrome, 811, 1103, *1104*, 1124
  light reflex, 1103, *1103*
  normal, 1124
  old age, 1103
  physiological changes, 1103
  pinpoint and fixed, 1124
  relative afferent defect, 1103
  unconscious patient, 1124
Pure red cell aplasia, 403
Purgative abuse, 308–9
Purine(s), 1098
Purine analogues, 76
Purine antagonists, 460
Purkinje fibres, 682
Purpura, 434, *1227*
  Henoch-Schönlein *see* Henoch-Schönlein purpura
  immune thrombocytopenic, 436
  infection-induced, 435
  post-transfusion, 437
  senile, 435
  steroid-induced, 435
Pus, 68
Pustular rashes, drug-induced, *1265*
Pustules, *1227*
Puumala virus, 116
PUVA, 1248
Pyelitis, 601
Pyelography
  antegrade, 580–1, *580*
  intravenous (IVP) *see* Excretion urography
  retrograde, 581
Pyelonephritis, 601
  acute, 600, *601*
    acute abdomen, *315*
  chronic (atrophic) *see* Reflux nephropathy
  xanthogranulomatous, 604
Pyloric stenosis
  congenital, 44
    inheritance, *44*
Pyoderma gangrenosum, 1250–1, *1251*
Pyogenic abscess, *91*, 360–1
  clinical features, 361
  investigations, 361
  management, 361
  prognosis, 361
Pyogenic granuloma, 1258
Pyomyositis, *86*
Pyramidal lesions *see* Upper motor neurone (UMN) lesions
Pyramidal system, *see* Corticospinal (pyramidal) system; Upper motor neurone (UMN) lesions
Pyrantal pamoate, *166*
Pyrazinamide
  side-effects, 867, *928*
  tuberculosis, 283, 866
  tuberculous meningitis, 1156
Pyrexia *see* Fever
Pyrexia of unknown origin (PUO), 91, 361
  causes, *91*
Pyridostigmine, myasthenia gravis, 1180
Pyridoxal *see* Vitamin B_6 (pyridoxine)
Pyridoxamine *see* Vitamin B_6 (pyridoxine)
Pyridoxine *see* Vitamin B_6 (pyridoxine)
Pyrimethamine
  side-effects, haemolysis in G6PD deficiency, *414*
  toxoplasmosis, 160, 201
Pyrimidine 5′ nucleotidase deficiency, 414
Pyrimidine antagonists, 460
Pyrogens, 427
Pyrophosphate
Pyrophosphate analogues, 101, *101*
Pyrophosphate arthropathy *see* Pseudogout
Pyrrolizidine alkaloids
  hepatotoxicity, *366*
  and veno-occlusive disease, 360
Pyruvate dehydrogenase, 221
Pyruvate kinase deficiency, 414
Pyruvic acid, 210
Pyuria, sterile, 602

Q fever, 139
  as biological agent, 960
  clinical features, 139
  management, 139
5q-syndrome, 422
QT prolongation *see* Electrocardiogram (ECG)
Quadrantanopia, *1101*
quadruple test, 45
Quality of life, 502
Query fever *see* Q fever
Questioning styles, medical interview, 11
Quinagolide, hyperprolactinaemia, 982
Quincke's sign, 767
Quinidine, 730
  genetic polymorphism in metabolism, *926*
  side-effects
    haemolysis in G6PD deficiency, *414*
    immune haemolytic anaemia, 417
    long QT syndrome, *727*
Quinine
  babesiosis, 160
  overdose, 940
  poisoning, *935*
  side-effects
    haemolysis in G6PD deficiency, *414*
    hepatotoxicity, *366*
    hypoglycaemia, 1060
Quinolones, *94*, 98–9
  cytochrome P450 inhibition, *925*
  indications, 98
  interactions, 98
  mechanism of action, 98
  resistance, 98, 866
  side-effects, haemolysis in G6PD deficiency, *414*
  structure, *98*
  toxicity, 99
  see also individual drugs
Quinsy, 1084
Quinupristin, 99, 124

Rabeprazole, 252
Rabies, *81*, *83*, 119–20
  aerophobia, 120
  clinical features, 119–20
  control, 120
  diagnosis, 120
  dumb, 120
  furious, 120
  hydrophobia, 120
  immunization, *102*, *103*, 120
  prevention, 120
  sylvan, 119
  treatment, 120
  urban, 119
Rabson-Mendenhall syndrome, *1035*
Rac, 24, 26
Radial artery, cannulation, 900, *901*
Radial nerve compression, 1172
Radiation
  and cancer, 450
  and leukaemia, 467
  respiratory reactions, 878
  see also Ionizing radiation; Radiotherapy
Radiation dose, 465
Radiation effect, 465
Radiation myelopathy, 1165
Radiation plexopathy, 1177
Radiation sickness
  acute, 958, *958*
  CNS syndrome, 958
  gastrointestinal syndrome, 958
  haemopoietic sickness, 958
  mild acute, 958
  treatment, 958–9
Radiculopathy, 1071
Radioactivity, 958
Radiocontrast nephropathy, 599
Radiofemoral delay, 777
Radioisotope imaging *see* Radionuclide imaging
Radiology *see* X-ray
Radionuclide imaging
  abdomen, 245–6
  biliary system, 369
  bone *see* Bone scans (scintigraphy)
  brain, 1120
  heart, 704–5
    myocardial infarction, 705
    myocardial perfusion/viability, 705, *705*
  image type, 705
  see also specific techniques
  kidney, 581
    dynamic scintigraphy, 581, *582*
    static scintigraphy, 582

lungs, 824
phaeochromocytoma, 1025
specific conditions
angina pectoris, 749
cancer, 457
Crohn's disease, 289
heart failure, 737
liver disease, 327
pulmonary embolism, 785
ulcerative colitis, 291
urinary tract infection, 602
urinary tract obstruction, 615
thyroid, 993
Radionuclides, 465
Radiotherapy
chemo-radiotherapy, 466
curative, *466*
external beam, 465
hyperfractionation, 465
intensity modulated, 466
palliative, *466*
side-effects, 466, *466*
specific conditions
acromegaly, 980
breast cancer, 488
bronchial carcinoma, 883
cancer, 465–6, *466*
hyperprolactinaemia, 982
pituitary tumours, 973
prostate cancer, 492
systemic radionuclides, 465
thyroid eye disease, 992
treatment planning, 465–6, *466*
Rai staging system, *474*
Raloxifene, osteoporosis, 563, *564*
Raltegravir, *193*
Ramipril
acute coronary syndromes, *754*
heart failure, *739*
hypertension, 804
myocardial infarction, 759
Ramsay Hunt syndrome, 1108
Randomization, 929
Randomized controlled trials, 924, 929–31
assessment, 929–30, *930*
cross-over design, 929
outcomes, 930
parallel group design, 929
Ranibizumab, age-related macular degeneration, 1094
Ranitidine, 252
H. pylori eradication, 261
side-effects, hepatotoxicity, *366*
RANKL, 452, 484, 557
Ranolazine, angina pectoris, 750
Ranson scoring system, 378
RANTES, 850
Rapamycin (sirolimus), 76
Rapeseed oil, and systemic sclerosis, 545
Rapid eye movement (REM) sleep, 1145, 1197, *1197*
Rappaport-Luebering shunt, *413*
ras, 24, 26, 49, 453
Rasburicase, 458
hyperuricaemic nephropathy, 606
Rashes, *88*
aspirate, *89*
drug-induced, 1265–6, *1265*, *1266*
erythematous, *88*
facial, *1245*, **1245–8**
macular/maculopapular, *88*
neonatal, 1267–8, *1268*
papulo-squamous/inflammatory, 1236–45
petechial/haemorrhagic, *88*, *89*
photosensitive, 1248
urticarial, *88*
vasculitic, *596*
vesicular, *88*, *89*
see also Petechiae; specific disorders/infections
Rashkind procedure, 780
Raspberry tongue, 126
Rate drop response, 717
Rathke's pouch, 970
Raynaud's phenomenon, 809, 1252
diagnosis, 809
management, 809
polymyositis/dermatomyositis, 547
symptoms, 809
systemic lupus erythematosus, 543
systemic sclerosis, 545, 546
Re-entry, 716–17, *716*
Reactive arthritis, 535–6
aetiology, 535
clinical features, 535–6, *536*
treatment, 536
Reactive oxygen species, ischaemia-reperfusion, *898*
Real-time PCR, 34
Reasonable doctor standard of disclosure, 4
Reasonable patient standard of disclosure, 4
Reboxetine, 1203
Receptors
abnormalities, 965
antigen, 63–4, *63*
bile acids/salts, 322
cell membrane, 22
cell surface, 966
dimeric transmembrane, 966
endocyte pattern recognition, 59
enzyme-linked surface, 22
extrarenal volume, 652
Fc, 415, 422
G-protein coupled, 22, 966
glycoprotein IIb-IIIa, 753
guanylyl cyclase-linked, 22
hormone, 966–7
hypothalamic-pituitary axis, *971*
membrane, 966
natural killer cells, *66*
nuclear, 966
osmoreceptors, 653, 654
pattern recognition, 55, 58, 59, *59*
serine/threonine kinase, 22
serpentine, 22
signalling, 59
T cell *see* T cell receptors
Toll-like *see* Toll-like receptors
transmembrane, *22*
volume, 651, 652, 654
see also individual receptor types
Reciprocal translocations, 39
Rectal cancer *see* Colorectal carcinoma
Rectal examination, 243–4, 314
Rectal prolapse, 300
Rectal varices, 349
Rectum, 294
physiology, *see* Defecation
role in defecation, 295
see also Colon
Recurrent laryngeal nerve lesions, 1110
Red blood cells (erythrocytes), **389–91**
alloantibodies, 425
atypical antibodies, 425
breakdown *see* Haemolysis
destruction, 628
fragmentation, 627
haemoglobin synthesis, 390–1, *390*
haemolytic transfusion reactions, 427
haemopoiesis (erythropoiesis), 389–90, *390*
see also Erythroblasts; Erythropoiesis; Erythropoietin (EPO)
immune destruction, 414, 415–16
lifespan, 387
haemolytic anaemia, 403
membrane defects, 404–6
hereditary elliptocytosis, 406
hereditary spherocytosis, 404–6
hereditary stomatocytosis, 406
metabolic disorders, 412–14
G6PD deficiency *see* Glucose-6 phosphate dehydrogenase (G6PD) deficiency
metabolism, 412
non-immune destruction, 414
origins of, 388
osmotic fragility, 405
phagocytosis, 416, *416*
production, *391*
sedimentation *see* Erythrocyte sedimentation rate (ESR)
splenic sequestration, 410
turnover, 387
urinary, 579, 627
see also Blood; Haemoglobin
Red cell concentrates, 429, 906
Red cell count (RCC), *389*
Red cell distribution width (RDW), 389, *389*
Red cell volume (RCV)
Red eye, *1092*
Red-cell casts, *579*, 627
Redox imbalance, 897
5α-Reductase deficiency, *1007*
Reduviid bugs, *82*, 158
Reed-Sternberg cells, 476, *476*
Refeeding, 216, *217*
Refeeding syndrome, 216, 234
Reference nutrient intake (RNI), 207
protein, 209
Reflex sympathetic dystrophy, 517
Reflexes, 1115, *1115*
Reflux nephropathy, 600–1, *601*, *605*, 676
chronic, 601
diagnosis, 601
Reflux (oesophageal), 245
Reflux oesophagitis *see* Oesophagitis, reflux
Refractive errors, 1086–7, *1086*
Refractory anaemia with excess blasts, 422
Refractory shock, definition, *899*
Refsum's disease, 1122, 1174
Regression, 1189
Regulatory T cells *see* T cell(s), regulatory (Treg)
Regurgitation, 249, 251
Rehabilitation
heart failure, 738
pulmonary, 840
Rehydration solution for malnutrition, 216
Reidel's lobe, 243, *331*
Reinke's oedema, 1083, *1083*
Reiter's disease, 536
Rejection, acute, 640–1, *641*
Relapsing fevers, *83*, 146–7
diagnosis and management, 147
prevention, 147
see also individual conditions
Relapsing polychondritis, 556–7
Relaxation techniques, 1208
REM sleep *see* Rapid eye movement (REM) sleep
Remnant hyperlipidaemia, 1064–5, *1065*
Remodelling response in asthma, 851–2, *851*
Renal arteriography, 581, *581*
renovascular disease, 608
Renal artery, 571, *572*
fibromuscular disease, 607, 608
occlusion, 609
stenosis, 607
Renal autoregulation, 619
Renal biopsy, 582, *582*
chronic renal failure, 628
complications, *582*
glomerular disease, *585*
Renal blood flow, 581
Renal calculi, 609–14
aetiology, 609–10, *609*
cystinuria, 611
drugs, 611
hypercalcaemia, 610
urinary tract infection, 610–11
autosomal-dominant polycystic kidney disease, 642
causes of, 612–13
clinical features, 611–12, *611*
formation, 677
hypercalciuria, 610
hyperoxaluria, 610
hyperuricaemia and hyperuricosuria, 610
idiopathic stone-formers, 613
investigations, 612, *612*
management, 612
nephrocalcinosis, 613–14, *613*, *614*
pathology, 611
prophylaxis, 613
staghorn calculus, 611
see also Urinary tract stones
Renal cell carcinoma, 644–5
clinical features, 644
diagnosis, 644
epidemiology, *449*
prognosis, 645
pyrexia, *91*
treatment, 644–5
Renal colic, *314*
Renal cortex
Renal disease, **571–647**
calculi *see* Urinary tract stones
diabetic *see* Diabetic kidney
drug handling in, 618–19, *618*
elderly patients, 646–7
glomerulonephritis *see* Glomerulonephritis
HIV-related, 189
and hypertension, 799
and hypoglycaemia, 1060
infections *see* Urinary tract infections (UTI)
investigations, biopsy *see* Renal biopsy
obstructive *see* Urinary tract obstruction
see also specific diseases
Renal failure
acute *see* Acute renal failure (ARF)
acute cortical necrosis, 624
autosomal-dominant polycystic kidney disease, 642
chronic *see* Chronic renal failure (CRF)
cirrhosis, 354
investigations, 627–8
serum prolactin, *981*
see also Renal replacement therapy; Uraemia
Renal function, drug-induced impairment, 617–19
Renal hypertension, 607
see also Renal artery stenosis
Renal natriuretic peptide, 577
Renal osteodystrophy, 628, *628*
Renal parenchymal disease, 620–1
Renal replacement therapy, 634–9
haemodialysis, *see* Haemodialysis
transplantation *see* Renal transplantation
see also Dialysis
Renal scintigraphy, 581
Renal stones *see* Renal calculi
Renal transplantation, 639–42, *639*
antibodies, 640
complications, 640–2
acute rejection, 640–1, *641*
acute tubular necrosis, 640
cardiovascular disease, 642
chronic allograft nephropathy, 641
disease recurrence, 642
infections, 641

Renal transplantation (*Continued*)
lymphoproliferative disorders, 641
malignancy, 641
osteoporosis, 642
technical failure, 640
donor kidney, 639–40
cadaveric donation, 639
living related donation, 639
factors affecting success, 639
matched donor and recipient, 639
immunosuppression for, 639–40
Renal tubular acidosis, 613–14, 676–7, *676*
diagnosis, *676*
and renal calculi, 611
type 1 (distal), 676–7, *676*
type 2 (proximal), 676
type 3, 676
type 4 (hyporeninaemic hypoaldosteronism), 676, *676*
Renal tubules, 572
acute necrosis *see* Acute tubular necrosis (ATN)
cellular recovery, 621
collecting ducts, 571, 653, 655
aldosterone-regulated transport, *659*
intercalated cells, 653, 673
principal cells, 653, 673
sodium transport, *652*
countercurrent mechanism, 573
defects *see* Renal tubular acidosis (RTA)
distal convoluted, 571, 657, *658*
magnesium reabsorption, 668
sodium transport, *652*
transport mechanisms, *658*
function, 575–6
acid-base balance, 575–6
drug and toxicant elimination, 576
investigation, 576
protein and polypeptide metabolism, 576
ischaemic injury, 621
loop of Henle, 571, 573, 574
sodium transport, *652*
transport mechanisms, *658*
proximal tubule
phosphate transport, *671*
sodium transport, *652*
Renal tumours
malignant, 644–5
nephroblastoma (Wilms' tumour), 645
renal cell carcinoma, 644–5
urothelial, 645
Renal vein thrombosis, 609
Renin, 576
Renin inhibitors, hypertension, 805, 1023
Renin-angiotensin system, 576, 799
glomerular dynamics, *632*
Renin-angiotensin-aldosterone system, 651, 652, *1023*
biochemistry and actions, 1023
heart failure, 655, 689
oedematous conditions, 653, 689
Renoprotection, 632, *632*
Renovascular disease, 607–8
atherosclerotic, 607, 608
mechanism of hypertension, 607
pathology, 607–8
prohibited drugs, *923*
screening for, 608
treatment, 608
Reoviruses, 85, *110*, 112
infection, 112
Repaglinide, 1039
Reperfusion injuries, 898, *898*
see also Ischaemia
Repetitive strain injury, 516–17
treatment, 517
Replicative senescence, 226
Report writing, complaints, 15
Repression, 1189
Reproduction, **994–1008**
disorders of *see* reproductive disorders
embryology, 994
females, 995–6, *995*
males, 994–5
menopause, 997–8
puberty, 996–7
see also Pregnancy
Reproductive disorders, 998, *998*
females, **1001–6**
males, **998–1001**
see also individual disorders
Reserpine, hypertension, 805
Residual volume (RV), 815, *816*, 825
β-Resistin-like molecule, 275
Resources, 10
health authorities, 7–8
limited, 7–8
Respiration
artificial *see* Mechanical ventilation; Respiratory support
ataxic, 1124
Cheyne-Stokes, 688, 1124
control of, 814–15
Kussmaul (acidotic) *see* Kussmaul respiration
unconscious patient, 1124
see also Breathing; Lung(s); Oxygen; Ventilation
Respiratory acidosis, 675, *675*, 893
Respiratory alkalosis, 675, *675*, 893–4
Respiratory bronchiolitis interstitial lung disease, 873
Respiratory centre, 814
Respiratory complications
HIV, 189
sickle cell anaemia, 410
Respiratory diseases, **811–18**
infections *see* Respiratory tract infections
investigations, **821–9**
chest X-ray, 822–3
computed tomography, 823
fibreoptic bronchoscopy, 828–9
haematological and biochemical tests, 826–7
intercostal drainage, 827–8
magnetic resonance imaging, 823
mediastinoscopy and scalene node biopsy, 828
pleural aspiration, 827
pleural biopsy, 827
positron emission tomography, 823–4
respiratory function tests, 824–6
scintigraphic imaging, 824
skin-prick tests, 829
transtracheal aspiration, 827
ultrasound, 824
video-assisted thoracoscopic lung biopsy, 829
lower respiratory tract, 835–46
acute bronchitis, 935
bronchiectasis, 842–3
chronic cough, 845–6
COPD *see* Chronic obstructive pulmonary disease (COPD)
cystic fibrosis, 844–5
nocturnal hypoxia, 839–42
obstructive sleep apnoea, 842
symptoms, 819–20
breathlessness, 819–20
chest pain, 820
cough, 819
nasal, 819
sputum, 819
wheezing, 820
upper respiratory tract, 831–5
acute epiglottitis, 834
acute laryngitis, 833
common cold, 831
foreign body inhalation, 834–5
influenza, 834
pharyngitis, 833
rhinitis, 831–3
see also entries beginning pulmonary; Lung disease
Respiratory distress *see* Adult respiratory distress syndrome (ARDS)
Respiratory failure, 910–17
acute, 912
causes, 910–11
COPD, 837
management, 911–17
non-invasive ventilation, 916
oxygen therapy, 912
respiratory support, 912–16
weaning from ventilation, 916–17
monitoring, 911
capnography, 911
pulse oximetry, 911
prohibited drugs, *923*
type I, 910
type II, 910–11
see also Mechanical ventilation; Ventilation
Respiratory function tests *see* Lung function tests
Respiratory muscles, 814
Respiratory stimulants, 841
Respiratory support, 912–16
see also Mechanical ventilation; Ventilation
Respiratory syncytial virus, *90*, *110*, 119
diagnosis and treatment, 119
prevention, 119
Respiratory tract
anatomy, 811–14
defence mechanisms, 818, *818*
humoral and cellular, 818–19
examination, 820–1, *820*, *821*
humidification, 818
particle removal, 818
physiology, 814–18
secretions, 818, *818*
see also Mucus; Sputum
tumours, 880–4
benign, 884
malignant, 880–4
see also Bronchial carcinoma; Lung cancer/tumours
Respiratory tract infections
bacterial, 125–8
lower respiratory tract, 127–8
upper respiratory tract, 125–7
pneumonia *see* Pneumonia
tuberculosis *see* Tuberculosis
Restless legs syndrome, 626, 630, 1197–8
Restriction enzymes, 33
Restriction fragment length polymorphisms, 33
Resuscitation, 709–12
advanced cardiac life support, 711
basic life support, 709–11, *710*
defibrillation, 711
therapeutic hypothermia, 712
unexpected cardiac arrest, 711, *711*
RET proto-oncogene, 965
Retching, 242
Rete ridges, 1225
Reteplase
arterial thrombosis, *443*, 444
myocardial infarction, 757
Reticulocyte count, 389, *389*, 393
aplastic anaemia, 403
Reticulocytes, 390, 403, *405*
Reticulocytosis, 401
Reticuloendothelial system, 323, *416*
Kupffer cells, 320, 322, 323
Retina, 1086, *1086*, 1093–4
age-related macular degeneration, 1093–4, *1094*
central retinal artery occlusion, 1093, *1093*
central retinal vein occlusion, 1093, *1093*
detachment, 1093, *1093*
lesions, 1101, *1102*
photoreceptors, 1086, *1086*
Retinal haemorrhage, and hypoxia, 956, *956*
Retinal vasculitis, 543
Retinal vein thrombosis, 1063
Retinaldehyde, 218
Retinitis, cytomegalovirus, 188, 202, *202*
Retinitis pigmentosa, *1035*
Retinoblastoma, 453, *453*
Retinoblastoma gene, 49–50
Retinocerebellar angiomatosis *see* Von Hippel-Lindau syndrome
Retinoic acid, 218
Retinoids
acanthosis nigricans, 1251
acne vulgaris, 1246, 1247
chronic discoid lupus erythematosus, 1252
hidradenitis suppurativa, 1230
psoriasis, 1242
side-effects, hepatotoxicity, *366*
storage of, 320
teratogenesis, *924*
Retinol *see* Vitamin A
Retinol palmitate, 219
Retinol-binding protein, 576
Retinopathy, diabetic *see* Diabetic retinopathy
Retinyl phosphate, 218
Retrobulbar neuritis, 1102, 1152
Retrograde pyelography, 581
Retroperitoneal fibrosis, 318, 616–17
aetiology, 616
causes, 616–17
clinical features, 617
imaging, 617
laboratory tests, 617
pathology, 616
treatment, 617
Retropulsion, 1097
Retroviruses, 120, *120*, 185
see also HIV
Revascularization, 740
Reverse transcriptase, 35, 185
Reverse transcriptase inhibitors, 192, 195
non-nucleoside, 192, *193*, 195
nucleoside, 195
see also individual drugs
Reye's syndrome, 364, 1159
Rh blood group, 424
incompatibility, 417
RhD immunization, 418
RhD-negative phenotype, 424
Rhabdomyolysis, 621, 627
overdose/poisoning, 936
Rhabdoviruses, *80*, *110*, 119–20
Rhetorical questions, medical interview, 11
Rheumatic autoimmune disorders, 283
Rheumatic diseases
gout, **536–9**
rheumatoid arthritis, 283, 501, **523–32**
autoantibodies *see* Autoantibodies
see also Back pain; Joint(s); specific diseases
Rheumatic fever, 136–7, *137*
children, 554
clinical features, 137

diagnosis, *137*
investigations, 137
pathology, 137
treatment, 137
Rheumatic heart disease
aortic regurgitation, *766*
aortic stenosis, 765
mitral stenosis, 760, *760*
mitral valve stenosis, *703*
see also Rheumatic fever
Rheumatism, palindromic, 526
Rheumatoid arthritis (RA), 283, 501, **523–32**
aetiology, 523–4
anti-citrullinated cyclic peptide, 523, 524–5
antinuclear antibodies, *504*
autoantigens, *74*
complications, 526, *526*
amyloidosis, 526
septic arthritis, 526
diagnosis, *69*
diagnosis and investigations, 528
elderly patients, 552
epidemiology, 523
familial, 523
genetic factors, 523
histology, *524*
HLA system in, *60*, 523
immunology, 523–4
joint involvement, 526–7, *526*, *527*
cervical spine, 527
elbows, 526
feet, 526, *527*
hands and wrists, 526, *526*
hips, 527
knees, 526–7
shoulders, 526
management, 528–32, *528*, *529*
corticosteroids, 529
DMARDs, 529–31, *530*
NSAIDs and coxibs, 529
physical measures, 532
surgery, 532
myopathy, 1179
non-articular manifestations, 527–8, *527*
eyes, 528
heart and peripheral vessels, 528
kidneys, 528
lungs, 527
nervous system, 528
soft tissue, 527
spleen, lymph nodes and blood, 528
vasculitis, 528
pathogenesis, 523–4
pathology, 524–5, *524*, *525*
prevalence, 523
prognosis, 532
proteomics, *48*
respiratory involvement, 871, *871*
rheumatoid factor, *503*, 523, 524–5
seronegative, 525
symptoms, 525
erosions, 524, *525*
hypersplenism, 423
pyrexia, *91*
triggering antigen, 524
typical presentation, 525, *525*
Rheumatoid factor, 503, *503*
IgM, 525
measurement, *69*
rheumatoid arthritis, *503*, 523, 524–5
Rheumatoid nodules, 871
Rhiniphyma, 1247
Rhinitis, **831–3**
food allergy, 237
investigations and diagnosis, 832
nasal obstruction, 1081
nasal polyps *see* Nasal polyps
pathogenesis, 832
perennial, 831–2, *831*
seasonal, 831, *831*
skin-prick tests, 832
treatment, 832–3
vasomotor, 832
Rhinoplasty, 1082
Rhinorrhoea, 819, 1081
Rhinosinusitis, 843
Rhinovirus, *110*, 112, 814
Rhizomelic dwarfism, 23
Rhizopus, 152
Rho, 24, 26
Rhodococcus equii, *191*
Rhodopsin, 218
Rib fractures, 884
Rib notching, 778
Ribavirin, *101*, 102, 117
hepatitis C, 341
Riboflavin (vitamin B_2), 221–2
deficiency, *218*
low intake, 221–2
LRNI, *218*
recommended IV intake, *234*
RNI, *218*
sources of, 221
Ribonucleic acid *see* RNA (ribonucleic acid)
Ribosomal subunits, *31*
Ribosomes, *19*
Ribovirin, side-effects, bone marrow aplasia, 402
Rickets, **567–8**
aetiology, 567, *567*
autosomal-dominant hypophosphataemic, 670
clinical features, 567
investigations, 567
myopathy, 1179
pathology, 567
treatment, 567–8
vitamin D-dependent
type I, 670
type II, 670
vitamin D-dependent type I, 567
vitamin D-dependent type II, 567
X-linked dominant hypophosphataemic, 670
X-linked hypophosphataemic, 567
Rickettsia, *80*, *82*
R. africae, *147*
R. akari, *147*
R. conorii, *147*
R. felis, *147*
R. prowazekii, *82*, *147*
R. rickettsii, *147*
R. typhi, *82*, *147*
Rickettsiae spotted fevers, rash, *88*
Rickettsial pox, *147*
Rickety rosary, 567
Riedel's thyroiditis, 993
Rifabutin, 204
H. pylori eradication, 261
tuberculosis, 866
Rifampicin, *93*
cytochrome P450 induction, *925*
drug interactions, 204
leprosy, *141*
primary biliary cirrhosis, 355
resistance, 866
side-effects, 866–7, *928*
acute tubulointerstitial nephritis, *604*
hepatotoxicity, *366*
immune haemolytic anaemia, 417
spotted fevers, 148
tuberculosis, 283, 866
tuberculous meningitis, 1156
Rift Valley fever, *110*, *114*, 116
Right atrium (atrial), enlargement, 693
Right ventricle (ventricular)
function curve, *902*
hypertrophy, 693
outflow, 699
systolic dysfunction, 736
Right-to-left shunts, 893
Rights, 1–2
see also Ethics
Rigidity, 1111, 1112
Parkinson's disease, 1146
Rimonabant, *232*
diabetes mellitus, 1039
Ring sideroblasts, 397
myelodysplasia, *422*
Ringworm *see Taenia*
Rinne test, 1078
Risedronate
osteoporosis, 563, *564*
Paget's disease, 566
Risk, consent, 4
Risperidone
genetic polymorphism in metabolism, *926*, *927*
schizophrenia, 1218
Risus sardonicus, 134
Ritonavir, *193*, 196
cytochrome P450 induction, *925*
Rituximab, *76*, *464*
acute lymphocytic leukaemia, 475
ANCA-positive vasculitis, 597
autoimmune haemolytic anaemia, 416
Burkitt's lymphoma, 483
cold haemagglutinin disease, 417
cryoglobulinaemic renal disease, 591
diffuse large B cell lymphoma, 482
follicular lymphoma, 481
hairy cell leukaemia, 475
lupus nephritis, 591
lymphoplasmacytic lymphoma, 481
membranous glomerulonephropathy, 588
pemphigus vulgaris, 1256
psoriatic arthritis, 535
rheumatoid arthritis, *530*, 531
systemic lupus erythematosus, 544
thrombocytic purpura, 436
thrombotic thrombocytopenic purpura, 437
Wegener's granulomatosis, 870
Rivaroxaban, 446
arterial thrombosis, *443*
atrial fibrillation, 725
Rizatriptan, migraine, 1137
RNA polymerases, 31
RNA (ribonucleic acid)
double stranded, *59*
microRNA, 29
ribosomal, 29
single stranded, *59*
transfer, 29, 30
RNA viruses, **109–21**, *110*
arenaviruses, 120–1
bunyaviruses, 116
caliciviruses, 112
coronaviruses, 119
filoviruses, 121
flaviviruses, 114–16
orthomyxoviruses, 116–17
paramyxoviruses, 117–19
picornaviruses, 109–12
reoviruses, 112
retroviruses, 120
rhabdoviruses, 119–20
togaviruses, 113–14
RNI *see* Reference nutrient intake
Robertsonian translocations, 39
Rocky Mountain spotted fever, *147*
Rodent ulcer, 1259, *1259*
Rofecoxib, 518
Rokitansky-Aschoff sinuses, 372
Rolling, 67, *68*
Romaña's sign, 158
Romano-Ward syndrome, *727*
see also Torsades de pointes
Romiplostim, thrombocytic purpura, 436
Ropinirole, Parkinson's disease, 1147
Rosacea, 1247, *1247*
Roscotivine, HIV-associated nephropathy, 597
Rose Bengal, 548
Roseola infantum, *105*, 108
Rosiglitazone, 1039
Ross River fever, 541, 1156
Rosuvastatin, hyperlipidaemia, *1066*
Rotator cuff (supraspinatus)
tear, 507
tendonosis, 507
Rotavirus, 85, *110*, 112, *112*
clinical features, 112
and coeliac disease, 278
diagnosis, 112
diarrhoea, *307*
immunization, *103*, 112
treatment and prevention, 112
Roth spots, infective endocarditis, *770*, *771*
Rotigotine, Parkinson's disease, 1147
Rotor's syndrome, 330
Roundworms *see* Nematodes (roundworms)
Roux-en-Y gastric bypass, 233
Rowell's syndrome, 1252
Royal College of Physicians, 2
Royal College of Surgeons, 2
Rubbish
Rubella, *83*, *110*, 113
arthritis in, 540
clinical features, 113
complications, 113
congenital rubella syndrome, 113
diagnosis and treatment, 113
patent ductus arteriosus, *774*, 777
prevention, 113
see also Measles-mumps-rubella (MMR) vaccine
rash, *88*, *113*
Rubeosis, 1052, 1053
Rugger jersey spine, 629
Rumination syndrome, 311
Ryanodine receptor, 683

26S proteosome, 20
S100, *493*
Saccular (berry) aneurysms, 1134, *1134*
Sacroiliac joints, 499
Sacroiliitis, 535
Sagramostim, Crohn's disease, *290*
St John's wort, 1145, 1203
St Louis encephalitis, 1156
St Vitus's dance *see* Sydenham's chorea (St Vitus's dance)
Salbutamol
asthma, 854, *854*
COPD, 839
genetic polymorphism in metabolism, *927*
hyperkalaemia, 668
pulmonary hypertension, 840
Salicylates, 87
overdose, *935*, 948–9
clinical features, 949
treatment, 949
side-effects
hepatotoxicity, *366*
hypoglycaemia, 1060
renal calculi, 611
Salicylic acid, viral warts, 1232
Saline, adenine, glucose, 906
Saline (salt solution)
distribution, 650, *650*
hypertonic
Saliva, hormone estimations, 968
Salivary glands, 248
neoplasms, 248
obstruction, 248
sarcoidosis, 248
sialadenitis, 248

Salivation, excessive, 248
Salmeterol, asthma, *854*
Salmonella, *70*, 85, *90*, *123*, *191*
  antibiotic resistance, *94*
  antibiotics, *132*
  diarrhoea, *307*
  food poisoning, *133*
  gastroenteritis, *128*, 129
  HIV-related, 204
  reactive arthritis, 535
  S. choleraesuis, 129
  S. enteritidis, *80*, 129
  S. paratyphi, 144
  S. typhi, 144
  S. typhimurium, 129
Salpingitis, 316
Salt *see* Saline; Sodium
Salt loading, 652
Sampling error, 933
San Filippo syndrome, 1071
Sanctuary sites
Sandflies, *82*, 158
Sapovirus, *110*, 112
SAPS scoring system, 920
Saquinavir, *193*, *195*, 196
Sarcocystis, *161*
  S. lindemanni, *1179*
Sarcoidosis, 556, 822, **868–70**, 1253
  clinical features, 869
    bilateral hilar lymphadenopathy, 869
    cardiomyopathy, *793*
    pulmonary infiltration, 869
    pyrexia, *91*
  epidemiology and aetiology, 868
  immunopathology, 868–9
  investigations, 870
  myopathy, 1179
  prognosis, 870
  renal, 606
  salivary glands, 248
  treatment, 556, 870
Sarcomere, 682
Sarcoplasmic reticulum, 684
Sarcoptes scabiei, 1234
Sarcosporidiosis, *1179*
Sarin, 946
Satavaptan, *657*, 663
Satisfaction
  clinician, 10
  patient, 9–10
Saxitoxin, 950
Scabies, 1234–5, *1234*
  HIV-related, 204
  Norwegian, 204, 1235
  treatment, 1234–5
Scalded skin syndrome, *86*
Scalene node biopsy, 828
Scaling, *1227*
Scalp, tenderness, 1096
Scaphoid fractures, 509
Scarlet fever, *83*, 126
  clinical features, 126
  diagnosis, 126
  prevention, 126
  rash, *88*
  staphylococcal, *86*
  treatment, 126
Scavenger-receptor B_1, 1062
Schatzki's ring, 255
Scheie syndrome, 1071
Schilling test, 400
Schirmer tear test, 548
Schistosoma, *163*
  geographical distribution, *169*
  S. haematobium, 168, 169, 361
    treatment, *170*
  S. intercalatum, 168
  S. japonicum, 168, 169, 361
    treatment, *170*
  S. mansoni, 168, 169, 361
    treatment, *170*
  S. mekongi, 168
Schistosomiasis, 80, **168–70**, 361
  clinical features, 169
  diagnosis, 169
  diarrhoea, *307*
  geographical distribution, *169*
  hepatomegaly, *331*
  management, 169, *170*
  parasitology and pathogenesis, 168–9
  portal hypertension, 349
  prevention, 169–70
  rash, *88*
  returning travellers, *90*
Schizophrenia, 1216–18
  causes, 1216
  chronic, 1216
  clinical features, 1216
  diagnosis, 48
  differential diagnosis, 1217
  inheritance, *44*
  negative, 1216
  positive, 1216
  prognosis, 1217
  treatment, 1217–18, *1217*
    see also Neuroleptics
Schmidt's syndrome, 1026
Schmorl's node, 510
Schoeber's test, 533
Schumm's test, 403
Schwachman-Diamond syndrome, *430*
Schwann cells, 1055
Schwannoma *see* Neurofibroma (Schwannoma)
Sciatica, 1176–7
Scintigraphy *see* Radionuclide imaging
Scintiscanning *see* Radionuclide imaging
Sclera, 1085, 1086
Scleritis, 528
Scleroderma, 1252
  autoantigens, *74*
  diffuse cutaneous, 546
  limited cutaneous, 546
  see also Systemic sclerosis (SSc)
Scleroderma renal crisis, 609
Scleromalacia perforans, 528
Sclerosing cholangitis
  primary *see* Primary sclerosing cholangitis
Sclerosing peritonitis, 638
Sclerostin, 558
Sclerotherapy, variceal, 350
Scombroid fish poisoning, 950, 1145
Scombrotoxin, *133*
Scorpion stings, 951
Scotoma, 1101, *1101*, *1102*
Screening
  autosomal-dominant polycystic kidney disease, 643
  Barrett's oesophagus, 253
  C. trachomatis, **174**
  cancer, 454
    lead time bias, 454, *454*
  colorectal carcinoma, 305
  cystic fibrosis, 845
  gastric cancer, 264–5, *264*
  genetic *see* Genetic screening
  gonorrhoea, **174**
  haemochromatosis, 357
  hepatitis viruses, 173
  hyperlipidaemia, 1065, 1067
  hypothyroidism, 987
  lung cancer, 884
  sight, 1052
Scrofuloderma, 1230
Scrub typhus, 147–8, *147*
Scrum pox, 1228
Scurvy, *218*, 223, *223*, *789*
  clinical features, *223*
  diagnosis, 223
  prevention, 223
  treatment, 223
Seasonal affective disorder, 1199
Seasonal rhinitis, 831, *831*
Sebaceous cyst *see* Epidermoid cyst
Sebaceous glands, 1227
Seborrhoeic eczema, 1238–9, *1239*
  aetiology, 1238
  clinical features, 1238–9
  HIV-related, 1269
  treatment, 1239
Seborrhoeic wart, 1257, *1257*
Secalins, 277
Secondary biliary cirrhosis, 356
Secondary messengers, 24
  phosphorylation, 24
Secretin, *272*, 322, 375
Secretin-cholecystokinin stimulation test, 376
Secretory diarrhoea, 306
See Chagas' disease
SeHCAT test, 308
Seizures, *1140*, 1161
  absence, 1140
  akinetic, 1140
  generalized, 1139, *1139*
  Jacksonian, 1139
  myoclonic, 1140
  partial, 1139, *1139*, 1140
  temporal lobe, 1140
  threshold, 1139
  tonic, 1140
  tonic-clonic, 1140
  see also Convulsions; Epilepsy
Seldinger technique, *902*
Selectins, *25*, 27
  E-selectin, 27, 67, 545, *686*
  L-selectin, 27, 67, *686*
  P-selectin, 27, 67, *686*
Selective oestrogen receptor modulators (SERMs), 463, 997
Selective serotonin reuptake inhibitors (SSRIs)
  anxiety disorders, 1209
  depression, 1201, 1202, *1202*
  genetic polymorphism in metabolism, *926*
  obsessive-compulsive disorder, 1211
  overdose, 949
  see also 5-Hydroxytryptamine (5-HT; Serotonin)
Selenium, 226
  deficiency, 226
  recommended IV intake, 233
  toxicity, 226
Selenium sulphide, 1234
Self-poisoning, 934
  examination, 934–5
  history, 934, *934*
  paracetamol, 934
  prevention, *934*
  see also Deliberate self-harm; Poisoning
Self-reference, 1188
Self/nonself recognition *see* Tolerance
Semicircular canals, 1077
Seminiferous tubule dysgenesis *see* Klinefelter's syndrome
Seminoma, 492
Senile eczema, 1239
Senile miosis, 1103
Senile purpura, 435
Senna, sickle cell anaemia, *411*
Sensorimotor polyneuropathies
Sensory ataxia, 1097
Sensory cortex, 1115, *1116*
Sensory loss/symptoms, *1117*
  dissociated, 1117
Sensory pathways, 1115–19
  lesions, 1116–18
    see also Sensory loss/symptoms
Sensory polyneuropathy, HIV-related, 188
Sentinel lymph nodes, 457
Sepsis, 135, 909–10
  acute tubular necrosis, 623
  adjunctive treatment, 909–10
  catheter-related, *136*, 235–6
  clinical features, 899
  definition, *899*
  diagnosis, 899
  management, 909–10
  mortality, 909
  nephrotic syndrome, 585
  severe, 135
    definition, *899*
  terminology, *899*
  see also Septicaemia; Septic shock
Septal defects
  ventricular *see* Ventricular septal defects (VSD)
Septata intestinalis, 201
Septic arthritis, 503, 504, 526, 539
  clinical features, 539
  elderly patients, 553
  infected prostheses, 539
  investigations, 539
  prognosis, 539
  treatment, 539
Septic discitis, 512
Septic shock, 135
  ADH deficiency, 895
  clinical features, 899
  definition, *899*
  genetic variation in response, 897
  haemodynamic changes, 897, *899*
  inflammatory mediators, 895–7
    adhesion molecules, 896
    arachidonic acid metabolism, 895
    complement activation, 895
    cytokines, 895
    endothelium-derived vasoactive mediators, 896–7
    lysosomal enzymes, 895–6
    platelet-activating factor, 895
    toxic products of microorganisms, 895
  microcirculatory changes, 897
  redox imbalance, 897
  see also Septicaemia (sepsis syndrome)
Septicaemia (sepsis syndrome), *82*, *86*, 135–6, *135*, *136*
  causes, *135*, *136*
  meningococcal *see* Meningococcal septicaemia
  vasodilatation, 660
  see also Sepsis; Septic shock
Septicaemic plague, 146
Septo-optic dysplasia, 974
Sequestosome p62, Paget's disease, 565
Serglycan, 500
Serine protease inhibitors (serpins)
Serine/threonine kinase receptors, *22*
Serious Hazards of Transfusion (SHOT) scheme, 425, *426*
Seroconversion illness, 108
Serotonin *see* 5-Hydroxytryptamine
Serotonin and noradrenaline reuptake inhibitors (SNRIs), 1202
Serotonin syndrome, 1202
Serpentine receptors, *22*
Serpins, 358, 433
Serratia
  S. marcescens, *70*
  septicaemia, *136*
Sertraline, overdose, 949
Serum amyloid A (SAA)
Serum amyloid protein, 528
Sevelamer, 633
Severe acute respiratory syndrome (SARS), 79, *80*, 91, *110*, 119, 859–60
  transfusion-transmitted, 428
Severe combined immunodeficiency (SCID), 46, 70
Sever's disease, 515, *515*

Sevoflurane
  side-effects, *928*
    hepatotoxicity, 365
Sex
  chromosomal *see* Sex chromosomes
  gonadal, 1008
  phenotypic, 1008
  social (gender), 1008
Sex chromosomes, 1008
  disorders, 1007–8, *1007*
  see also individual disorders
Sex differences *see* Gender
Sex hormones *see* Androgen(s); Oestrogen(s); Testosterone
Sex-limited inheritance, 43
Sex-linked disorders, *42*, 43
  X-linked dominant disorders, 43
  X-linked recessive disorders, 43
  Y-linked, 43
Sex-linked genetic disorders, *see* X-linked genetic disorders
Sexual abuse, 1148, 1201
Sexual activity, heart failure, 738
Sexual assault, 173–5
  examination, 174
  history, 173–4
  investigations, 174–5
  management, 175
Sexual deviation, 1221
Sexual differentiation disorders, 1007–8, *1007*
  see also individual disorders
Sexual disorders, 1221, *1221*
Sexual dysfunction, male *see* Erectile dysfunction
Sexual orientation, 1008, 1221–2
Sexually-transmitted diseases (STDs), **172–205**
  AIDS *see* AIDS; HIV infection
  approach to patient, 172–3
  control, 173
  examination, 173
  genital ulceration, *172*
  history, 172–3
  infestations, 183–4
  investigations, 173
    asymptomatic screening, 173, **174**
    symptomatic patients, 173, **174**
  prevention, 173
  treatment, 173
  urethral discharge, *172*
  vaginal discharge, *172*
  see also specific infections
Sézary syndrome, 483
Shagreen patches, 1253
Sheehan's syndrome, 974
Shellfish allergy, 73
Shigella, *123*
  antibiotic resistance, *94*
  antibiotics, *132*
  diarrhoea, *307*
  food poisoning, *133*
  gastroenteritis, *128*, 129–30
  reactive arthritis, 535
  S. dysenteriae, 129
  S. flexneri, 129
  S. sonnei, 129
Shingles *see* Herpes zoster (shingles)
Shock, **894–910**
  anaphylactic, 897, 899
  cardiogenic, *894*, 898
  causes, 894, *894*
  clinical features, 898–9
  coagulation system activation, 897–8
  distributive, *894*
  heart sounds, 692
  hypovolaemic, *894*, 898
  inflammatory mediators, 895–7
  management, 904–10, *905*
    myocardial contractility and inotropic agents, 907–8
    preload and volume replacement, 904–6, *905*
    sepsis and multiple organ failure, 909–10
    vasodilator therapy, 908–9
  monitoring of critically ill patients, 899–904
    blood pressure, 900, *900*
    cardiac output, 903–4
    central venous pressure, 901–2, *901*, *902*
    left atrial pressure, 901–2, *901*, *902*
    tissue perfusion, 900, *901*
  neuroendocrine response, 895
  obstructive, *894*, 899, 904
  pathophysiology, 894–8
  refractory, *899*
  septic *see* Septic shock
  sympatho-adrenal response, 894–5, *894*
  see also Post-traumatic stress disorder (PTSD)
Shock wave lithotripsy, 370, 381, 612
Short QT syndrome, 728, *789*
Short stature, 978–9, *979*
Short-bowel syndrome, 281
Shoulder, *507*
Shoulder girdle pain, 505–6
Shoulder impingement syndrome, 507
Shoulder pain, 507–8
  bursitis, 507–8
  calcific tendonosis, 507
  differential diagnosis, *507*
  frozen shoulder, 507, 508
  rheumatoid arthritis, 526
  rotator cuff tendonosis, 507
  torn rotator cuff, 507
SHOX gene, 978
Shrinking lung syndrome, 543
Shuffling gait, 1097
Shunt nephritis, 593
Shwachman's syndrome, 71
Shy-Drager syndrome, 1148
  postural hypotension, *659*
Sialadenitis, 248
Sialyl-Lewis^x, 67
Siberian ginseng, 1145
Sibutramine, 232
Sicca syndrome, 528
Sick euthyroid syndrome, 984
Sick role, 1191
Sick sinus syndrome, 717, 792
  pacing, 717
Sickle cell anaemia, 409–11
  clinical features, 409–10
    anaemia, 410
    bone marrow aplasia, 410
    splenic sequestration, 410
    vaso-occlusive crises, 409–10
  complications, *411*
  investigations, 410
  long-term problems, 410
  management, 411, *411*
  prognosis, 411
  renal tubular acidosis, *676*
  splenic atrophy, 423
Sickle cell crisis
  acute abdomen, *315*
  joint pain, 556
Sickle cell trait, 409, 411–12
  pregnancy, 412
Sickle cells, *409*
Sickle nephropathy, 597
Sickle solubility test, 410
Sickle syndromes, 409–12
  genetics
  pathogenesis, 409
  see also individual syndromes/traits
Sideroblastic anaemia, 393, *396*, 397, 945
  classification, 397
  treatment, 397
Sideroblastic granules, 422
Sigmoidoscopy, 243
  acute abdomen, 315
  flexible, 243
Signalling receptors, 59
Signet ring cells, 304
Sildenafil
  erectile dysfunction, 1000, 1056
  pulmonary hypertension, 783
Silica dust, and systemic sclerosis, 545
Silicosis, 878
SIM1, 229
Simvastatin
  acute coronary syndromes, *754*
  hyperlipidaemia, 746, *1066*
  myocardial infarction, 759
Sin Nombre, 116
Single nucleotide polymorphisms, 35–6
  linkage disequilibrium, 36
  lod score, 36
Single-gene disease, 42–4
  autosomal dominant disorders, 42, *42*
  autosomal recessive disorders, *42*, 43
  imprinting, 44
  mitochondrial disease, 44
  sex-limited inheritance, 43
  sex-linked disorders, *42*, 43
  triplet repeat mutations, 43–4, *44*
Single-photon-emission computed tomography *see* SPECT
Sinoatrial disease, 717
Sinoatrial node, 682
Sinus arrhythmia, 716
Sinus bradycardia, 716, 717
  extrinsic causes, 717
  intrinsic causes, 717
  treatment, 717
Sinus node, 681–2, *682*, 715–17, *715*
  blood supply, 685
Sinus rhythms
Sinus squeeze, 957
Sinus tachycardia, 716, *721*
  inappropriate, 720–1
Sinusitis, 1081–2, *1082*
  treatment, 1081–2
  types of, *1082*
Sirolimus
  autosomal-dominant polycystic kidney disease, 644
  renal transplantation, 640
Sitagliptin, 1039
Sitaxentam
  pulmonary hypertension, 783
  systemic sclerosis, 547
Sjögren's syndrome, 548
  antinuclear antibodies, *504*
  autoantigens, *74*
  dry mouth, 248
  management, 548
  pathology and investigations, 548
  renal tubular acidosis, *676*
  rheumatoid factor, *503*
Skeletal dysplasias, 568
Skeletal muscle, 501
Skew deviation, 1105
Skill, lack, clinical communication, 11
Skin
  basement membrane zone, 1227, *1227*
  care
  dermis, 1227
  in diabetes mellitus, 1057
  diseases *see* Skin diseases
  epidermis, 1225–6
  function, *1227*
  innervation, 1227
  specific conditions *see* Skin diseases
  structure, 1225–7, *1226*
  subcutaneous layer, 1227
  sun damage, *see* Melanoma(s); Skin tumours
  tumours/malignancy *see* Skin tumours
Skin biopsy, *1228*
Skin diseases, **1225–71**
  bullous, 1255–7, *1255*
  cutaneous signs of systemic disease, 1250–5
  eosinophilia, *431*
  examination, 1227–8, *1227*
  history, 1227
  HIV-related, 1268–9
  infections *see* Skin and soft tissue infections
  investigation, 1228, *1228*
  malignancy *see* Skin tumours
  of pregnancy, 1269–70, *1269*
  see also Petechiae; Rashes; specific diseases
Skin scrapes, *1228*
Skin and soft tissue infections, 123–5, **1228–35**
  bacterial, 1228–30
  deep, 125
    gas gangrene, 125
    necrotizing fasciitis, 125
  developing/tropical countries, 140–3
    anthrax, *80*, *83*, 142
    bejel, 142
    endemic treponematoses, 142–3
    leprosy, 140–1, *140*, *141*
    mycobacterial (Buruli) ulcer, 142
    pinta, 142–3
    yaws, 142
  fungal, 1232–4
  infestations, 1234–5
  mycobacterial infections, 1230
  papulo-squamous/inflammatory rashes, 1236–45
  predisposing factors, *124*
  superficial, 123–5
    cat scratch disease, 125
    MRSA, 124
    pasteurellosis, 124–5
    Staphylococcus aureus, 123–4, *124*
  toxin mediated, 125
    staphylococcal scalded skin syndrome, 125
    toxic shock syndrome, 125
  viral, 1230–2
  see also Dermatological disorders; specific infections
Skin swabs, *1228*
Skin tumours, **1257–61**
  benign, 1257–8
  cutaneous lymphoma, 483
  malignant, 1259–61
  potentially pre-malignant, 1258–9
  see also Melanoma; Naevus; Squamous cell carcinoma (SCC): and individual tumours
Skin tumours ultraviolet light, 450
Skin-associated lymphoid tissue (SALT), 53
Skin-prick tests, 829
  asthma, 853
  rhinitis, 832
Skinfold thickness, 230
Skull
  basal impression (platybasia), 1169
  base, 1107
  pepperpot, 628
  sutures, 499
Skull fractures, 1119, 1163
  see also Skull X-rays
Skull X-rays, 1119
Slapped cheek disease, 109, 540, 1230
Slapping gait, *1096*, 1097

Sleep, 1123, 1145
  deprivation, and epilepsy, 1141
  hormone secretion during, 968
  seizures, 1145
Sleep disorders, 1145, 1197–8, *1197*
  cataplexy, 1145
  central sleep apnoea syndrome, 688, 842, 1145
  and depression, 1200–1
  familial fatal insomnia, 145
  insomnia, 1145, 1197
  narcolepsy, 1145
Sleep-talking, 1198
Sleep-walking, 1198
Sleeping sickness *see* African trypanosomiasis
Slow reacting substance of anaphylaxis, *72*
Slow-transit constipation, 296
Smac protein, 28
Small airways, 812
Small bowel enema, 245
Small bowel follow-through, 245, 276–7
Small intestinal disease, 282–4
  investigations, 276–7, *276*
    absorption tests, 277
    biopsy, 277
    blood tests, 276
    CT scanning, 277
    14C-glycocholic acid breath test, 277
    hydrogen breath test, 277
    MRI enteroclysis, 277
    small bowel enema *see* Small bowel enema
    small bowel follow-through *see* Small bowel follow-through *see also* Sigmoidoscopy
  presenting features, 275–6
  tumours *see* Small intestine tumours
  see also Malabsorption; specific diseases
Small intestinal pseudo-obstruction, 317
Small intestine, **270–85**
  absorption, 271
    active transport, 271
    carbohydrate, 271–2, *273*
    facilitated diffusion, 271
    fat, 273–4, *273*
    protein, 272–3
    simple diffusion, 271
    vitamins and minerals, 274
    water and electrolytes, 274
  commensal bacteria, 275
  disease *see* Small intestinal disease
  hormones, *272*
  innervation, 271
  motility, 271
  neuroendocrine peptides, 271
  physiology, 271–4
  response to antigens and pathogens, 274–5
    cellular defence, 275
    innate chemical defence, 274–5
    innate immunological defence, 275
    physical defence, 274
  structure, 270–1, *270*
Small intestine tumours, 284–5
  adenocarcinoma, 284
  carcinoid tumours, 284–5
  lymphoma, 284
  Peutz-Jeghers syndrome, 285
  predisposing factors
    coeliac disease, 284
    Crohn's disease, 284
    immunoproliferative small intestinal disease, 284
  treatment, 284
Small-volume pulse, 690
Smallpox, *83*, 109
  as biological agent, 959–60
Smell, 1080, 1100
  sensorineural loss, 1082
Smoke, 830, 959
Smoking, 829–30
  airflow effects, *836*
  and bone density, 563
  and COPD, 835, 837
  and coronary artery disease, 745–6
  cough, 819
  dangers of, 829, *829*
  effects on lung, *829*
  and gastric cancer, 264
  and lung cancer, 449
  passive, 829
  and polycythaemia, 421
  prevalence, 829
  sputum production, 819
  toxic effects, 829
Smoking cessation, 830
  COPD, 838
  heart failure, 737
Snail track ulcers, 178
Snake bites, 951, *951*
  clinical features, 951, *951*
  management, 951–2
SNARE proteins, 21, *21*
Sneezing, 818, 819, 832
Snoring, 842, *842*, 1084, *1084*
Social outcomes, effect of communication, 9
Social phobia, 1208
Social treatments
  depression, 1204
  schizophrenia, 1218
Sodium
  24-hour output, 576
  concentration disorders *see* Hypernatraemia; Hyponatraemia
  dialysate
    haemodialysis, *634*
    peritoneal dialysis, *637*
  dietary intake, 207, 212
    high, 652
    and hypertension, 798–9
    low, 652
  excretion, 650–1
  extracellular volume regulation, 650
  fractional excretion (FENa), 620
  and hypertension, 207
  input, *295*
  interstitial fluid, 649, *650*
  intracellular fluid, *650*
  losses from gut, *662*
  output, *295*
  plasma/serum, *650*
  renal transport, *652*
  restriction, 656
  retention, 351, 653, 656
    heart failure, 734
    ventilation-related, 914
  RNI, 207
  urinary, 620
Sodium aurothiomalate *see* Gold
Sodium bicarbonate, *660*
Sodium calcium edentate, *937*
Sodium chloride, *660*
  retention, 655
  see also Saline (salt solution); Sodium
Sodium citrate, enema, *297*
Sodium cromoglicate, 850
  allergic conjunctivitis, 1089
  asthma, *854*, 855
  rhinitis, 833
Sodium fusidate
  erythrasma, 1228
  folliculitis, 1228
  pitted keratolysis, 1230
Sodium nitrite, *937*, 943
Sodium nitroprusside
  hypertension, 805
  vasodilation, 909
Sodium taurocholate co-transporting polypeptide, 321
Sodium thiosulphate, *937*, 943
Sodium valproate
  diabetic neuropathy, 1055
  epilepsy, *1143*
    dose, *1143*
    therapeutic range, *1143*
  mania, 1206
  overdose, 940
  side-effects, *1143*
    acute tubulointerstitial nephritis, *604*
    hepatotoxicity, *366*
  teratogenesis, *924*
Sodium-channel blockers, supraventricular tachycardia, 723
Sodium-chloride cotransporter, *658*, 667
Sodium/potassium-ATPase pump, 649, 673
Soft chancre *see* Chancroid
Soft tissue infections *see* Skin and soft tissue infections
Solar (actinic) keratoses, 1258, *1258*
Solar urticaria, 1249
Solitary rectal ulcer syndrome, 300
Solvent abuse, *1114*, 1214
Soman, 946
Somatic mutation theory of ageing, 226
Somatization disorder, 1195
Somatoform disorders, 1195
Somatostatin, 229, 258, *272*, 322, 376, 970, *971*
  control of growth hormone secretion, 977
  variceal bleeding, 350
Somatostatin receptor agonists
  acromegaly, 980
  pituitary tumours, *973*
  see also individual drugs
Somatostatinoma, 385
Sorafenib, *464*
  renal cell carcinoma, 645
Sorbitol, 655
Sotalol, *731*
  arrhythmogenic right ventricular cardiomyopathy, 791
  side-effects, long QT syndrome, *727*
  supraventricular tachycardia, 723
Sound
  A-weighted, 959
  bone conduction, 1078
  intensity, 959
  pressure, 959
South American trypanosomiasis *see* Chagas' disease
Southern blotting, 34
Space of Disse, 320
  cirrhosis, 345
Spasm
  hemifacial, 1108
  torticollis, 1148, 1150
  trigeminal neuralgia, 1106–7
Spastic diplegia, 1169
Spastic paraparesis, *1112*, 1117, 1152
Spasticity, *1096*, 1097, 1112
SPECT
  brain, 1120
  heart failure, 737
  infections, 87
  myocardial infarction, 704, *705*
Spectinomycin, gonorrhoea, 176
Spectrin, 404, 405
  cross-linking, 412
Speech, 1187
  Parkinson's disease, 1146
  poverty of, 1187
  pressure of, 1187
Speech therapy, 1133
Spherocytes, 405, *405*
Spherocytosis, *416*
  hereditary, 368, 404–6
    causes, *405*
    clinical features, 405
    investigations, *405*
    treatment, 405
Sphincter pupillae, 1103
Sphincters
  disturbance of, 1117
  of Oddi, 320, 322
Sphingolipids, Fabry's disease, 597
Sphingomyelin cholesterol lipidoses, 1071
Spider bites, 951
Spider naevi, 328
Spina bifida, 1169
  with meningomyelocele, 1169
  occulta, 1169
Spinal canal stenosis, 511
Spinal cord, 1115
  arterial supply, *1127*
  diseases/lesions *see* Spinal cord diseases/lesions
  posterior columns, 1115
  spinothalamic tracts, 1115
  subacute combined degeneration (SACD), 1175
  see also Spine
Spinal cord compression, 458, 1117–18, *1117*, 1163–4
  causes, 1164, *1164*
  management, 1164
Spinal cord diseases/lesions, 1116–18, **1163–6**, *1164*
  compression *see* Spinal cord compression
  disc lesions
    cervical *see* Cervical discs
    prolapse/protrusion, 506, 510-11, *511*
    see also Peripheral nerve disease
  epidural abscess, 1159–60
  hemiparesis, 1112
  HIV-related, 1158
  inflammatory, 1165
  metabolic, 1165
  multiple sclerosis, 1152
  posterior column, 1116–17
  toxic, 1165
  vascular, 1165
  see also Demyelination; Paraparesis; specific diseases
Spinal cord stimulation, 752
Spinal cord tumours, 1164, *1164*
Spinal muscular atrophies, 1167
Spinal reflex arc, 1115, *1115*
Spinal root(s), *1116*
  dorsal, lesions, 1116, *1116*
  lesions, 1116
  pain, 1116
  see also Nerve root
Spinal stenosis, 1177
Spine
  bamboo, 534, *534*
  cervical, rheumatoid arthritis, 527
  examination, *502*
  imaging
    angiography, 1120
    X-rays, 1119
  osteoporotic crush fracture, 512
  rugger jersey, 629
  see also Spinal cord
Spinocerebellar degeneration, 1170
Spinothalamic tracts, 1115
  lesions, 1117, *1117*
Spiral computed tomography
  common bile duct stones, 371
  pancreas, 376
  pancreatic carcinoma, 383, *383*
  pancreatitis, 378, *378*
  renovascular disease, 608
  urinary tract obstruction, 615

Spirometra mansoni, *163*
Spirometry, 825, *825*, *826*
  asthma, 852
Spironolactone, 653, *657*
  ascites, 352
  clinical uses, 657
  heart failure, *739*
  hirsutism, 1005
  hyperaldosteronism, 1024
  hypertension, 804
  myocarditis, 788
  side-effects, gynaecomastia, 352
Spleen, 387, **422–3**
  atrophy, 423
  enlargement *see* Splenomegaly
  functions, 422–3
    blood pooling, 423
    extramedullary haemopoiesis, 422
    immunological, 423
    sequestration and phagocytosis, 422
  palpable *see* Splenomegaly
  red pulp, 422
  removal *see* Splenectomy
  rheumatoid arthritis, 528
  white pulp, 422
Splenectomy, 69, 423
  autoimmune haemolytic anaemia, 416
  complications, 423
  haematological features after, 423
  hereditary spherocytosis, 406
  infection prophylaxis, 423
  myelofibrosis, 422
  pyruvate kinase deficiency, 414
  thrombocytic purpura, 436
Splenic infarction
Splenic sequestration, 410
Splenomegaly, 329, 331, 423
  causes, 423
  massive, 423
  myelofibrosis, 421, 423
  portal hypertension, 329, 423
  sickle cell anaemia, 410
Splicing mutations, 41
Splinter haemorrhages, 690
  infective endocarditis, *770*, *771*
Spondylitis, 535
Spondyloarthropathy
  seronegative, 522, **532–6**
    aetiology, 532
    ankylosing spondylitis *see* Ankylosing spondylitis
    non-articular problems, *533*
    psoriatic arthritis, 534–5, *534*
    reactive arthritis, 535–6
Spondylolisthesis, 511–12
  degenerative, 512
Spondylosis, 505, 1176
  cervical, 506, 1176
  lumbar, 510, 1176
Spongiform encephalopathy, transmissible *see* Prion diseases
Spontaneous bacterial peritonitis (SBP), 353
Sporothrix, *S. schenckii*, *82*, 149, 152
Sporotrichosis, *150*, 152
  lymphocutaneous, *82*
  plaque, 152
Spotted fevers, *147*
  diagnosis, 148
  treatment and prevention, 148
Spur cells, 325
Sputum, 819, 827
  asthma, 853
  bronchiectasis, 843
  colour, 819, 827
  in COPD, 837, 838
  cytology, 827
  tuberculosis, 865–6
Sputum culture, *89*
Squamous cell carcinoma (SCC), 1259, *1259*
  HIV-associated, 205
  larynx, 1083, *1083*
  oesophagus, 256
Squashed stomach syndrome, 496
Squint (strabismus), 1105
  cover test, 1105
  non-paralytic, 1105
  paralytic, 1105
Staghorn calculus, 611
Staging, chronic renal failure (CRF), *626*
Stamping gait, *1096*, 1097, 1117
Standards of conduct, General Medical Council (GMC), 1
Standards of disclosure, 4
  doctor, 4
  patient, 4
Stanols, 1065
Stapes, 1077
Staphylococcal food poisoning, *86*
Staphylococcal infections, *86*
  host factors, *86*
  see also specific staphylococcal species
Staphylococcal scalded skin syndrome, 125, 1228–9
Staphylococcal scarlet fever, *86*
Staphylococcus, *1179*
  S. aureus, *70*, 85, *90*, *123*, *191*
    antibiotic resistance, *94*
    bacterial meningitis, 134
    blepharitis, 1090
    bronchiectasis, 843
    CAPD peritonitis, *637*
    conjunctivitis, 1088
    food poisoning, *133*
    gastroenteritis, *128*, 130–1
    meticillin-resistant *see* MRSA
    osteomyelitis, 410
    pericarditis, 795
    pneumonia, *858*, 860
    pyogenic abscess, 361
    septic arthritis, 526, 539
    septicaemia, *135*, *136*
    sialadenitis, 248
    skin and soft tissue infections, 123–4, *124*
    vancomycin-insensitive *see* VISA
    vancomycin-resistant *see* VRSA
    ventilator-associated pneumonia, 914
  S. epidermidis, 70, *70*, 85, *90*, *123*
    CAPD peritonitis, *637*
    septicaemia, *136*
    urinary tract infection, *600*
  S. pneumoniae, osteomyelitis, 410
  S. saprophyticus, urinary tract infection, *600*
Starch, 211
Starling's law *see* Frank-Starling relationship
Starvation, 216
  adaptive processes, 213
  leptin levels in, 230
  metabolic response to, 213
  pathophysiology, 213–14, *213*
  thiamin deficiency, 221
  trauma/shock, 213
  see also Malnutrition; Undernutrition
Static myoclonic encephalopathy, 1150
Statins
  acute coronary syndromes, *754*
  contraindications, *1066*
  diabetes mellitus, 1044
  genetic polymorphism in metabolism, *926*
  hyperlipidaemia, 746, 1065, *1066*
  mechanism of action, *1066*
  myocardial infarction, 759
  renovascular disease, 608
  safety, *1066*
  side-effects, *1066*
  see also individual drugs
Statistical analyses (clinical trials), 932–3
  average, 932–3
  hypothesis testing, 933
Statistical significance, 932–3
Stature
  short *see* Growth, failure
Status asthmaticus, 856–7
Status epilepticus, 1142, *1142*
Stavudine, *192*
  lipoatrophy, 198
Stearic acid, *210*
Steatohepatitis, 344
  non-alcoholic, 344
Steatorrhoea, 244, 275, 277, 281
  pancreatic insufficiency, 381
Steatosis, alcoholic liver disease, 358
Steele-Richardson-Olzewski syndrome, 1148
Stein-Leventhal syndrome, *228*, 1003
Stellate cells, 320
  alcohol effects, 358
  cirrhosis, 345
  liver injury, 349
Stem cell diseases, 389
Stem cell factor, haemopoiesis, 388
Stem cell therapy, 47
Stem cell transplantation (SCT)
Stem cells, 29, 37
  adult, 29
  cord blood, 29
  disorders
  embryonic, 29
  multipotent, 29
  pluripotent, 29, 387–8
    self-renewal, 388
  totipotent, 29
  unipotent, 29
STEMI *see* Myocardial infarction, ST-elevation (STEMI)
Stenosis
  aortic *see* Aortic stenosis
  mitral *see* Mitral stenosis
  pulmonary *see* Pulmonary stenosis
Stents
  coronary artery, 750–1, *751*, 758
  drug-eluting, 750–1, *751*
  endovascular, 808
  tracheobronchial, 883
'Step-down' units, 889
see also Intensive care
Stercobilinogen, 323
Sterility, chemotherapy-induced, 462
Sternoclavicular joint, rheumatoid arthritis, 527
Steroid acne, 1246
Steroids
  side-effects, hepatotoxicity, 365
  specific conditions
  topical
    atopic eczema, 1237–8
    classification, *1237*
    safety of, 1270
    see also Corticosteroids; Glucocorticoids; individual drugs
Stevens-Johnson syndrome, clinical spectrum, *1251*
Stiff person syndrome, *455*, 1182
Stiffness, 501
Still's disease
  adult, 553
    pyrexia, *91*
  children, 553
Stings
  insects, 951
  jellyfish, 950
  marine animals, 950
  scorpions, 951
Stix testing, 478–9
Stokes-Adams attacks, 689, 719
Stomach, **258–67**
  acid secretion, 258, *259*
  anatomy, 258
  antral mucosa, 258
  function, 258
  mucosal barrier, 258
  mucosal lining, 258
  mucus-secreting cells, 258
  tumours *see* Gastric cancers/tumours
  ulcers *see* Gastric ulcers
  see also entries beginning gastric/gastro-
Stomatitis, *105*, 246
  angular, 246
Stomatocytosis, hereditary, 406
Stool
  cultures
  examination, 244
  rabbit dropping, 244
  ribbon-like, 244
  watery, *see* Cholera; Diarrhoea *see also* Faecal
Stool antigen test, 260
Storage pool disease, 437
Stratum corneum, 1225
Strawberry cervix, 182
Strawberry naevus, 1267–8, *1268*
Strawberry tongue, 126
Streptococcal infections
Streptococci
Streptococcus, *70*, *1179*
  S. agalactiae, *123*
  S. milleri, pyogenic abscess, 360
  S. mutans, dental caries, 248
  S. pneumoniae, *70*, 84, *90*, *123*, *191*
    antibiotic resistance, *94*
    bacterial meningitis, 134
    conjunctivitis, 1088
    haemolytic uraemic syndrome, 599
    pneumonia, *858*
    respiratory infection, 410
    septicaemia, *135*
    sialadenitis, 248
  S. pyogenes, 85, *123*, 136
    antibiotic resistance, *94*
    sialadenitis, 248
    skin and soft tissue infections, *124*
Streptogramins, 124
Streptokinase
  arterial thrombosis, 443–4, *443*
  myocardial infarction, 757
  pulmonary embolism, 786
Streptomycin, 96, *96*
  side-effects, 867, *928*
    acute tubular necrosis, 617
Streptozotocin, pancreatic neuroendocrine tumours, 384
Stress, 799
  related hormones, 967, 968
  serum prolactin, *981*
Stress cardiomyopathy, 794, *794*
Stress echocardiography, heart failure, 737
Stress (march) fractures, 515
Stress reactions, 1209–10
  acute, 1209
  adjustment disorder, 1209
Stress ulceration of stomach, 936
Striae, 1263
Striatonigral degeneration, 1148
Stridor, 819, 833, 1083
Stroke, **1126–36**
  completed, 1126
  definition, 1126
  elderly patients, 1133
  in evolution, 1126
  imaging, 1131
  immediate care, 1130, *1130*
  incidence, *1096*
  investigations, 1130
  management, 1131–3, *1132*
  minor, 1126
  pathophysiology, 1126

Stroke (*Continued*)
prevention, 1126–7
prognosis, 1133
rehabilitation, 1133
risk factors, 1126–7, *1127*
surgery, 1133
thrombolysis, 1130, *1131*
vascular anatomy, 1127–8, *1127*
Stroke volume, 684, 891–2
afterload, 892, *892*
myocardial contractility, 891, *892*
preload, 684, 891, *891*
Stroke work, 684
Stromal cells, 387
Stromelysins, 500
Strongyloides
eosinophilia, 875
HIV-related, 204
rash, *88*
S. stercoralis, 80, *163*, 167
Strongyloidiasis, 167
treatment, *163*
Strontium ranelate, osteoporosis, 563, *564*
Stropharia, 952
Student's *t*-test, 933
Stupor, 1123
causes, 1123–4, *1123*
Sturge-Weber syndrome, 1170
Subacute sclerosing panencephalitis, 1158
Subarachnoid haemorrhage (SAH), *1102*, 1134–5
causes, 1134–5, *1134*
complications, 1135
differential diagnosis, 1135
investigations, 1135, *1135*
management, 1135
Subcortical encephalopathy, 1168
Subdural empyema, 1159
Subdural haematoma, 1135
incidence, *1096*
Subfertility, 963, 1006–7
causes, 1006, *1007*
clinical assessment, 1006
female factors, 1006, *1007*
investigations, *1007*
male factors, 1006, *1007*
treatment, 1007
Sublimation, 1189
Submucosal plexus, 271
Subphrenic abscess, 317
Substance P, *272*, 814, 1098, 1113
asthma, 852
Substantia nigra, 1113, *1113*
Subthalamic nuclei, 1113, *1113*
Succinylcholine, genetic polymorphism in metabolism, *926*
Succussion splash, 243, 262
Sucralfate, variceal bleeding, 351
Sucrose, 211
Sudden death, 759–60
causes, *711*, *760*
see also Cardiac arrest; Cardiac failure
Sudden infant death syndrome, *37*, 119
Sudek's atrophy, 517
Suicide, 1206, *1206*
self-poisoning, *see* Self-poisoning
Sulbactam, 95
Sulconazole, 100, *100*
Sulphadiazine, *93*
nephrotoxicity, 189
side-effects, renal calculi, 611
toxoplasmosis, 160, 201
Sulphapyridine (sulphasalazine)
psoriatic arthritis, 534
rheumatoid arthritis, 530, *530*
side-effects
haemolytic anaemia, 415
hepatotoxicity, *366*
ulcerative colitis, 291–2
Sulphasalazine *see* Sulphapyridine (sulphasalazine)
Sulphate
interstitial fluid, *650*
intracellular fluid, *650*
plasma, *650*
Sulphonamides, *94*, 98
cytochrome P450 inhibition, *925*
indications, 98
interactions, 98
mechanism of action, 98
resistance, 98
side-effects, *928*
acute tubulointerstitial nephritis, *604*
haemolysis in G6PD deficiency, *414*
hepatotoxicity, *366*
rash, *1265*
respiratory, *877*
specific conditions, dermatitis herpetiformis, 1256
structure, 98
toxicity, 98
see also individual drugs
Sulphonylureas
diabetes, 1038–9, *1038*
side-effects
endocrine disease, *964*
hypoglycaemia, 1060
rash, *1265*
Sulphur dioxide, 830, *830*, 959
Sulphuric acid, 959
Sumatriptan, migraine, 1137
Sun protection factor, 1248
SUNA, 1138
Sunbeds, 1248
Sunblocks, 1248
SUNCT, 1138
Sunitinib, *464*
renal cell carcinoma, 645
Sunlight, 1248
see also Ultraviolet light
Superior vena caval obstruction, 458
Superiority trials, 929
analysis, 930
Superoxide dismutase, 20, *224*
Superoxide radical, 20
Suppositories, *297*
Supraventricular tachycardias (SVTs), 715, 720–6
acute management, 722–3
see also Antiarrhythmic drugs
AV nodal re-entry, *703*, 721, *721*, *722*
AV reciprocating, 721–*2*, *721*
with bundle branch block, *726*
causes, *721*
long-term management, 723
sinus *see* Sinus tachycardia
symptoms, 722
see also Atrial tachyarrhythmias; Ventricular tachyarrhythmias
Suramin, 158
side-effects, neuropathy, *1174*
Surfactant, 811, 817
Surfactant protein A, 818
Surfactant replacement therapy, ALI/ARDS, 919
Surgery
bariatric, 232–3
blood ordering, 425
breast cancer, 488
bypass, 751–*2*, *751*, 758
cardiac, infective endocarditis, 773
in diabetes mellitus, 1058
eye, 1087
functional endoscopic sinus, 1081
gastrointestinal
bariatric, 232–3, 1039–40
colorectal carcinoma, 304–5
Crohn's disease, 290
gastro-oesophageal reflux, 252
peptic ulcer disease, 262
ulcerative colitis, 292–3, *292*
hyperprolactinaemia, 982
metabolic response, 898
musculoskeletal
osteoarthritis, 522
Paget's disease, 566
rheumatoid arthritis, 532
nervous system *see* Neurosurgery
ovarian cancer, 492
psychosurgery, 1203, 1211
respiratory, COPD, 840–1
steroid cover, *1016*
stroke, 1133
testicular cancer, 492
thyroid eye disease, 992
see also specific procedures
Suxamethonium
side-effects, *928*
hypokalaemia, 667
Swallowing, 248, 249
difficulty, *see* Dysphagia
Swan-Ganz catheter, 714, 902
Sweat electrolytes, 843, 844
Sweat glands, 1227
apocrine, 1227
eccrine, 1227
Sweating, 954
inhibition of, 954
Swelling
joints, 501
see also Oedema
Swimmer's itch, 169
Sycosis barbae, 1228
Sydenham's chorea (St Vitus's dance), 137, 1149
Synacthen test, Addison's disease, 1011
Synapse, 1098, *1098*
Synaptic transmission, 1098
Synaptopodin, 597
Syncope, 689, 717, 1144
cardiac, 1144
cardiovascular causes, *689*
carotid sinus, 689
causes, *1144*
congenital heart disease, 774
cor pulmonale, 783
exercise-induced, 1144
aortic stenosis, 765
Fallot's tetralogy, 778
investigation, 1144
management, 1144
micturition, 689, 1144
pulmonary stenosis, 768
vasovagal, 717
ventricular tachycardia, 726
Syndecans, 500
Syndesmophytes, ankylosing spondylitis, 533, 534
Syndrome of apparent mineralocorticoid excess, 1024
Syndrome of inappropriate ADH secretion (SIADH), *455*, 654, 663, 1018–19
causes, *1018*
clinical features, 1018
diagnosis, 1018
treatment, 1018–19
Syndrome X (metabolic syndrome) *see* Metabolic syndrome (syndrome X)
Synergy, 54
Synovial fluid, 499
aspiration, 504, 522
examination, 504
microscopy, 538
Synovial inflammation, 520
Synovial joints, *500*
disease, *see* Arthritis; Gout; Rheumatic diseases
Synoviocytes, 499
Synovitis
chronic inflammatory, 518
crystal, 518
hip, 554
Synovium, 499
Syphilis, **177–9**
cardiovascular, 809
classification, *178*
clinical features, *178*
congenital *see* Congenital syphilis
diagnostic investigations, 178–9
endemic (bejel), 142
genital ulceration, *172*
Jarisch-Herxheimer reaction, 179
latent, 178
neurological disease *see* Neurosyphilis
rash, *88*
rheumatoid factor, *503*
secondary, 178, *178*
serology, *179*
tertiary, 178
transfusion-transmitted, 428
Syphilitic arthritis, 540
Syringobulbia, *1110*, 1164–5
Syringomyelia, 1117, 1164–5
Charcot's joints, 555
Syrinx, 1165, *1165*, 1169
System theories of ageing, *227*
Systemic amyloid, 1254
Systemic amyloidosis, postural hypotension, *659*
Systemic inflammatory response syndrome (SIRS), 135, 899, *900*
definition, *899*
Systemic inflammatory vasculitis, **548–52**
large vessel, 549–50
medium-sized vessel, 550–1
sites of, *549*
small vessel, 551–2
types of, *549*
see also individual vasculitides
Systemic lupus erythematosus (SLE), 435, **541–4**, 1225, 1252–3
aetiology, 541–2
antinuclear antibodies, *504*
autoantibodies, 543, *544*
autoantigens, *74*
clinical features, 542–3, *542*
butterfly rash, 542–3, 1252, *1253*
eyes, 543
gastrointestinal system, 543
heart and cardiovascular system, 543
joints and muscles, 542
kidneys, 543
lungs, 543
nervous system, 543
pulmonary hypertension, 783
pyrexia, *91*, 542
renal tubular acidosis, *676*
skin, 542–3
course and prognosis, 544
diagnosis, *69*
disease associations, *544*
epidemiology, 541
investigations, 543
lupus glomerulonephritis, 590–1, *590*
management, 544
pathogenesis, 542
pathology, 542
in pregnancy, 544
respiratory involvement, 871
rheumatoid factor, *503*
Systemic sclerosis (SSc), 283, *545*, **545–7**, 1252
antinuclear antibodies, *504*
autoantibodies, 546
clinical features, 546
environmental risk factors, 545
hypertension, 608–9
investigations, 546
management, 546–7
oesophagus in, 254
pathology and pathogenesis, 545–6, *545*
prognosis, 547

respiratory involvement, 871–2
rheumatoid factor, *503*
see also Scleroderma
Systemic vasculitides
Systemic/multisystem infections, 135–40
bacteraemia and sepsis syndrome, 135–6, *135*, *136*
brucellosis *see* Brucellosis
leptospirosis *see* Leptospirosis
listeriosis, 138–9
Lyme disease *see* Lyme disease
meningococcal septicaemia *see* Meningococcal septicaemia
Q fever, 139
rheumatic fever, 136–7, *137*
tularaemia, 139
Systolic blood pressure

T cell(s), 53
activated, 64–6
dendritic cells, 59, 65
rheumatoid arthritis, 523
antigen receptors, 63
CD4 *see* T cell(s), helper
CD8 *see* T cell(s), cytotoxic
cytotoxic, 58, *64*
antigen presentation, *62*
functions, 66
killing mechanisms, 66
deficiency, *69*, 70
development, 64–6, *64*
helper, 58, *64*, *65*, 185
in autoimmune disease, 74
chronic inflammation, 68
functions, 65
licensing, 65
measurement, *69*
Th1, *65*
Th2, *65*
Th17, *65*
HIV replication in, *186*, 190
interactions with antigen-presenting cells, *69*
naive, 65
origins of, 388
regulatory (Treg), 65–6, *65*
in autoimmune disease, 74
subtypes, 66
small intestine, 275
thymic selection, 64
tolerance *see* Tolerance
see also HLA system
T cell lymphoma, 483
cutaneous, 1260
enteropathy-associated, 279
T cell receptors, 62, 64
in autoimmune disease, 73–4
negative selection, 64
positive selection, 64
T-helper cells *see* T cell(s), helper
T-lymphocytes *see* T cell(s)
Tabes dorsalis, 1157
Tabun, 946
Tachy-brady syndrome, 717
Tachycardia cardiomyopathy, 794
Tachycardias, 659, 891
accelerated junctional, *721*
atrioventricular junctional, 721–3, *722*
definition, 715
paroxysmal, 689
sinus *see* Sinus tachycardia
supraventricular *see* Supraventricular tachycardias (SVTs)
ventricular, 715
Tachypnoea, 688, 820
Tacrolimus, 76
atopic eczema, 1238
autoimmune hepatitis, 344
renal transplantation, 640
side-effects, haemolytic uraemic syndrome, 598
vitiligo, 1264
Tadalafil, erectile dysfunction, 1000, 1056
Taenia
eosinophilia, 875
T. saginata, *163*, 170
T. solium, *163*, 170–1, 1157
Takayasu's arteritis, *549*, 550
Takayasu's disease, 809
Tako-tsubo/octopus pot cardiomyopathy, 794, *794*
Tall stature, 979
Tamm-Horsfall glycoprotein, 599
Tamm-Horsfall protein, 600, 610
Tamoxifen, 463, 488
retroperitoneal fibrosis, 617
side-effects, 463, 488, *488*
Tamsulosin, benign prostatic hypertrophy, 646
Tangier disease, 1064, *1064*
Tapeworms *see* Cestodes (tapeworms)
Tardive dyskinesia, neuroleptic-induced, 1148
Target cells, 325, *409*, 423, *423*
Tarsal tunnel syndrome, 515, *515*
rheumatoid arthritis, 528
Tarui disease, *1069*
TATA box, 30
Taurine, 655
Taxanes, 460
side-effects, neurotoxicity, 462
Tazarotene, psoriasis, 1242
TBid, 28, *28*
Tear film, 1085
Teeth, 248
decay *see* Dental caries
erosion, 248
Tegafur, 460
Teicoplanin, 99
Telangiectasia, 546, *1227*, 1247
Teletherapy, 465
Telithromycin, 97
Telmisartan, hypertension, 804
Telogen effluvium, 1267
Telomerase, 226
Telomeres, 37
loss of, 226
Telopeptides, 569
Temocillin, 95
Temozolomide, 460
Temperature
body core, 954
environmental, 953
in poisoning, 934
unconscious patient, 1124
Temporal arteritis *see* Giant-cell arteritis
Temporal artery biopsy, 550
Temporal lobe seizures *see* Epilepsy
Temporomandibular joint, rheumatoid arthritis, 527
Temporomandibular pain dysfunction syndrome, 517
Temsirolimus, renal cell carcinoma, 644
Tenase, 432
Tendon reflexes, 1112
Tendon sheath, 499
Tendon xanthomata, 556
Tendonitis, Achilles tendon, *515*
Tendons, 500
Tenecteplase
arterial thrombosis, *443*, 444
myocardial infarction, 757
Tennis elbow, 508, *508*
Tenofovir, *192*, *195*
and Fanconi's syndrome, 189
hepatitis B, 339
Tenosynovitis, 508
rheumatoid arthritis, 527
Tensilon (edrophonium) test, 1180
Tension headache, 1136
Tension pneumothorax, 914
Teratogens, 199, 219, *924*, 1217
Teratoma, 492
Terbinafine, 100, *100*
ringworm, 1233
Terbutaline, asthma, 854, *854*
Teriparatide, osteoporosis, 563, *564*
Terlipressin, variceal bleeding, 350
Termination mutations, 42
Testes/testicles
Testicular cancer, 492
histology, *492*
investigations and surgery, 492
symptoms and signs, 492
treatment, 492
Testicular failure, *964*, *965*
Testicular feminization, *1007*
Testicular synthetic defects, *1007*
Testosterone, *971*
chronic renal failure, 630
PCOS, 1004
replacement therapy, *976*, *999*
Tetanus, *82*, *83*, 134–5
cephalic, 134
clinical features, 134
diagnosis, 134
immunization, *102*, *103*, 135
localized, 134
management, 134–5
mortality, *80*
neonatal, 134
prevention, 135
Tetany, *1022*
Tetracyclines, *94*, 96–7, *97*
aphthous ulcers, 247
balantidiasis, 162
Chlamydia, 176
gastroenteritis, *132*
H. pylori eradication, 261
indications, 97
interactions, 97
rosacea, 1247
side-effects
hepatotoxicity, *366*
rash, *1265*
respiratory, *877*
uraemia, 619
spotted fevers, 148
structure, *97*
toxicity, 97
see also individual drugs
Tetrahydrofolate (THF), 400
see also Folates/folic acid
Tetraparesis, 1112, 1117, *1128*
Tetrazines, 460
Tetrodotoxin, 271
[^{99}Tc]Tetrofosmin, 705, *705*
Thalamic lesions, 1118
Thalamus, 1113, *1113*
Thalassaemia(s), 41, 393, **407–9**
α-thalassaemia, 409, *409*
four-gene deletion, 409
one-gene deletion, 409
three-gene deletion, 409
two-gene deletion, 409
β-thalassaemia, 407–9, *407*
clinical syndromes, 407–8, *408*
haemoglobin patterns, *407*
homozygous, 407
management, 408–9
molecular genetics, 407
hepatomegaly, *331*
intermedia, 407–8
major (Cooley's anaemia), 408
back pain, 556
minor *see* Thalassaemia trait
molecular genetics, 42
Thalassaemia trait, *396*, 407
Thalidomide
aphthous ulcers, 247
Behçet's disease, 552
chronic discoid lupus erythematosus, 1252
familial Mediterranean fever, 556
myeloma, 486
Thallium poisoning, antidotes, *937*
Thelarche, premature, 996
Theophylline
asthma, *854*
overdose, *935*, 949
side-effects, renal calculi, 611
therapeutic drug monitoring, *927*
Therapeutic drug monitoring, 92, 927, *927*
Therapeutic ratio, 925
Thermogenesis
dietary induced, *208*
obesity, 230
Thevetia peruviana, 952
Thiamazole, 989
Thiamin diphosphate *see* Thiamin pyrophosphate (TPP)
Thiamin pyrophosphate (TPP), 220–1
Thiamin (vitamin B_1), 220–1
deficiency, 221
alcohol dependence, 221, 238–9, *239*
see also Wernicke-Korsakoff syndrome
function, 220–1
recommended IV intake, *234*
RNI, *218*
side-effects, respiratory, *877*
sources of, 221
structure, *220*
Thiazide diuretics, *657*
clinical uses, 657, *658*
nephrotic syndrome, 585
diabetes insipidus, *976*, 1018
hypercalciuria, 613
hypertension, 803–4
side-effects
acute tubulointerstitial nephritis, *604*
rash, *1265*
renal calculi, 611
Thiazolidinediones
diabetes mellitus, 1039
menstrual disorders, 1005
non-alcoholic fatty liver disease, 345
overdose, 940
sodium retention, 656
6-Thioguanine, 460
Thiopurine methyltransferase, 290, 926
genetic polymorphism, *926*
Thioridazine, poisoning, *935*
Thiotepa, bladder tumours, 645
Third parties in communication, 18
Thirst, 1036
Thirst axis, 1016–19
Thomsen's disease, 1182
Thoracic discs, protrusion, 1177
Thoracic outlet syndrome, 1177
Thoraco-abdominal aneurysm, 808, *808*
Thought block, 1187
Thought broadcast, 1187
Thought content, 1187
Thought disorders
form of thought, 1187
obsessional rumination, 1187
stream of thought, 1187
Thought insertion, 1187
Thought stopping, 1211
Thought withdrawal, 1187
Threadworm (*Enterobius vermicularis*), *163*, 167, 299
Threonine, 210
Throat, **1082–5**
anatomy and physiology, 1082
disorders, 1082–5
dysphagia *see* Dysphagia
glandular fever, 1083
hoarseness, 1082–3
pharyngitis, 1083–4
quinsy, 1084
snoring, 1084, *1084*
stridor, 1083
tonsillitis, 1083
examination, 1082
see also specific components
Throat swabs, *89*
Thrombin, 433

Thrombin activatable fibrinolysis inhibitor, 434
Thrombin inhibitors, 445
Thrombin time (TT), 435
  DIC, 441
Thromboangiitis obliterans, 806, 809
Thrombocythaemia
  essential, 421, 438
    splenic atrophy, 423
  musculoskeletal problems, 503
Thrombocytopenia, *88*, 435–6
  alcohol-induced, 325
  autoimmune, 416
  causes, *436*
  chronic renal failure, 627
  drug-induced, 436–7
  heparin-induced, 444–5
  HIV-related, 189
  immune thrombocytopenic purpura, 436
    clinical features, 436
    investigation, 436
    treatment, 436
  liver disease, 440
  musculoskeletal problems, 503
  neonatal alloimmune, 437
Thrombocytopenic purpura *see* Immune thrombocytopenic purpura
Thromboembolic disease, 442
Thromboembolism
  tachy-brady syndrome, 717
  see also Thrombosis
β-Thromboglobulin, 432
Thrombolysis In Myocardial Infarction (TIMI) score
  acute coronary syndromes, 752, *753*
  ST elevation myocardial infarction, *756*
Thrombolytic therapy
  arterial thrombosis, 443–4, *443*
  Budd-Chiari syndrome, 360
  contraindications, *758*
  deep vein thrombosis, 810
  stroke, 1130, *1131*
  venous thrombosis, 444–7
  see also Anticoagulants/anticoagulation
Thrombomodulin, 431, 687
Thrombophilia, 44, 442
Thrombophlebitis
  migratory, *455*
  superficial, 809
Thrombopoietin, 389
Thromboprophylaxis, 446–7
Thrombosis, 442–7
  anti-thrombotic mechanisms, 687
  arterial *see* Arterial thrombosis
  atherosclerotic plaques, *744*, *745*
  cortical venous, 1136
  drug treatments, *see* Anticoagulants/anticoagulation; Thrombolytic therapy
  dural venous sinus, 1136
  pro-thrombotic mechanisms, 687, *744*
  venous *see* Venous thrombosis
Thrombospondin, 432
Thrombotic disease, 807
Thrombotic thrombocytopenic purpura, 437, 599
Thromboxane(s)
  allergic response, *72*
  haemostasis, 432
Thromboxane A2, 577
Thrombus, 442
  red, 442
  white, 442
  see also Thromboembolism; Thrombosis
Thunderclap headache, 1135
Thymectomy, 1180
Thymic education, 74
Thymic selection, 64
Thymine, 29, *30*
Thymoma, 888
Thymus, 387
  T cell development, 64, *64*
Thyroglobulin, *493*, 983
Thyroid acropachy, 988, 1254
Thyroid antibodies, 993
Thyroid carcinoma, *993*, **993–4**
  anaplastic, *993*, 994
  follicular, *993*, 994
  lymphoma, *993*, 994
  medullary, *993*, 994
  papillary, *993*, 994
Thyroid crisis/storm, 990
Thyroid disease, 963
  cancer *see* Thyroid carcinoma
  eye *see* Thyroid eye disease
  myopathy, 1179
  neuropathy, 1174
  see also Goitre; Hyperthyroidism; Hypothyroidism
Thyroid eye disease, 990, 991–2
  clinical features, 991
  investigations, 991
  pathophysiology, 991
  treatment, 991–2
Thyroid function tests, 984, *984*
  goitre, 993
  growth failure, 978
  interpretation of, 984
Thyroid gland, **982–94**
  anatomy, 982–3
  colloid *see* Thyroglobulin
  enlargement *see* Goitre
  overactivity *see* Hyperthyroidism
  physiology, 983–4, *983*
  retrosternal/intrathoracic, 887–8
  underactive *see* Hypothyroidism
Thyroid hormones, 967
  drugs affecting, 984
  metabolism, 983, *983*
  physiological effects, 983–4
  in pregnancy, 984
  resistance to, 991
  in serious illness, 984
  synthesis, 983, *983*
    defects, 985
  see also Thyroid-stimulating hormone (TSH); Thyroxine (T4); Tri-iodothyronine (T3)
Thyroid peroxidase, 984
Thyroid transcription factor, *493*
Thyroid-binding prealbumin (TBPA), 983
Thyroid-stimulating hormone (TSH), 967, 970, *971*, 983
  deficiency, thyroid function tests, *984*
  hyperthyroidism, 988
  hypothyroidism, 985
  measurement, 984
  receptor, IgG antibodies, 984
Thyroid-stimulating immunoglobulin (TSI), 989
Thyroidectomy, 990
Thyroiditis, 992
  de Quervain's, 988, 992
  Hashimoto's, *74*, 541, 985
  postpartum, 985
  renal tubular acidosis, *676*
  Riedel's, 993
  see also Graves' disease
Thyrotoxicosis *see* Hyperthyroidism (thyrotoxicosis)
Thyrotoxicosis factitia, *964*
Thyrotrophin-releasing hormone test, 984
Thyrotrophin-releasing hormone (TRH), 967, 970, *971*, 983
Thyroxine (T4), 967, *971*
  metabolism, 983, *983*
  side-effects, endocrine disease, *964*
  synthesis, 983, *983*
Thyroxine-binding globulin (TBG), 983
Tiabendazole, *166*, 168
Tiagabine, overdose, 940
Tic douloureux see Trigeminal neuralgia
Ticarcillin, 95
Tick-borne encephalitis, *82*, *110*, 116
Tick-borne flavivirus encephalitis, 1156
Tick-borne relapsing fever, *82*, 147
Ticlopidine
  acute coronary syndromes, 753, *754*
  side-effects, thrombotic thrombocytopenic purpura, 437
Tics, 1150
Tidal volume, respiratory failure, 911
Tietze's disease, 516
Tigecycline, 124, 128
Tight junctions, *25*, 26
Tilt testing, 697, *698*
Tiludronate, Paget's disease, 566
Time, clinical communication difficulties, 11
TIMI risk score, *756*
Timolol, glaucoma, 1091
Tinea
  incognito, 1233, 1261
  T. capitis, 1223
  T. corporis, 1232
  T. cruris, 1232, *1232*
  T. faciei, 1232
  T. manuum, 1223
  T. pedis, 1232
  T. unguium, 1233, *1233*
  treatment, 1233
Tinel's sign, 508, 515, 1116
Tingling, 1116
Tinidazole, 99
  amoebiasis, 161
  giardiasis, 162
Tinnitus, 1080, 1108
Tinzaparin, venous thromboembolism, 447
Tioconazole, 100, *100*
Tiopronin, 613
Tiotropium, COPD, 839
Tipranavir, *193*, 196
Tirofiban
  acute coronary syndromes, *754*
  angina pectoris, 750
  arterial thrombosis, *443*
Tissue dysoxia, 897
Tissue factor, 433, 897
Tissue factor pathway inhibitor, 433
Tissue inhibitors of metalloproteinases (TIMPs), 325, 345, 452, 501
  osteoarthritis, 520
Tissue perfusion, 900, *900*
Tissue-type plasminogen activator, 434
  arterial thrombosis, *443*, 444
  myocardial infarction, 757
Titanium dioxide, 1248
Titin, 788
Titratable acid, 674, *674*
Titubation, 1114
Tobacco, and cancer, 449
Tobramycin, 96, *96*
Tocilizumab, rheumatoid arthritis, *530*, 531
α-Tocopherol, 220
Tocopherols, 220
Tocotrienoles, 220
Todd's paralysis, 1140
Togaviruses, *110*, 113–14
Tolbutamide
  diabetes, *1038*
  genetic polymorphism in metabolism, *926*
  side-effects, bone marrow aplasia, 402
Tolerance
  loss *see* Autoimmunity (type II hypersensitivity)
  oral, 275
  restoration of, 76
Toll-like receptors, *59*, 86, 275, 819, 847
  TLR2, *59*
  TLR3, *59*
  TLR4, *59*, 895
  TLR7, *59*
  TLR9, *59*
Tolnaftate, 100, *100*
Tolvaptan, *657*, 663
  autosomal-dominant polycystic kidney disease, 644
  SIADH, 1019
Tone
  asymmetry, 1125
  increased, 1112
Tongue, 247
  black hairy, 247
  geographic, 247
  glossitis, 247
    folate deficiency, 401
    pernicious anaemia, 399
  macroglossia, 1072
  raspberry, 126
  strawberry, 126
Tonsillar biopsy, 1121
Tonsillectomy, 1084, *1084*
Tonsillitis, 833, 1083
Tophi, 538, *539*
Topical therapy, 1270
Topiramate
  epilepsy, *1143*
  overdose, 940
    metabolic acidosis, *939*
Topoisomerase II inhibitors, side-effects, leukaemia, 467
Topoisomerases, 460
Topotecan, 460
Torasemide, *657*
Torcetrapib, 1065, *1066*
Torsades de pointes, 97, *728*
Torticollis, spasmodic, 1148, 1150
Total body water, 649
Total cholesterol *see* Cholesterol, total
Total iron-binding capacity, 346
  microcytic anaemia, *396*
Total lung capacity (TLC), 815, 825
Total proteins, 324
Totipotentiality, 47
Toxic confusional state *see* Delirium
Toxic epidermal necrolysis, 1229
  clinical spectrum, *1251*
  drug-induced, *1265*, 1266
Toxic erythema of newborn, 1268
Toxic megacolon, 291, *291*
Toxic neuropathies, 1174, *1174*
Toxic shock syndrome, *86*, 125
  rash, *88*
Toxocara canis, *163*, 168
Toxocariasis, *80*, 168
  rash, *88*
Toxoplasma
  in organ rejection, 75
  T. gondii, *70*, 85, *90*, 160, *190*, *1179*
Toxoplasma myositis, *1179*
Toxoplasmosis, *80*, 160
  clinical features, 160
  diagnosis, 160
  hepatitis, 342
  HIV-associated, 201
  management, 160
  pyrexia, *91*
Trace elements *see* Minerals
Trachea, 811–12
  rupture, 884
Tracheal carcinoma, 883

Tracheostomy, 913–14, 1083, *1083*
  complications, *914*
  decannulation, 917
  laryngitis, 833
Trachoma, 143
TRAIL, 28
Tramadol, pancreatitis, 379, 381
Trandolapril, hypertension, 804
Tranexamic acid, 428
  hyperfibrinolysis, 441
Tranquillizers
  misuse, 1215
  see also individual drugs
Trans-migration, 67
Trans-septal balloon valvotomy, 762
Transaminases *see* Aminotransferases
Transbronchial biopsy
  sarcoidosis, 870
  tuberculosis, 865
Transcobalamin, 399
Transcobalamin II, 399
Transcranial magnetic stimulation (TMS), 1203
Transcription, 29–30, *31*
Transepidermal electrical nerve stimulation (TENS)
  angina pectoris, 752
  diabetic neuropathy, 1055
  pain, 1118–19
Transfer coefficient, 826
Transfer factor, 826
Transferrin, 321, *390*, 395
Transferrin receptors, 356
  microcytic anaemia, 396, *396*
Transforming growth factor (TGF-β), *65*, 66, 451
  asthma, 852
  cirrhosis, 345
  haemopoiesis, 389
  osteoarthritis, 520
Transforming growth factor (TGF-α), asthma, 852
Transfusion haemosiderosis, 408
Transfusion reactions
  haemolytic, 427
  non-haemolytic (febrile), 427
Transfusion-associated graft-versus-host disease, 426
Transfusion-related acute lung injury (TRALI), 426, 427
Transient cerebral ischaemia, 1144
Transient global amnesia, 1128
Transient ischaemic attacks (TIAs), 1126, 1128–9
  clinical findings, 1128–9, *1129*
  differential diagnosis, 1129
  features, 1128
  pathophysiology, 1126
  prognosis, 1129
Transjugular intrahepatic portocaval shunt (TIPS), 351, 353
  Budd-Chiari syndrome, 360
Transketolase, 221
Translation, 29–30, *31*
Translocations, 39
  reciprocal, 39
  Robertsonian, 39
Transmembrane receptors, 22
Transmissible spongiform encephalopathies (TSE), **121–2**, 1073, 1158
  see also Prion diseases
Transmyocardial laser revascularization, 752
Transoesophageal echocardiography, 702
Transplantation
  allogeneic stem cell, 462–3
  bone marrow *see* Bone marrow transplantation
  cardiac *see* Cardiac transplantation
  heart-lung, 846
  liver *see* Liver transplantation
  lung, 846
  organ rejection, 75, *75*
  renal *see* Renal transplantation
Transport, cell membrane, 21–2
Transposition of great arteries, *774*, 779–80, *779*
  clinical features, 779
  investigations, 780
  treatment, 780
Transthoracic echocardiography, 701
Transthyretin, 1072
Transtracheal aspiration, 827
Trastuzumab, *464*, 923
  breast cancer, 488, 489
Trauma
  brain *see* Brain damage/injury
  chest wall, 884
  cornea, 1090, *1090*
  mechanical ventilation, 913
  metabolic response, 214, 898
  noise-related, 1079
  see also Abuse
Travel, 960
  jet-lag, 960
  motion sickness, 960
Travellers
  immunization, 103
  infections, 89–90, *90*
Travellers' diarrhoea, *90*, 131, 307
  antibiotics, *132*
  causes, *131*
Travoprost, glaucoma, 1091
Trazodone, 1202
Treatment adherence, 9
Treatment evaluation, 8
Trefoil peptides, 275
Treg *see* T cell(s), regulatory (Treg)
TREM-1, 59, 895
Trematodes (flukes), *163*, 168–70
  food-borne flukes, 170
  liver flukes, 170
  water-borne flukes, 168–70
    schistosomiasis, 168–70
Tremor, 1111, 1112, 1114
  essential, *1114*, 1149
  initiation, 1114
  intention, 1114
  Parkinson's disease, 1146
  postural, 1114
  rest, 1114
Trench fever, 148, *148*
Trench foot, 955
Treponema
  T. carateum, 142
  T. endemicum, 142
  T. pallidum, *123*, 177
    urethral discharge, *172*
    see also Syphilis
  T. pertenue, 142
Treprostinil, pulmonary hypertension, 783
Triage, 8
Triamcinolone, 247
  asthma, 855
  keloid scars, 1264
Triamterene, 656, *657*
  clinical uses, 657
  side-effects, renal calculi, 611
Triazoles, 100, *101*
Trichiasis, 143
Trichinella
  eosinophilia, 874
  T. spiralis, *163*, 168, *1179*
Trichinosis, *80*, 168, *1179*
Trichloroacetic acid, human papillomavirus, 182
Trichloroethylene, and systemic sclerosis, 545
Trichomonas vaginalis, 182
  diagnosis and treatment, 182
  testing for, 175
  urethral discharge, *172*
  urethritis, 176
  vaginal discharge, *172*
Trichophyton, 153
  T. tonsurans, 1233
  T. violaceum, 1233
Trichuris
  eosinophilia, 874
  T. trichiura, *163*, 167
Triclabendazole, trematode infections, *170*
Tricuspid atresia, *774*
Tricuspid regurgitation, 768
  heart sounds, 692
  organic, 768
  symptoms and signs, 768
  treatment, 768
Tricuspid stenosis, 767–9
  investigations, 769
  pathophysiology, 769
  signs, 769
  symptoms, 769
  treatment, 769
Tricuspid valve, 684
Tricyclic antidepressants (TCAs), 1201, 1202
  diabetic neuropathy, 1055
  narcolepsy/cataplexy, 1145
  overdose, *935*, 949–50
    clinical features, 949
    metabolic acidosis, *939*
    treatment, 949–50
  pain relief, *495*
  pancreatitis, 381
  side-effects, *928*, *1203*
    elderly patients, *924*
    long QT syndrome, *727*
    postural hypotension, *659*
Trientine, Wilson's disease, 357
Trifluoperazine, schizophrenia, 1217
Trigeminal nerve, 1106–7, *1106*
  lesion, 1106
  sensory input, *1106*
Trigeminal neuralgia, 1106–7
  features, 1106–7
  incidence, *1096*
  secondary, 1107
  treatment, 1107
Trigeminal neuropathy, idiopathic, 1107
Triglycerides (TGs), 321, 1061, 1062–3
  absorption, 273–4
  synthesis, 210
Triiodothyronine (T3), 967, *971*
  metabolism, 983, *983*
  synthesis, 983, *983*
Trilostane, Cushing's syndrome, 1014
Trimethoprim, *94*, 98
  acne vulgaris, 1246
  prostatitis, 603
  urinary tract infections, 602, 603
Triple X syndrome, *40*
Triplet repeat mutations, 43–4, *43*
Trismus, 1148
Trisomy 13, *40*
Trisomy 18, *40*
Trisomy 21, *30*, 38, *40*, *41*
Trochanteric bursitis, 512
Trochlear nerve, 1105
Tropheryma whipplei, 281
Tropical ataxic neuropathy, 1165
Tropical chronic pancreatitis, 380
Tropical countries *see* Developing/tropical countries, bacterial infections
Tropical pulmonary eosinophilia, 164, *875*, 876
Tropical splenomegaly syndrome, 155
Tropical sprue, 280
  aetiology, 280
  clinical features, 280
  diagnosis, 280
  treatment and prognosis, 280
Tropoelastin, 500
Tropomyosin, 682
Troponin, 683
  acute coronary syndromes, 752
Troponin C, 752
Troponin I, 684, 752, 788
Troponin T, 752, 788
Trousseau's sign, 1021, 1022
TRPC6, 583
Truncal ataxia, 1111, 1114
Trypanosoma, 85
  T. brucei, *82*
  T. brucei gambiense, 157
    see also African trypanosomiasis
  T. brucei rhodesiense, 157
    see also African trypanosomiasis
  T. cruzi, *82*, 158, *1179*
  see also Chagas' disease
Trypanosomal chancre, 157
Trypanosomiasis, 157–8
  African sleeping sickness *see* African trypanosomiasis
  mortality, *80*
  South American (Chagas' disease) *see* Chagas' disease
Trypsin, 273
Trypsinogen, 375, 380
Tryptophan, 210
Tryptophan malabsorption syndrome, 1071
Tsetse flies, *82*, 157
Tuberculides, 1230
Tuberculoid leprosy, 140
Tuberculoma, 866
Tuberculosis, 80, *83*, **863–8**
  adult post-primary, 865, *865*
  clinical features and investigations, 864, *864*
  developing/tropical countries, 145, *145*
  epidemiology, 863
  geographical distribution, *145*
  HIV-infected persons, 865–6
  lymph node presentation, 865
  management, 866–7
    chemoprophylaxis, 867
    DOTS, 866
    drug resistance, 866
    follow-up, 867
    longer regimes, 866
    six-month regimen, 866
    unwanted drug effects, 866–7
  miliary, 822, 864–5
  mortality, *80*
  multi-drug-resistant, 145
  pathology, 863–4
  prevention, 867–8
    BCG vaccination, 867
    contact tracing, 867–8
  pyrexia, *91*
  returning travellers, *90*
  rheumatoid factor, *503*
  skin manifestations, 1230
  small intestine, 282–3
    clinical features, 282
    diagnosis, 282–3
    treatment, 283
  spinal cord compression, 1164
  urinary tract, 603, *604*
Tuberculosis verrucosa cutis, 1230
Tuberculous arthritis, 540
Tuberculous meningitis, *1155*
  management, 1155–6
Tuberculous osteomyelitis, 541
Tuberculous pericarditis, 795
Tuberculous peritonitis, 318
Tuberous sclerosis (epiloia), 1170, 1253–4
Tubulointerstitial nephritis (TIN), **604–6**
  acute, 604–5
    drug-inducd, 604, *604*
    infection causing, 604, *604*
    in multisystem inflammatory diseases, 605
  chronic, 605–6
    causes, *605*
  granulomatous, *606*
Tubulointerstitial nephritis with uveitis (TINU) syndrome, 605
Tularaemia, 139

Tumour angiogenesis, 50
Tumour markers, 456, *456*, *493*
Tumour necrosis factor, 56, 57, 61, 66, 68, 214, 323, 655
  in diabetes, 1034
  haemopoiesis, 389
  osteoarthritis, 520
  response to shock, 895
  rheumatoid arthritis, 524
Tumour necrosis factor (TNF) blockers
  ankylosing spondylitis, 534
  juvenile idiopathic arthritis, 554
  psoriasis, 1242
  rheumatoid arthritis, *530*, 531
  side-effects, 531
Tumour suppressor genes, 49–50
  microsatellite instability, 50
  mode of action, 50
  viral inactivation, 50
  see also p53 gene
Tumours
  adrenal gland, 1014
  biliary tract, 373–5
  bladder, 645
  immunology, 451
  liver, 363–4, *364*
  Pancoast's, 811, 1177
  pancreatic, 384–5
  pelvic, 645
  renal, 644–5
  respiratory tract, 880–4
  skin, **1257**–**61**
  ureteric, 645
  see also Cancer; Malignancy/malignant disease; specific tumours
Tunga penetrans, 171
Tunica intima, 685
Tunica media, 685
Tunnel vision, 1101
Turcot syndrome, *301*, 302
Turgor, 659
Turner's syndrome, 39, *40*, 1002, *1007*
  coarctation of aorta, 777
  short stature, 978
Twin studies
  ankylosing spondylitis, 533
  diabetes, 1034
  inflammatory bowel disease, 286
  obesity, 228
  osteoarthritis, 520
  psoriasis, 1240
  systemic lupus erythematosus, 541–2
Tylosis, malignancy-related, *1255*
Tympanic membrane, 1077
  perforation, 1078
Tympanostomy tube, 1079
Type 1 errors, 933
Type 2 errors, 933
Type I hypersensitivity *see* Allergic disease (type I hypersensitivity)
Type II hypersensitivity *see* Autoimmunity (type II hypersensitivity)
Type III hypersensitivity *see* Immune complex disease (type III hypersensitivity)
Type IV hypersensitivity *see* Delayed hypersensitivity (type IV)
Type V hypersensitivity *see* Hypersensitivity reactions
Typhoid fever, *83*, 144
  immunization, *103*
  rash, *88*
  returning travellers, *90*
Typhus, *80*, *82*, *83*, 147–8, *147*
  African tick, *147*
  endemic (murine), 147, *147*
  epidemic, 147, *147*
  returning travellers, *90*
  scrub, 147–8, *147*
Tyrosinase deficiency, *1070*
Tyrosine, 210
Tyrosine kinase(s), 1031
Tyrosine kinase inhibitors
  renal cell carcinoma, 645
  see also individual drugs
Tyrosine kinase receptors, 22
Tyrosine phosphatase receptors, 22

Ubiquitin, 20, 44, 1168
Ubiquitin-proteasome pathway, 214
Ubiquitination, 576
UDP-glucuronosyl transferase, 329
  genetic polymorphism, *926*
Ulcer(s) (ulceration)
  aphthous *see* Aphthous ulcers/ulceration
  chancroid, 180
  Curling, 262
  decubitus *see* Pressure sores (decubitus ulcers)
  foot *see* Foot
  Hunner's, 602
  leg *see* Leg(s)
  mycobacterial (Buruli), 142
  oral *see* Oral ulceration
  skin, *1227*
  snail track, 178
  see also specific diseases/conditions
Ulcerative colitis, **290**–**4**
  clinical features, 290–1
  course and prognosis, 293
  examination, 291
  extragastrointestinal manifestations, *288*
  histology, *287*
  investigations, 291
    blood tests, 291
    colonoscopy, 291
    imaging, 291
    radionuclide scans, 291
    stool cultures, 291
  medical management, 291–2
    left-sided proctocolitis, 292
    proctitis, 292
    total colitis, 292
  microscopic, 294
  microscopic changes, 287
  mucosa in, 287
  severe attack, *291*
  splenic atrophy, 423
  surgical management, 292–3, *292*
Ulnar nerve compression, 1172
Ultrafiltration rate, 572
Ultrasonography *see* Ultrasound
Ultrasound
  abdominal, 244
    common bile duct stones, 370
    gallstones, 369, *369*
  coronary intravascular, 704
  Doppler *see* Doppler ultrasound
  endoanal, 244
  endorectal, 244
  endoscopic (EUS) *see* Endoscopic ultrasound (EUS)
  kidney, 579–80
  lungs, 824
  pancreas, 376
  specific conditions
    acute abdomen, 315
    autosomal-dominant polycystic kidney disease, *643*
    cancer, 457
    cholestatic jaundice, 331, *331*
    chronic renal failure, 628
    cirrhosis, 346
    Crohn's disease, 289
    gall bladder disease, *326*
    gastric cancer, 265
    goitre, 993
    liver disease, 325–6, *326*
    musculoskeletal problems, 505
    osteoporosis, 562
    pancreatic carcinoma, 383
    pancreatitis, 378
    PCOS, 1004, *1004*
    pulmonary embolism, 785
    pyogenic abscess, 361
    urinary tract infections, 602
    urinary tract obstruction, 615
    see also Echocardiography
Ultraviolet light
  PUVA, 1248
  side-effects on skin, 450, 1248
  see also Skin cancer
  UVA, 1248
  UVB, 1248
  see also Sunlight
Umbilical cord blood, 47
Unconscious patient, 1124–6
  assessment, 1124–5
  breath, 1124
  lateralizing signs, 1124–5
  respiration, 1124
  skin, skull, spine and injuries, 1124
  temperature, 1124
  see also Coma; Unconsciousness
Unconsciousness, 1122–6
  actions, *1124*
  disturbances
    syncope, 689, 717
      see also Syncope
  see also Coma; Unconscious patient
Uncontrolled observational studies, 931–2
Undernutrition, 207
  children, 215
  see also Malnutrition
Underweight *see* Weight, loss
Undulant fever *see* Brucellosis
Unguentum Merck, *1270*
Unithiol, *937*
Upper limbs
  chronic work-related pain syndrome, 516–17
  examination, *502*
  pyramidal drift, 1111
Upper motor neurone (UMN) lesions, 1107, 1111, *1111*, 1112
  loss of bladder control, 1119
Upper respiratory tract infection, returning travellers, *90*
Uraemia *see* Renal failure
Uraemic acidosis, 677
Uraemic neuropathy, 1174
Uraemic pericarditis, 794
Urate
  myelofibrosis, 421
  renal calculi, 613
Urea, 574
  myeloma, 485
  non-Hodgkin's lymphoma, 480
  osmotic activity, 649
  plasma/serum, *585*
    Addison's disease, 1011
    causes of alteration, *620*
    chronic renal failure, 627
    factors affecting, *574*
    infective endocarditis, 772
    pancreatitis, *378*
    systemic lupus erythematosus, 543
    systemic sclerosis, 546
  urinary, 210
  see also Electrolytes
Urea reduction ratio, 635
Ureagenesis, 674
Ureaplasma, *70*, *123*
  U. urealyticum
    non-specific urethritis, 535
    urethral discharge, *172*
    urethritis, 176
Ureter
  colic, 611–12
  tumours, 645
Ureterography
  antegrade, urinary tract obstruction, 615
  retrograde, urinary tract obstruction, 615
Urethra, bleeding from, 578
Urethral discharge, *172*
Urethral syndrome, 602
Urethritis, 176–7
  clinical features, 176
  diagnosis, 176–7
  recurrent/persistent, 177
  treatment, 177
Urethrography, urinary tract obstruction, 615
Urethroscopy, urinary tract obstruction, 615
Uric acid
  excretion, 537, *538*
  levels
    gout, *537*
    Hodgkin's lymphoma, 477
    non-Hodgkin's lymphoma, 480
  myeloma, 485
  renal calculi, 613
  stones *see* Uric acid stones
  synthesis, 537
Uric acid stones, *609*, 613
Uridine diphosphate galactase-4-epimerase, 1068
Uridine diphosphoglucuronosyl transferase, 323
Urinalysis
  chronic renal failure, 627
  neurology, *1122*
Urinary anion gap, 677
Urinary incontinence, *1099*
  elderly patients, 647
  functional, 647
  overflow, 614, 647
  stress, 647
  urge, 647
Urinary retention
Urinary tract
  abnormal, 600
  infections *see* Urinary tract infections (UTI)
Urinary tract infections (UTI), 135, *314*, **599**–**604**, 677
  acute pyelonephritis, 600, *601*
  aetiology and pathogenesis, 599–600, *600*
  bacterial prostatitis, 603
  bladder outflow obstruction, *see* Bladder outflow obstruction
  complicated vs uncomplicated, 600, *600*
  diagnosis, 602, *602*
  elderly patients, 647
  indwelling catheters, 603
  innate host defence, 600
  malakoplakia, 604
  natural history, 600–1, *600*, *601*
  pregnancy, 603
  reflux nephropathy, 600–1, *601*
  reinfection vs relapse, 601, *601*
  and renal calculi, 610–11
  renal carbuncle, 603
  returning travellers, *90*
  special investigations, 602
  symptoms and signs, 601–2
  treatment, 602–3
    recurrent infection, 602–3
    single isolated attack, 602
  tuberculous, 603, *604*
  virulence, 599
  xanthogranulomatous pyelonephritis, 604
Urinary tract obstruction, **614**–**17**
  aetiology, 614, *614*
  bladder outflow *see* Bladder outflow obstruction
  causes, *614*
    benign prostatic hypertrophy, 617
    calculi *see* Renal calculi
    obstructive megaureter, 616
    pelviureteric junction obstruction, 616, *616*
    retroperitoneal fibrosis, 616–17, *617*
  clinical features, 614–15

investigations, 615, *615*
pathophysiology, 614
prognosis, 617
renal scintigraphy, 581
signs, 614–15
treatment, 615–16
Urinary tract stones
formation, 677
uric acid *see* Uric acid stones
Urine
alkalinization, 677, 938
appearance, 578
bacteria in *see* Bacteriuria
biochemistry, chronic renal failure, 627
bloody *see* Haematuria
concentration, 573–4
electrolytes, 627
electrophoresis, 627
examination, 578–9
glycosuria, 579
immunofixation, 627
microalbuminuria, 579
microscopy, 579, *585*
chronic renal failure, 627
nitrogen excretion, 210
osmolality, 578, 620, 653
and host defence, 600
renal failure, 627
pH, 578
and host defence, 600
proteins in *see* Proteinuria
red cells in, 579, 627
sediments
sodium, 653
specific gravity, 578, 620
tests
liver disorders, 325
Stix, 578–9
toxicology screens, 1215–16, *1216*
volume, 578
Urine culture, *89*
Urine flow, and host defence, 600
Urine tests
glucose, 1044
urobilinogen, 325
Urobilinogen, 323
dipstick tests, 325
Urodilatin, 577
Uroepithelium, 600
Urography
antegrade/retrograde, *see* Antegrade urography
excretion, 580
magnetic resonance, 580
Uroporphyrins, 1073
Urothelial tumours, 645
Ursodeoxycholic acid
gallstone dissolution, 370
intrahepatic cholestasis of pregnancy, 362
primary biliary cirrhosis, 355
Urticaria, 72, **1243–4**
aetiology, 1243
cholinergic, 1243, 1253
clinical features, 1243, *1243*
cold, 1243
drug-induced, *1265*
food allergy, 237
investigations, 1243
papular, arthropod bites, 171
physical, 1243
plasma protein incompatibility, 427
pressure, 1243
prognosis, 1243
solar, 1249
Urticaria pigmentosa, 1265
Urticarial vasculitis, 1243–4
Usual interstitial pneumonia, 872–3, *872*
Uterine cancer, epidemiology, *449*
Uthoff's phenomenon, 1152
Uveitis, 1092–3
and ankylosing spondylitis, 533
anterior, 188, 1092–3, *1092*
and sarcoidosis, 869
symptoms, 1092
with tubulointerstitial nephritis, 605
Uveoparotid fever, 869

Vaccines
hepatitis A, 173, 335
herpes zoster, 1157
Hib, 834
human papillomavirus, 182
influenza, 834
see also Immunization; specific vaccines
Vaccinia virus, *104*, 109
Vacuolating toxin, 258–9
Vagina
Vaginal discharge, *172*
Vaginal examination, 314
Vaginosis, bacterial *see* Bacterial vaginosis
Vagus nerve, 1110
lesions, 1110
stimulation, 1203
Valaciclovir, *101*, 107
Bell's palsy, 1108
genital herpes, 181
herpes simplex, 1231
shingles, 1231
Valganciclovir, *101*
cytomegalovirus, 107, 202
Valgus deformity
Validation, 12
Valine, 210
Valproate *see* Sodium valproate
Valsalva manoeuvre, 722, *723*
Valsartan
heart failure, *739*
hypertension, 804, 1023
myocardial infarction, 759
Valvotomy
closed, 762–3
open, 763
trans-septal, 762
Valvular calcification, 693
Valvular heart disease, 742, **760–9**
calcific, 765
regurgitation, 703
stenosis, 703
see also Heart valves; and individual valves
Vanadium, *226*
Vancomycin, *94*, 99
gastroenteritis, *132*
infective endocarditis, *773*
therapeutic drug monitoring, *927*
Vanilloid receptor-1, 251
Vapours, 849
Vaptans, cirrhosis, 353
Vardenafil, erectile dysfunction, 1000, 1056
Varenicline, 830
Variant Creutzfeldt-Jakob disease (vCJD), 79, *81*, 122, 1158
proteolysis resistant proteins, *see* Prion protein (PrP)
tonsillar biopsy, 1121
transfusion-transmitted, 428
Variceal banding, 350
Variceal haemorrhage, management
acute episodes, 350–1
acute rebleed, 351
balloon tamponade, 350
endoscopy, 350, *350*
injection sclerotherapy, 350
resuscitation, 350
variceal banding, 350
vasoconstrictors, 350
see also Somatostatin; Terlipressin
Varicella (chickenpox), *105*, 106, *106*
clinical features, 106, *106*
prophylaxis and treatment, 106–7
Varicella zoster virus, *104*, *105*, *191*, 1231
HIV-related, 203
immunization, *102*, *103*
Varicose veins, 809
Variegate porphyria, 1074, *1074*
Variola *see* Smallpox
Variola virus, *104*
Vasa recta, 571, 573
Vascular cell adhesion molecule *see* VCAM-1
Vascular compliance, reduced, 631
Vascular disease/disorders, 435, *435*
gene therapy, 47
see also Vasculitis
Vascular endothelial growth factor *see* VEGF
Vascular endothelium
functions of, 685–7
see also Endothelium
Vasculitis, 1263
ANCA-positive, 596–7, *596*
diagnosis, *69*
drug-induced, *1265*
leucocytoclastic, 1263
rheumatoid arthritis, 528
systemic inflammatory *see* Systemic inflammatory vasculitis
see also Polyarteritis nodosa (PAN); specific disorders
Vaso-occlusion
Vasoactive intestinal peptide, *272*, *375*, 384, 814, 970
asthma, 852
VIPoma, 385
Vasoactive mediators, endothelium-derived, 896–7
Vasodilator washout effect, 686
Vasodilators
ACE inhibitors *see* ACE inhibitors
heart failure, 739, *739*
shock, 908–9
see also Beta(β)-blockers; Nitric oxide (NO)
Vasomotor control, 686–7
Vasomotor rhinitis, 832
Vasoplegia, 897
Vasopressin *see* Antidiuretic hormone (ADH; Vasopressin)
Vasopressin receptor antagonists, autosomal-dominant polycystic kidney disease, 643
Vasopressor agents, heart failure, *739*, 740
Vasovagal attack, 689
VCAM(s), *686*
VCAM-1, *68*, 409, 545
rheumatoid arthritis, 524
systemic sclerosis, 545
Vector-borne disease, 82, *82*
Vectors, 35
Vegetative state, *1123*
differential diagnosis, *1125*
VEGF, 50, 451, 452, 686, *686*
osteoarthritis, 520
pancreatic neuroendocrine tumours, 384
placental development, 597–8
VEGF-receptor, 598
Veins *see also* entries beginning venous
Venesection
haemochromatosis, 357
polycythaemia vera, 420
Venezuelan equine virus, 1156
Venlafaxine, 1202
genetic polymorphism in metabolism, *926*
Veno-occlusive disease, 360
Venodilators *see* Vasodilators
Venography
Venomous animals, 951–2
Venous eczema, 1239
Venous outflow occlusion, 283
Venous return *see* Preload
Venous thromboembolism
established, 447
prophylaxis, 446–7
treatment, 444–6
see also Thrombolytic therapy
Venous thrombosis, 442–3, 809–10
coagulation abnormalities, 442–3
deep vein *see* Deep venous (vein) thrombosis (DVT)
investigations, 443
renal vein, 609
risk factors, *442*
thrombophilia, 442
Venous ulcers, 1261, *1261*
Ventilation
control of, *815*
extracorporeal gas exchange, 916
hyperventilation, 688, 817, 820
'lung-protective', 915–16
non-invasive, 916, *916*
poisoning, 934, 935
ventilatory forces, *816*
see also Mechanical ventilation
Ventilation/perfusion mismatch, 817, 893
COPD, 836
Ventilation/perfusion relationships, 817, *817*
Ventilator-associated lung injury, 914
Ventilatory failure, acute, 912–13
Ventricles (heart)
depolarization, 684
isovolumetric contraction, 684
isovolumetric relaxation, 684
myocyte action potentials, 682, *682*
performance *see* Frank-Starling relationship
pressure-volume loop, 684, *685*
see also Left ventricle (ventricular); Right ventricle (ventricular)
Ventricular arrhythmias *see* Ventricular tachyarrhythmias
Ventricular assist devices, 743, *743*
Ventricular dysfunction
Ventricular ectopic beats, *729*
Ventricular end-diastolic volume, 891
Ventricular end-systolic volume, 891
Ventricular fibrillation, 711, 726–7, *727*
post-STEMI, 759
Ventricular function curve, *684*
Ventricular premature beats, 729, *729*
Ventricular septal defects (VSD), 774–5
auscultation, *775*
clinical features, 774–5, *775*
investigations, 775
pathophysiology, *775*
post-STEMI, 759
right bundle branch block, *720*
treatment, 775
Ventricular tachyarrhythmias, 726–9
Brugada syndrome, 727, *727*
life-threatening, 726
fibrillation *see* Ventricular fibrillation
management, *730*
long QT syndrome, 727–8, *728*, *1078*
management, 729–33
short QT syndrome, 728
tachycardia *see* Ventricular tachycardia

Ventricular tachycardia, 711
  electrocardiogram, 726, *727*
  non-sustained, 729
    treatment, *730*
  normal, 728–9
    treatment, *730*
  post-STEMI, 759
  pulseless, 711
  sustained, 726
Venturi mask, *841*
Verapamil, 731, *731*
  angina pectoris, 750
  atrial fibrillation, 725
  genetic polymorphism in metabolism, *926*
  hypertrophic cardiomyopathy, 790
  side-effects, hepatotoxicity, *366*
  supraventricular tachycardia, 723
Verbal memory, 1188
Veretebral artery dissection, pathophysiology, 1126
Vermis, 1114
Verner-Morrison syndrome, 306, 385
Verocytotoxin, 598
Verrucae, 1231
Verruga peruana, 148
Vertebrae
  discs *see various types*
  fractures *see* Vertebral fractures
Vertebral artery
Vertebral fractures, 486, 512, 561, 562, 564
Vertigo, 1079–80, 1097, 1108, *1128*, 1145
  benign paroxysmal positional, 1079–80
  causes, *1108*
  labyrinthine/central causes, 1080
  Ménière's disease, 1080
  and seizures, 1109
Very low-density lipoprotein (VLDL), 321, 1061, *1061*
  disorders of, 1063
  nephrotic syndrome, 584
Very low-density lipoprotein (VLDL) particles *see* Triglycerides
Vesical calculi *see* Renal calculi
Vesicles, *1227*
Vestibular nerve, 1108
  investigations, 1109
  lesions, 1109
Vestibular neuritis, 1109
Vestibulo-cochlear nerve, 1108–9
Vestibulo-ocular reflexes, 1124
Vibrio
  gastroenteritis, 130
  V. cholerae, 85, *123*, *128*, 143
    antibiotic resistance, *94*
  V. parahaemolyticus, 130
    food poisoning, *133*
  V. vulnificans, skin and soft tissue infections, *124*
Video capsule endoscopy, 246
  GI bleeding, 270
Video-assisted thoracoscopic lung biopsy, 829
Vigabatrin
Vildagliptin, 1039
Villi, 270
  atrophy *see* Villous atrophy
Villous atrophy, 39, 161, 189, *276*, 278, *279*, 280, 1256
Vimentin, *25*
Vinblastine, 24, 460
  Hodgkin's lymphoma, *461*
Vinca alkaloids, 460
  side-effects, neurotoxicity, 462
  thrombocytic purpura, 436
Vincent's angina, 247
Vincristine, 460
  acute lymphoblastic leukaemia, 471
  Burkitt's lymphoma, 483
  diffuse large B cell lymphoma, 482
  non-Hodgkin's lymphoma, *461*
  side-effects
    neuropathy, *1174*
    neurotoxicity, 462
Vinorelbine, 460
  breast cancer, 489
Vinyl chloride, and systemic sclerosis, 545
VIPoma, 385
Viral exanthem, 1230
Viral haemorrhagic fever, *83*
  as biological agent, 960
Viral infection(s), **104–21**
  arthritis in, 540–1
  DNA viruses, **104–9**
  HIV-associated, 201–3
  outcomes, 104
  RNA viruses, **109–21**, *110*
  skin, 1230–2
  transfusion-transmitted, 428
  see also specific infections/diseases
Viral markers, 325
Viral pericarditis, 794
Viral pneumonia, 859–60
Viral stimulation, 49
Viral warts *see* Warts
Virchow's node, 265
Viridans streptococci, *123*
Virilization, *995*, 1003
Virilizing tumour, *1007*
Viruses, 80
  DNA *see* DNA viruses
  leukaemia association, 80
  RNA *see* RNA viruses
VISA, 124
Visceral larva migrans, 168
Visceral leishmaniasis (kala azar) *see* Leishmaniasis, visceral (kala azar)
Visilizumab, Crohn's disease, *290*
Vision
  loss *see* Visual impairment/loss
Visual acuity, 1101
Visual cortex, infarction, 1130
Visual fields
  acromegaly, 979
  defects, *1099*, 1101–2, *1102*
  hyperprolactinaemia, 981
  loss, *1101*
  pituitary tumour, 972
Visual impairment/loss, 1094, *1094*
  xerophthalmia *see* Xerophthalmia
  see also Blindness
Visual system, 1100–4, *1101*
Visuospatial defects, *1099*
Vital capacity, respiratory failure, 911
Vitamin(s), **217–25**
  absorption, 274
  deficiency, 275–6, 1174–5, *1174*, *1175*
    causes, *218*
    see also under specific vitamins
  fat-soluble, 218–20, *218*
  parenteral nutrition, 236
  recommended IV intake, *234*
  storage in liver, 321
  water-soluble, *218*, 220–5
  see also individual vitamins
Vitamin A, **218–19**
  deficiency, 218–19, *218*
    diagnosis, 219
    malnourished children, 219
    prevention, 219
    treatment, 219
  function, 218
  LRNI, *218*
  overdose, *1102*
  recommended IV intake, *234*
  RNI, *218*
  side-effects, 219
    hepatotoxicity, *366*
  teratogenesis, 219
Vitamin B_1 *see* Thiamin (vitamin B_1)
Vitamin B_2 *see* Riboflavin (vitamin B_2)
Vitamin B_6 (pyridoxine), 223
  deficiency, *218*, 1175
  LRNI, *218*
  recommended IV intake, *234*
  RNI, *218*
  tuberculosis, 866
Vitamin B_{12} (cobalamin), 223, 398–9
  abnormal utilization, *399*
  absorption, 399
  absorption tests, 400
  deficiency, *218*, 281, 397, 399, 1165, 1175
    treatment, 401
    see also Pernicious anaemia
  dietary intake, 398
  function, 398–9
  LRNI, *218*
  pernicious anaemia, 400
  pyogenic abscess, 361
  recommended IV intake, *234*
  RNI, *218*
  structure, 398–9, *398*
  synthesis, 398
  transport, 399
  see also Folates/folic acid
Vitamin C, 223
  deficiency *see* Scurvy
  LRNI, *218*
  recommended IV intake, *234*
  RNI, *218*
  side-effects, renal calculi, 611
Vitamin D
  deficiency, 275, 558, 559, 567, 670, 1019, 1021, 1022
    see also Rickets
  metabolism, 559, *559*, 577
  osteoporosis prevention, 563
  recommended IV intake, *234*
  RNI, *218*
  side-effects
    endocrine disease, *964*
    renal calculi, 611
    uraemia, 619
  see also 1, 25-Dihydroxycholecalciferol (1, 25(OH)2D3); 25-Hydroxy vitamin D (25-hydroxycholecalciferol); Vitamin D-resistant rickets
Vitamin D-dependent rickets
  type I, 670
  type II, 670
Vitamin E, 220
  deficiency, *218*, 220
  epidemiological data, 220
  function, 220
  RNI, *218*
  supplementation, 220
  see also Antioxidants
Vitamin K, 219–20, *937*
  deficiency, *218*, 219–20, 440
    activated partial thromboplastin time, *439*
    blood changes, *439*
    causes, 440
    cholestatic jaundice, 229
    concomitant vitamin K antagonists, 229
    newborns, 219–20
    prothrombin time, 324, *439*
  functions, *see* Coagulation
  haemolysis in G6PD deficiency, *414*
  RNI, *218*
Vitamin K antagonists, 229
Vitaxin, 50
Vitiligo, *965*, 1264, *1264*
  autoantigens, *74*
  treatment, 1264
Vocal cords, 1082
  nodules, 1083, *1083*
Vocal resonance, 821
Vogt-Koyanagi-Harada syndrome, 1159
Volkmann's canals, 557
Voltage-gated ion channels
Volume receptors, 654
  extrarenal, 652
  intrarenal, 651
  see also Osmoreceptors
Volume replacement, 904–5
  blood, 905–6
  blood substitutes, 906
  crystalloids and colloids, 906
  red cell concentrates, 906
Vomiting, 242, *1128*
  acute abdomen, 314
  causes of, *242*
  chemotherapy-induced, 461
  faeculent, 242
  functional, 311
  gastric outlet obstruction, 262
  haematemesis, 241, 242
  opiate therapy, 495, 496
  in pregnancy, 242
  projectile, 242
  protracted, 241
  see also Nausea
Von Gierke's disease, *1069*
Von Hippel-Lindau protein, 576
Von Hippel-Lindau syndrome, *453*, 1170
Von Recklinghausen's disease *see* Neurofibromatosis
Von Willebrand disease, 440
  blood changes, *439*
  clinical features, 440
  treatment, 440
Von Willebrand factor, 429, 432, 433, *686*
  ADAMTS 13, 437
  deficiency *see* Von Willebrand disease
  haemophilia A, *439*
Von Willebrand factor gene, 440
Voriconazole, 100, *100*
  chromoblastomycosis, 153
VRSA, 124
Vulnerable patients, respect for autonomy, 5–6

Waddling gait, *1096*, 1097
Waist circumference, 1034
Waist/hip ratio, 746
Waiting lists, 8
Waldenström's macroglobulinaemia, 481, 1072
Wall motion stress echocardiography, 702
Wallenberg's syndrome, 1129
Wallerian degeneration, 1171
Wangella, 153
Warfarin
  arterial thrombosis, *443*
  atrial fibrillation, 725
  decreased effect, 446
  deep vein thrombosis, 810
  genetic polymorphism in metabolism, *926*
  IgA nephropathy, 594
  increased effect, 446
  metabolism, 926
  overdose, antidotes, *937*
  pulmonary hypertension, 783
  side-effects, bleeding, 446
  teratogenesis, *924*
Warts
  anogenital, 181–2, 1231
  common, 1231, *1231*
  filiform, 1231
  plane, 1231
  plantar, 1231
  seborrhoeic, 1257, *1257*
Wasp venom allergy, 72, 73
Wasting *see* Cachexia
Water
  absorption, 274
  adult requirement, 234
  composition, **649–61**
  distribution, 650, *650*
  double-labelled technique, 208

excretion
regulation of, 653–4
fluoridation, 226
input, *295*
loss of, 660, 661
from gut, *662*
output, *295*
retention, 351
heart failure, 734
ventilation-related, 914
see also Ascites; Oedema
Water balance, 207
Water blisters *see* Molluscum contagiosum
Water hammer pulse, 690
Water-deprivation test, *1017*
Watershed (borderzone) infarction, 1130
Weakness *see* Muscle weakness
Weals, *1227*
Weber test, 1078
Weber's syndrome, 1130
Wegener's granulomatosis, *549*, 596, 870
ANCA antibodies in, 596
autoantigens, *74*
clinical features, *550*
diagnosis, *550*
gingival swelling, 248
midline granuloma, 870
pyrexia, *91*
Weibel-Palade bodies, 27, 687
Weight, 209
and amenorrhoea, 100
and coronary artery disease, 746
gain, 208
excessive *see* Obesity
and hirsutism, 1004
loss, 208, 214, 241, 243, 275
diabetes mellitus, 1036
gastric cancer, 265
in obesity, *230*, 232, 1065
see also Anorexia; Diet; Obesity
Wellcome Trust Case Control Consortium, 36
Wenckebach (Mobitz type I) block, *718*
Werdnig-Hoffman disease, 1167
Werner's syndrome, 226
Wernicke-Korsakoff syndrome, 221, 238, 1174–5
prevention, 359
Wernicke's aphasia, 1099, *1099*
West Nile virus, 79, *82*, *110*, 116, 1156
transfusion-transmitted, 428
Wheat sensitive enteropathy *see* Coeliac disease
Wheezing, 820, 821
Whiplash injury, 506–7
Whipple's disease, 281–2, 1159
rheumatological problems, 555
Whipple's triad, 1060
Whipworm *see Trichuris*, *T. trichiura*
White blood cells *see* Leucocytes
White cell count (WCC), 389, *389*, 393
musculoskeletal problems, 503
White-cell casts, 579
Whiteheads, 1246
Whole blood, 429
Whole bowel irrigation, 938
Whooping cough *see* Pertussis (whooping cough)
Wickham's striae, 1244
Widow's stoop, 512
Wilms' tumour *see* Nephroblastoma (Wilms' tumour)
Wilson's disease (hepatolenticular degeneration), 225, 357–8, 1148
aetiology, 357
children, 1148
clinical features, 357
investigations, 357
pathology, 357
prognosis, 357–8
treatment, 357
Winemaker's lung, *876*
Winter eczema, 1239
Wireless capsule endoscopy, coeliac disease, 279
Wiskott-Aldrich syndrome, 70
Withdrawing treatment, 919–20
Withholding treatment, 919–20
WNK proteins, 667
wnt, 558
Wnt co-receptor, 484
Wolff-Parkinson-White syndrome, 721, *721*, *722*
atrial fibrillation, *722*
management, *730*
Wolfram syndrome, *1035*
Women *see* Females
Wood dust allergy, *848*
Wood's light, *1228*
Woolsorter's disease, 142
World Health Organization (WHO) classification schemes
Wrist pain, 508–9
causes, *508*
ganglion, 509
inflammatory arthritis, 509
osteoarthritis, 509
rheumatoid arthritis, 526
scaphoid fracture, 509
Writer's cramp, 1150
WT1, **33**
Wuchereria bancrofti, *82*, *163*, *164*, 876

X-chromosome, 37
X-linked agammaglobulinaemia, 71
treatment, 71
X-linked dominant hypophosphataemic rickets, 670
X-linked genetic disorders
dominant, 43
recessive, 43
see also specific diseases/disorders
X-ray
abdominal *see* Abdominal X-ray
bone, 555, 560
chest *see* Chest X-ray
kidney, 579, *580*
sinus, 843
skeletal/bone abnormalities
skull, 1119
specific conditions
ankylosing spondylitis, 533–4, *533*
cancer, 457
musculoskeletal problems, 504–5
osteoarthritis, 521
Paget's disease, 565, *566*
poisoning, 939
rheumatoid arthritis, 528
rickets/osteomalacia, 567
spinal, 1119
thoracic inlet, 993
Xanthelasma, 1064, 1254
primary biliary cirrhosis, 355
Xanthogranulomatous pyelonephritis, 604
Xanthoma, 329, 1065, *1065*, 1254
Xeroderma pigmentosa, 453, *453*
Xerophthalmia, 218, 1094
classification, *219*
Xerostomia, 248
Xylometazoline, 1081
rhinitis, 833

Y-linked disorders, 43
Yaws, 142
Yeast artificial chromosomes, 35
Yellow fever, *81*, *82*, *83*, *110*, 114
clinical features, 114
diagnosis and treatment, 114
haemorrhagic manifestations, *114*
hepatitis, 342
immunization, *103*
prevention and control, 114
Yellow-nail syndrome, 1266
Yersinia
Y. enterocolitica, 130
food poisoning, *133*
reactive arthritis, 535
transfusion-transmitted, 428
Y. pestis, *80*, *82*, *123*, 146
see also Plague
Yersiniosis, 130

Z bands, 682
Zafirlukast
asthma, *854*, 856
rhinitis, 833
Zaleplon, insomnia, 1198
Zanamivir, 101–2, *101*
influenza, 834
Zebra bodies, 597
Zellweger's syndrome, 23
Zenker's diverticulum, 254
Zidovudine, *192*, *195*
in pregnancy, 198
side-effects
lipoatrophy, 198
myelotoxicity, 189
rash, 1269
Ziehl-Neelsen stain, 603, *604*
Zieve's syndrome, 359
Zileuton, asthma, *854*
Zinc, 225
absorption deficiency *see* Acrodermatitis enteropathica
deficiency, 225
recommended IV intake, 233
RNI, 225
toxicity, 225
Zinc acetate, Wilson's disease, 357
Zinc finger, *32*, **33**
Zoledronate
breast cancer, 489
hypercalcaemia, 458
osteoporosis, 563, *564*
prostate cancer, 492
Zollinger-Ellison syndrome, 260, 385, 679
and malabsorption, 282
Zolmitriptan, migraine, 1137
Zolpidem, insomnia, 1198
Zona occludens toxin, *306*
Zonula adherens, *25*, 26
Zonula occludens, *25*, 26
Zonula occludens-1, 583
Zoonoses, 81, *81*
Zopiclone, insomnia, 1198
Zoster immunoglobulin, 106
Zygomycosis, *150*
invasive, 152
subcutaneous, *150*, 152–3